НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2025 Том 25 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2025 Vol. 25 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

ОПТИЧЕСКИЕ СИСТЕМЫ И ТЕХНОЛОГИИ

OPTICAL ENGINEERING

doi: 10.17586/2226-1494-2025-25-6-1003-1013
УДК 532.211; 544.032.65

Абляция объемного халькогенидного стекла As_2Se_3 при воздействии фемтосекундными лазерными импульсами с высокой частотой следования

Александра Андреевна Шамова[1]✉, Дмитрий Сергеевич Поляков[2], Артур Джуракулович Долгополов[3], Дмитрий Васильевич Панькин[4], Галина Дмитриевна Шандыбина[5]

[1,2,3,5] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
[4] Санкт-Петербургский государственный университет, Научный парк, Ресурсный центр «Оптические и лазерные методы исследования вещества», Санкт-Петербург, Петергоф, 198504, Российская Федерация

[1] aashamova@itmo.ru✉, https://orcid.org/0000-0001-8728-260X
[2] polyakovdmitry1988@gmail.com, https://orcid.org/0000-0002-6240-4057
[3] addolgopolov@itmo.ru, https://orcid.org/0000-0002-9548-791X
[4] dmitrii.pankin@spbu.ru, https://orcid.org/0000-0003-3114-3410
[5] corchand@gmail.com, https://orcid.org/0000-0003-3501-1925

Аннотация
Введение. Фемтосекундная лазерная обработка халькогенидных стекол является перспективным методом высокоточной модификации их структуры и свойств для разработки оптических элементов инфракрасной фотоники. Одной из актуальных задач остается повышение скорости обработки стекла при обеспечении высокой пространственной точности и минимальных термических повреждений. При использовании режимов лазерного облучения с высокой частотой следования импульсов, обеспечивающих повышение производительности технологий, изменяются механизмы фазово-химических превращений и возрастает вклад накопительного нагрева. Однако динамика этих процессов в объемном материале остается недостаточно изученной. В работе предложено определение механизма изменения фазового и химического составов объемного халькогенидного стекла из селенида мышьяка (As_2Se_3) под действием серии фемтосекундных лазерных импульсов в режимах интенсивной абляции. **Метод.** В качестве объекта исследования используются пластины халькогенидного стекла As_2Se_3, подвергнутые воздействию фемтосекудных лазерных импульсов с длиной волны 515 нм и частотой следования вплоть до 1 МГц. Диагностика образцов проводится методами цифровой оптической микроскопии и рамановской спектроскопии. Теоретический анализ включает в себя как расчеты фотовозбуждения и нагрева полупроводника одиночным лазерным импульсом, так и расчеты накопительного нагрева поверхности образца с учетом трехмерного теплоотвода. **Основные результаты.** При частоте следования лазерных импульсов 1 кГц установлен порог одноимпульсной лазерной абляции и определены параметры степенной зависимости порога абляции от количества импульсов («инкубационный эффект»). Проведен детальный анализ морфологии облученных образцов и химического состава продуктов лазерной абляции, в ходе которого установлено образование аморфного селена (*a*-Se) и кристаллов арсенолита (As_2O_3). Теоретический анализ позволил оценить степени нагрева и фотовозбуждения халькогенидного стекла As_2Se_3 одиночным лазерным импульсом и выявил существенный вклад эффекта накопления тепла в нагрев поверхности при частоте следования импульсов 1 МГц. На основе совокупности полученных результатов установлен парофазный механизм преобразования фазового и химического состава в объемном халькогенидном стекле As_2Se_3 при лазерной абляции с частотой следования фемтосекундных импульсов 1 МГц. **Обсуждение.** Полученные результаты открывают перспективы для разработки высокопроизводительных технологий фемтосекундного лазерного микроструктурирования халькогенидных материалов в фотонике и сенсорике.

Ключевые слова
объемное халькогенидное стекло, селенид мышьяка, фемтосекундная лазерная абляция, рамановская спектроскопия

© Шамова А.А., Поляков Д.С., Долгополов А.Д., Панькин Д.В., Шандыбина Г.Д., 2025

Благодарности
Работа выполнена в рамках государственного задания (проект № FSER-2025-0007).
Авторы выражают благодарность Ресурсному центру «Оптические и лазерные методы исследования вещества» (проект № 125021902439-8) Научного парка Санкт-Петербургского государственного университета.
Авторы также благодарны сотруднику Института лазерных технологий Университета ИТМО М.М. Сергееву за помощь в проведении экспериментов.

Ссылка для цитирования: Шамова А.А., Поляков Д.С., Долгополов А.Д., Панькин Д.В., Шандыбина Г.Д. Абляция объемного халькогенидного стекла As_2Se_3 при воздействии фемтосекундными лазерными импульсами с высокой частотой следования // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1003–1013. doi: 10.17586/2226-1494-2025-25-6-1003-1013

Ablation of bulk chalcogenide glass As_2Se_3 by high-repetition-rate femtosecond laser pulses

Alexandra A. Shamova[1]✉, Dmitry S. Polyakov[2], Arthur D. Dolgopolov[3], Dmitrii V. Pankin[4], Galina D. Shandybina[5]

[1,2,3,5] ITMO University, Saint Petersburg, 197101, Russian Federation
[4] St. Petersburg State University, Research Park, Center for Optical and Laser Materials Research, Saint Petersburg, Peterhof, 198504, Russian Federation

[1] aashamova@itmo.ru✉, https://orcid.org/0000-0001-8728-260X
[2] polyakovdmitry1988@gmail.com, https://orcid.org/0000-0002-6240-4057
[3] addolgopolov@itmo.ru, https://orcid.org/0000-0002-9548-791X
[4] dmitrii.pankin@spbu.ru, https://orcid.org/0000-0003-3114-3410
[5] corchand@gmail.com, https://orcid.org/0000-0003-3501-1925

Abstract
Femtosecond laser processing of chalcogenide glasses is a promising method for high-precision modification of their structure and properties for the development of optical elements for infrared photonics. One of the key challenges is to increase the processing speed while maintaining high spatial accuracy and minimal thermal damage. When using laser irradiation modes with a high pulse repetition rate, which ensures increased productivity of technologies, the mechanisms of phase-chemical transformations change and the contribution of accumulative heating increases. However, the dynamics of these processes in bulk material remains insufficiently studied. This paper studies the mechanism of phase and chemical composition transformation of As_2Se_3 bulk chalcogenide glass under the action of femtosecond laser pulses in the intense ablation modes. The objects of study are plates of chalcogenide glass As_2Se_3 irradiated by femtosecond laser pulses with a wavelength of 515 nm at repetition rates up to 1 MHz. The irradiated samples are analyzed using digital optical microscopy and Raman spectroscopy. Theoretical analysis includes both calculations of photoexcitation and heating of the semiconductor by a single laser pulse as well as calculations of accumulative heating of the sample surface, taking into account three-dimensional heat removal. The single-pulse laser ablation threshold was established at a laser pulse repetition rate of 1 kHz and the parameters of the power-law dependence of the ablation threshold on the number of pulses (incubation effect) were determined. A detailed analysis of the morphology of the irradiated samples and the chemical composition of the laser ablation products was carried out revealing the formation of amorphous selenium (*a*-Se) and arsenolite crystals (As_2O_3). Theoretical analysis allowed us to estimate the degree of heating and photoexcitation of As_2Se_3 chalcogenide glass by a single laser pulse and revealed a significant contribution of the heat accumulation effect to the surface temperature rise at a pulse repetition rate of 1 MHz. Based on the combined experimental and theoretical results obtained, a vapor-phase mechanism of phase and chemical transformation in As_2Se_3 bulk glass has been established during femtosecond laser ablation with a pulse repetition rate of 1 MHz. These findings open up prospects for the development of high-performance technologies for femtosecond laser microstructuring of chalcogenide materials in photonics and sensing applications.

Keywords
bulk chalcogenide glass, arsenic selenide, femtosecond laser ablation, Raman spectroscopy

Acknowledgements
The authors acknowledge the financial support of the Ministry of Science and Higher Education of the Russian Federation (No. FSER-2025-0007).
The authors acknowledge the Center for Optical and Laser Materials Research (Project No. 125021902439-8) of Research Park of Saint Petersburg State University.
The authors are also grateful to M.M. Sergeev, a staff member of the Institute of Laser Technologies at ITMO University, for assistance with the experiments.

For citation: Shamova A.A., Polyakov D.S., Dolgopolov A.D., Pankin D.V., Shandybina G.D. Ablation of bulk chalcogenide glass As_2Se_3 by high-repetition-rate femtosecond laser pulses. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1003–1013 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1003-1013

Введение

Халькогенидные стекла, благодаря возможности реализации в них сверхбыстрых фазовых переходов и высокому показателю преломления, являются превосходным материалом для фотонных приложений, могут служить основой для получения новых материалов с уникальными оптическими функциями и широким спектральным диапазоном прозрачности — от ультрафиолетовой до длинноволновой инфракрасной областей [1]. Формирование таких свойств возможно за счет фемтосекундной лазерной обработки, обеспечивающей высокоточную модификацию материала на микро- и наноуровнях без значительного теплового воздействия.

В последние годы получен значительный экспериментальный материал по лазерной обработке халькогенидных стекол. Например, осаждение из паровой фазы многокомпонентных халькогенидных тонкопленочных стекол с последующей термической обработкой позволяет реализовать селективную кристаллизацию фаз и обеспечивает высокую фоточувствительность в широком спектральном диапазоне [2]. Оптические компоненты, состоящие из поверхностных структур микронного размера, полученных методом пространственно контролируемой прямой лазерной записи с последующим селективным травлением, позволяют управлять различными параметрами оптического волнового фронта — фазой, амплитудой, поляризацией и дисперсией [3]. Кроме того, массивы наноотверстий, сформированные на большой площади поверхности халькогенидного стекла из селенида мышьяка (As_2Se_3), демонстрируют потенциал для реализации структурного цвета [4]. Несмотря на достигнутый прогресс, ключевыми задачами остаются повышение производительности технологий фемтосекундной лазерной обработки и обеспечение устойчивости халькогенидных стекол к воздействию интенсивного лазерного излучения. Известно, что для металлов и полупроводников при низких частотах следования фемтосекундных лазерных импульсов определяющее значение отводится накоплению структурных изменений поверхности от импульса к импульсу [5]. При переходе к частотам следования лазерных импульсов порядка 10^5–10^6 Гц, представляющих технологический интерес, важную роль начинает играть эффект накопления тепла [6]. Механизмы фазовых превращений в халькогенидном стекле As_2Se_3 при воздействии серии фемтосекундных лазерных импульсов в режимах абляции на сегодняшний день остаются предметом активных исследований.

Целью работы является определение механизма изменения фазового и химического состава объемного халькогенидного стекла As_2Se_3 под действием серии фемтосекундных лазерных импульсов в режимах интенсивной абляции. Для этого проведен комплексный анализ, основанный на экспериментальных данных, полученных при лазерном облучении образцов в видимом диапазоне спектра при различных частотах следования импульсов, на результатах диагностики облученных образцов методами цифровой оптической микроскопии и рамановской спектроскопии, а также на теоретических оценках температуры, достигаемой при воздействии одиночного импульса, «инкубационного эффекта» и накопительного нагрева поверхности халькогенидного стекла.

Методы и материалы

В качестве образцов использовались пластины неполированного халькогенидного стекла As_2Se_3 толщиной около 3 мм, полученные методом ультразвуковой резки из массивного стекла, изготовленного прямым сплавлением материалов в соответствии со стехиометрическим составом в кварцевых ампулах. Обработка образцов проводилась в воздушной атмосфере при нормальном падении линейно поляризованного излучения фемтосекундного волоконного лазера ANTAUS-20W-20u/1M (ООО «АвестаПроект», Россия) на длине волны излучения $\lambda = 515$ нм с длительностью импульса $\tau = 224$ фс, энергией импульсов до 20 мкДж и частотой их следования f до 1 МГц. Излучение направлялось на поверхность образца линзой с фокусным расстоянием 45 мм с диаметром пятна излучения $d_0 \approx 30$ мкм (по уровню e^{-1}). Диаметр пятна рассчитан по известной методике [7]. Измерение средней мощности/энергии лазерного излучения проводилось с помощью измерителя Gentec-EO SOLO2 (Gentec Electro-Optics, Inc., Канада). Оптические изображения образцов были получены с помощью цифрового оптического микроскопа Axio Imager A1m (Carl Zeiss Microscopy GmbH, Германия).

Структурные исследования химического состава были проведены с использованием метода спектроскопии комбинационного рассеяния света (КРС), также известной как рамановская спектроскопия, в геометрии обратного рассеяния света на экспресс-рамановском спектрометре SENTERRA (Bruker, Германия). Возбуждение КРС осуществлялось с использованием полупроводникового лазера с длиной волны 785 нм. Мощность лазерного излучения под $10^{\times}$ объективом с числовой апертурой 0,25 составляла 0,3 мВт, диаметр лазерного пятна в фокальной плоскости — примерно 10 мкм (по уровню e^{-2}). Спектры регистрировались в диапазоне 80–1545 см$^{-1}$ с использованием дифракционной решетки 1200 шт/мм. В экспресс-рамановском спектрометре SENTERRA размеры апертуры составляли 25 × 1000 мкм. Время накопления сигнала для съемки подбиралось индивидуально и варьировалось от 50 до 100 с. Каждое измерение повторялось дважды. Калибровка спектрометра перед измерениями и в ходе эксперимента проверялась по линии КРС кремния при 520,7 см$^{-1}$.

Аналитические расчеты накопительного лазерного нагрева выполнялись в программе Mathcad Prime 7.0, а численное моделирование процессов фотовозбуждения и нагрева объемного халькогенидного стекла As_2Se_3 одиночным фемтосекундным лазерным импульсом проводилось в программе, разработанной в рамках настоящей работы на языке программирования C (Си), которая не имеет аналогов в готовых математических пакетах.

Экспериментальные результаты

При фиксированной частоте следования фемтосекундных лазерных импульсов на поверхности халькогенидного стекла As_2Se_3 создается массив кратеров, в котором по вертикали изменяется количество лазерных импульсов N, а по горизонтали — плотность энергии Q_0. Затем частота f изменяется, и формируется следующий массив кратеров. Оптические изображения типичных кратеров представлены на рис. 1. Вплоть до $f = 1$ кГц наблюдаются кратеры с ровным краем с образованием валика и с незначительным осаждением продуктов испарения (рис. 1, *a*). Повышение частоты следования импульсов до 1 МГц приводит к образованию кратеров, которое сопровождается факелом и изменением цвета поверхности как внутри кратеров, так и вокруг них (интенсивный красный цвет) (рис. 1, *b*).

Для определения порога абляции при низкой частоте следования импульсов ($f = 1$ кГц) и заданном их количестве N при гауссовом распределении интенсивности по сечению пучка воспользуемся методикой [7], учитывая отклонение формы кратеров от округлой [8]. На рис. 2, *a* представлены зависимости квадратов диаметров эллиптического кратера вдоль большой и малой осей от логарифма падающей энергии лазерного импульса для $N = 20$. Аналогичные зависимости были построены и для других значений N вплоть до 200. В результате аппроксимации полученного набора экспериментальных точек степенной функцией с помощью программы Microsoft Excel 2016 получено следующее уравнение: $y = 100{,}5 \times x^{0{,}94}$ (рис. 2, *b*). Путем сопоставления этой зависимости с известной моделью инкубационного эффекта $NQ_{th}(N) = Q_{th}(1)N^S$ [9] определены порог одноимпульсной абляции $Q_{th}(1)$ и коэффициент инкубации S. Для халькогенидного стекла As_2Se_3 порог одноимпульсной абляции составил примерно 0,1 Дж/см^2, что сопоставимо с данными, полученными в работе [10]. Это значение соответствует плотности мощности излучения $4{,}5 \cdot 10^{11}$ Вт/см^2 при длительности импульса $\tau = 224$ фс. Рассчитанный из экспериментальной зависимости коэффициент инкубации S для халькогенидного стекла As_2Se_3 при $f = 1$ кГц равен 0,94. Таким образом, с ростом количества лазерных импульсов от 1 до 200 наблюдается снижение порога абляции примерно на 25 % (рис. 2, *c*), что подтверждает наличие выраженного инкубационного эффекта. В его основе лежит процесс накопления дефектов в материале, приводящий к структурным изменениям на поверхности стекла. Эти изменения, происходящие от импульса к импульсу, повышают поглощательную способность материала и, как следствие, снижают порог абляции.

Для проведения исследования химического состава продуктов абляции, осажденных на поверхности халькогенидного стекла, методом спектроскопии комбинационного рассеяния света сформирован кратер при плотности энергии $Q_0 \approx 0{,}7$ Дж/см^2 (что соответствует плотности мощности излучения $3{,}1 \cdot 10^{12}$ Вт/см^2 при $\tau = 224$ фс) и $N = 1000$. Зондирование проводилось по точкам на поверхности образца от периферии к центру кратера, включая области внутри и вокруг него.

Спектры КРС в первой области (вокруг кратера) и соответствующие им оптические изображения областей осажденных продуктов абляции при перемещении от периферии (точка *1*) в направлении к кратеру (точки *2–5*) представлены на рис. 3.

В спектре *1* отмечается только вклад от As_2Se_3, положение максимума около 225 см$^{-1}$. В спектре *2* основной вклад вносят As_2Se_3, а также кристалл редкого

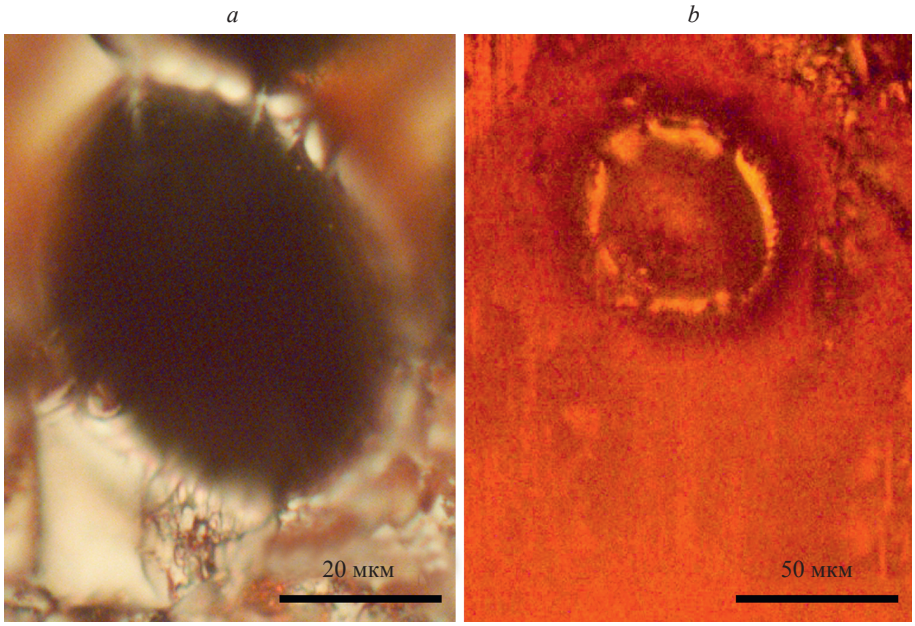

Рис. 1. Оптические изображения типичных кратеров на поверхности халькогенидного стекла As_2Se_3, сформированных при облучении фемтосекундными лазерными импульсами: $Q_0 \approx 0{,}7$ Дж/см^2, $f = 1$ кГц (*a*); $Q_0 \approx 5$ мДж/см^2, $f = 1$ МГц (*b*) при $\lambda = 515$ нм, $\tau = 224$ фс, $N = 1000$

Fig. 1. Optical images of typical craters on the surface of As_2Se_3 chalcogenide glass formed by irradiation with femtosecond laser pulses at: $Q_0 \approx 0.7$ J/cm^2, $f = 1$ kHz (*a*); $Q_0 \approx 5$ mJ/cm^2, $f = 1$ MHz (*b*) ($\lambda = 515$ nm, $\tau = 224$ fs, $N = 1000$)

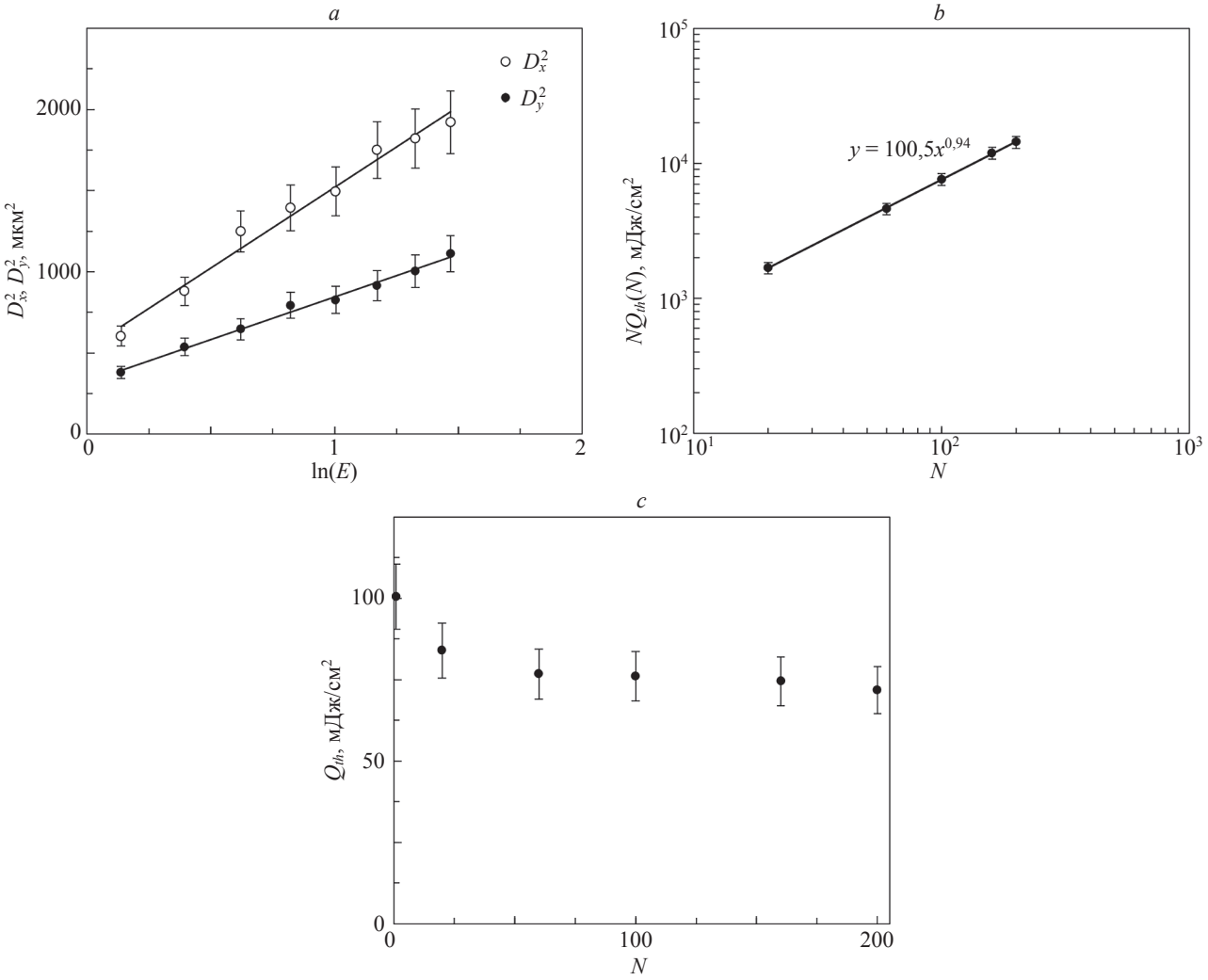

Рис. 2. Экспериментальная зависимость квадратов диаметров эллиптического кратера вдоль большой и малой осей от логарифма падающей энергии фемтосекундного лазерного импульса E при $f = 1$ кГц, $N = 20$ (*a*). Аппроксимация экспериментальных данных степенной функции, подтверждающая наличие инкубационного эффекта (*b*). Зависимость порога абляции халькогенидного стекла As_2Se_3 от количества лазерных импульсов (*c*)

Fig. 2. Experimental dependence of the squares of the elliptical crater diameters along the major and minor axes versus the logarithm of the incident femtosecond laser pulse energy (E) at $f = 1$ kHz, $N = 20$ (*a*). The power-law fit of the experimental data confirming the incubation effect (*b*). Dependence of the ablation threshold of As_2Se_3 chalcogenide glass on the number of laser pulses (*c*)

и ядовитого минерала арсенолита — оксида мышьяка (As_2O_3) (пики 370, 470 и 558 см$^{-1}$). Референсный спектр КРС As_2O_3 (номер 050383 в базе данных The RRUFF™ Project[1]) приведен на рис. 3, *b*. Спектр *3* близок к спектру аморфного селена (*a*-Se), основная широкая полоса около 250 см$^{-1}$, и достигается наибольшая относительная концентрация As_2O_3. Спектр *4* подобен спектру *a*-Se, и присутствуют небольшие пики от As_2O_3. В точке *5* спектр близок к спектру *a*-Se. Таким образом, по мере движения от необлученной области к участкам с осажденными продуктами абляции в спектрах КРС прослеживается следующая закономерность: $As_2Se_3 \rightarrow As_2Se_3 + As_2O_3 \rightarrow As_2Se_3 + As_2O_3 + a\text{-Se} \rightarrow a\text{-Se} + As_2O_3 \rightarrow a\text{-Se}$.

Во второй области (внутри кратера) характерной особенностью спектров КРС является преимущественный вклад от модифицированного лазером As_2Se_3 в аморфизованном стеклообразном состоянии. Под модификацией подразумевается изменение соотношения структурных единиц в стекле и/или появление новых связей по сравнению с исходным стеклом. Это проявляется в сдвиге максимума пика до 232 см$^{-1}$ и появлении явного плеча около 250 см$^{-1}$, что может быть следствием кольцевых структур Se_8 [11]. Для точек внутри кратера отмечается относительно бóльшая интенсивность рассеянного света в области 230–280 см$^{-1}$ по сравнению с немодифицированным стеклом.

Прежде чем выдвинуть обоснованную гипотезу о механизме наблюдаемых химических превращений, проведем анализ накопительного теплового

[1] The RRUFF™ Project database [сайт] Tucson [Электронный ресурс]. URL: https://rruff.info/arsenolite/display=default/R050383 (дата обращения: 07.07.2025).

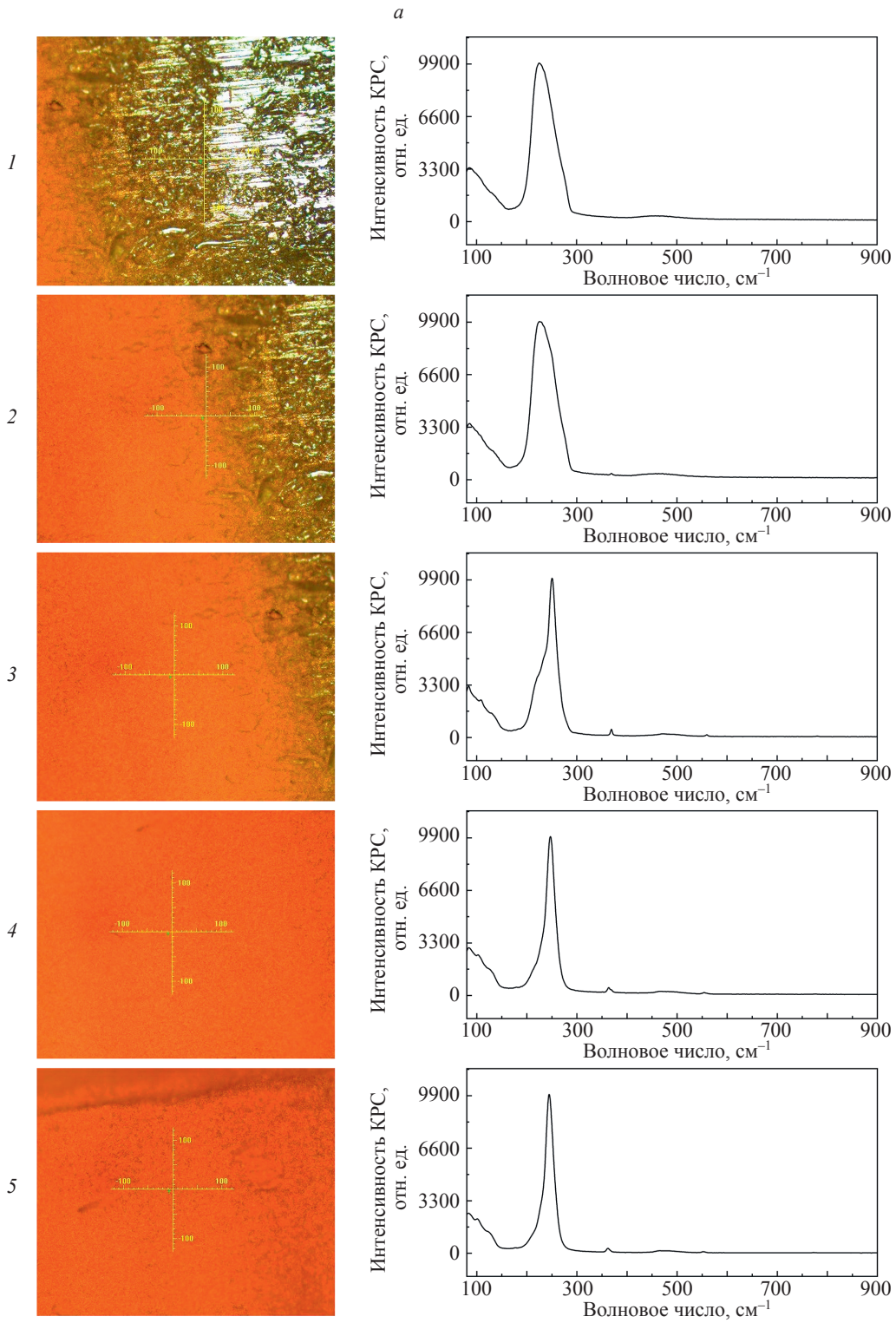

Рис. 3. Оптические изображения областей осажденных продуктов лазерной абляции при перемещении зонда от периферии (точка *1*) к кратеру (точки *2–5*) и соответствующие им спектры комбинационного рассеяния света в диапазоне 80–900 см$^{-1}$, нормированные на 10 000 отн. ед. (λ = 515 нм, τ = 224 фс, $Q_0 \approx$ 0,7 Дж/см^2, N = 1000, f = 1 МГц) (*a*). Спектры комбинационного рассеяния света в диапазоне 300–900 см$^{-1}$ и референсный спектр комбинационного рассеяния света минерала арсенолита по данным базы The RRUFF™ Project[1] (*b*)

Fig. 3. Optical images of the areas with deposited laser ablation products obtained by moving the probe from the periphery (point *1*) to the crater (points *2–5*), and the corresponding Raman spectra in the range of 80–900 cm^{-1}, normalized to 10000 a.u. (λ = 515 nm, τ = 224 fs, $Q_0 \approx$ 0.7 J/cm^2, N = 1000, f = 1 MHz) (*a*). Raman spectra in the range of 300–900 cm^{-1} and the reference Raman spectrum of the mineral arsenolite from The RRUFF™ Project database[1] (*b*)

[1] The RRUFF™ Project database [сайт] Tucson [Электронный ресурс]. URL: https://rruff.info/arsenolite/display=default/R050383 (дата обращения: 07.07.2025).

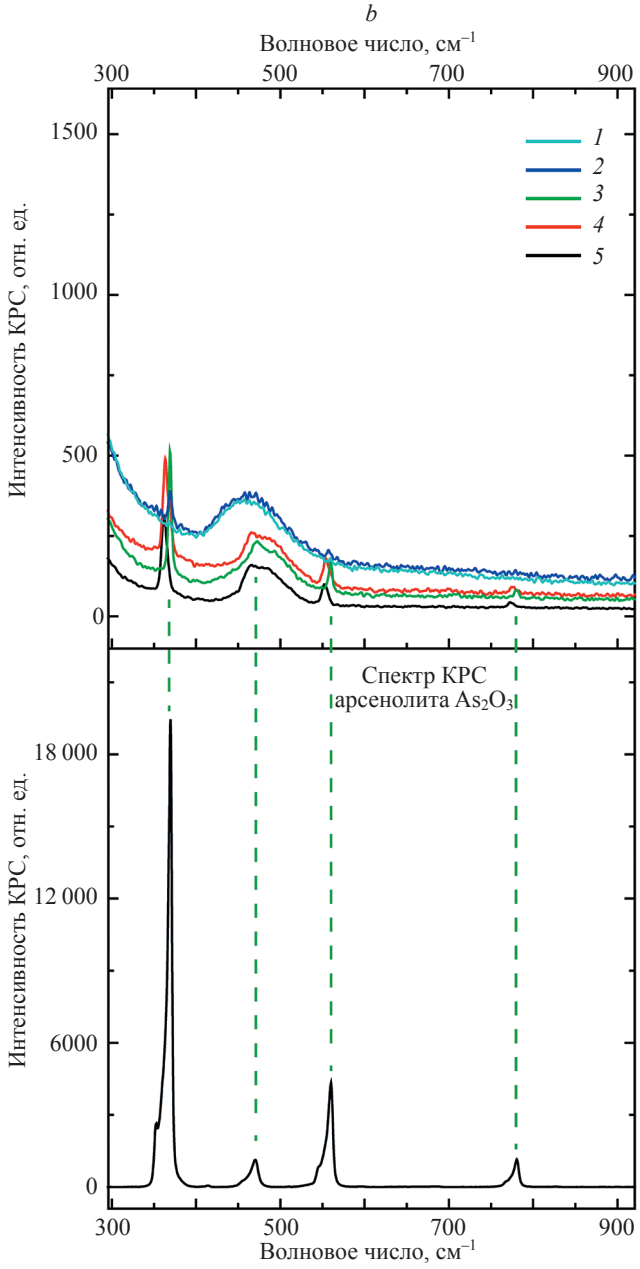

Рис. 3. Продолжение

Fig. 3. Continued

эффекта, который может сопровождать воздействие фемтосекундных лазерных импульсов на халькогенидное стекло As_2Se_3 при частотах их следования более 1 кГц.

Математическое моделирование

Теоретический анализ нагрева и фотовозбуждения халькогенидного стекла одиночным ультракоротким лазерным импульсом строится на основе модели, рассмотренной в [12] для случая облучения тонкой пленки As_2Se_3 ультракоротким лазерным импульсом на длине волны 1030 нм, которая в настоящей работе была адаптирована для случая воздействия на длине волны 515 нм на объемный материал. В модель добавлен учет однофотонного поглощения с коэффициентом поглощения $\beta_1 = 10^5$ см$^{-1}$ [10]. Значение коэффициента двухфотонного поглощения для длины волны 515 нм $\beta_2 = 0{,}14$ см/ГВт [13]. При расчетах использовались значения показателя преломления невозбужденного полупроводника $n = 3{,}4$ [14] и ширины запрещенной зоны $E_g = 1{,}85$ эВ [10]. Остальные параметры соответствуют использованным в работе [12].

Анализ накопительного нагрева поверхности материала при многоимпульсном лазерном облучении проводится с учетом трехмерного теплоотвода в условиях, когда длительность теплового действия импульса определяется временем перехода поглощенной энергии лазерного излучения в тепловую энергию (примерно 10^{-12}–10^{-11} с) и мала по сравнению с периодом следования фемтосекундных лазерных импульсов, а глубина фотовозбужденной области существенно меньше дли-

ны теплопроводности $(4a/f)^{1/2}$, где a — коэффициент температуропроводности материала.

Воспользуемся моделью теплового источника в форме диска [15] (рис. 4). Тогда остаточная температура (т. е. температура поверхности материала после воздействия N импульсов в момент прихода $N+1$-го импульса) в центре пучка на поверхности материала определяется согласно следующему выражению

$$T_N = \frac{(1-R)Q_0}{c_i}\sqrt{\frac{f}{\pi a}}\sum_{i=1}^{N}\frac{1}{\sqrt{i}(\psi i+1)} + T_0, \qquad (1)$$

где $R = 0{,}25$ — коэффициент отражения халькогенидного стекла As_2Se_3 [16]; $a = 1{,}9\cdot10^{-3}$ см²/с — коэффициент температуропроводности халькогенидного стекла As_2Se_3, рассчитанный по данным [17–19]; $\psi = \frac{4a}{fr_0^2}$ — безразмерный параметр; $T_0 = 300$ К — начальная температура.

Как видно из рис. 4, *a* при $Q_0 = 0{,}1$ Дж/см² (соответствует порогу одноимпульсной абляции) в течение лазерного импульса достигаются высокие концентрации фотовозбужденных носителей n_e около $6\cdot10^{21}$ см$^{-3}$, а их температура T_e примерно 40 кК. Дальнейшая передача энергии в кристаллическую решетку за счет электрон-фононного взаимодействия приводит к выравниванию электронной и решеточной (T_i) температур на значении примерно 1900 К, что существенно превосходит как температуру плавления, так и кипения As_2Se_3 (примерно 1120 К [20]). Известно, что основными механизмами абляции твердых тел ультракороткими импульсами является «откольная» абляция, реализуемая при относительно малых плотностях энергии импульса, и «фазовый взрыв», реализующийся при дальнейшем повышении плотности энергии [21].

В настоящей работе абляция вблизи пороговых значений плотности энергии может быть ассоциирована с первым механизмом. При снижении плотности энергии до 5 мДж/см² (что соответствует плотности

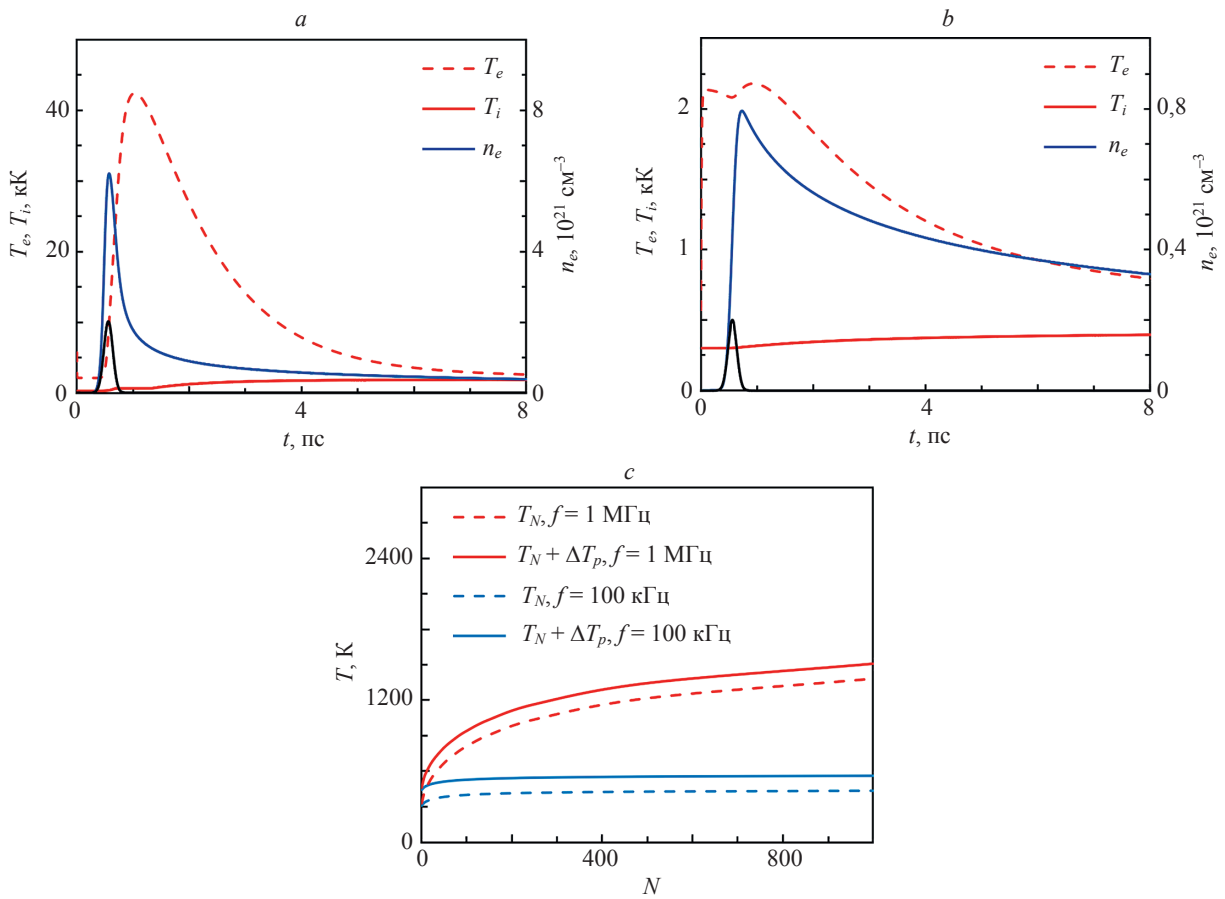

Рис. 4. Теоретические зависимости: электронной T_e и решеточной T_i температур, а также концентрации фотовозбужденных носителей n_e на поверхности халькогенидного стекла As_2Se_3 от времени для $Q_0 = 0{,}1$ Дж/см² (*a*) и $Q_0 = 5$ мДж/см² (*b*); остаточной температуры T_N и температуры $T_N + \Delta T_p$ в центре лазерного пятна на поверхности халькогенидного стекла As_2Se_3 от количества импульсов при разных частотах их следования (*c*), при $\lambda = 515$ нм, $\tau = 224$ фс, $d_0 = 30$ мкм, $Q_0 = 5$ мДж/см². Черная кривая (рис. 4, *a*, *b*) — форма импульса; ΔT_p — приращение температуры поверхности за импульс

Fig. 4. Theoretical dependences of: the electron (T_e) and lattice (T_i) temperatures as well as the concentration of photoexcited carriers (n_e) on the surface of As_2Se_3 chalcogenide glass on time for $Q_0 = 0.1$ J/cm² (*a*) and $Q_0 = 5$ mJ/cm² (*b*) (the black curve represents the pulse shape); the residual temperature T_N and temperature $T_N + \Delta T_p$ (where ΔT_p is the surface temperature increase per pulse) in the center of the laser spot on the surface of As_2Se_3 chalcogenide glass on the number of pulses at different repetition rates (*b*) ($\lambda = 515$ nm, $\tau = 224$ fs, $d_0 = 30$ μm, $Q_0 = 5$ mJ/cm²)

мощности излучения $2{,}2\cdot10^{10}$ Вт/см^2 при $\tau = 224$ фс) концентрация фотовозбужденных носителей составляет около $0{,}8\cdot10^{21}$ см$^{-3}$, а их температура достигает примерно 2000 К (рис. 4, *b*). После передачи энергии в кристаллическую решетку через электрон-фононное взаимодействие температура решетки повышается до примерно 430 К (т. е. приращение температуры за импульс $\Delta T_p = 130$ К), что, очевидно, недостаточно для абляции материала за 1 импульс. При многоимпульсном лазерном облучении с той же плотностью энергии ($f = 1$ МГц и $N = 1000$) остаточная температура поверхности T_N, определяемая согласно выражению (1), может превышать температуру кипения As_2Se_3 (рис. 4, *c*). Таким образом, в этом случае за счет накопительного нагрева может реализовываться механизм нормального испарения, не характерный для воздействия одиночных лазерных импульсов, что соответствует условиям образования большого количества паров селена (Se) с последующим осаждением аморфной фазы, наблюдаемым в эксперименте (рис. 1, *b*). При этом снижение частоты следования импульсов до 100 кГц приводит к снижению остаточной температуры до значений около 500 К, а при частоте 1 кГц значения остаточной температуры не превосходят 1 К (не показано на рис. 4), т. е. в этом случае эффект накопления тепла пренебрежимо мал.

Важно отметить, что при плотностях энергии $Q_0 > 0{,}1$ Дж/см^2 в условиях многоимпульсного облучения ($f = 1$ МГц, $N = 1000$) накопительный нагрев способствует достижению высоких температур, при которых, кроме *a*-Se, также наблюдается интенсивное образование кристаллов As_2O_3 (рис. 3).

Обсуждение

В работе [22] был предложен механизм реакции фоторазложения As_2Se_3 под действием лазерного излучения видимого диапазона, согласно которому поглощение фотона с энергией, превышающей запрещенную зону, приводит к разрыву связей и образованию атомов мышьяка (As), которые могут свободно диффундировать в стекле и в присутствии водяного пара (H_2O) (в качестве катализатора) участвовать в образовании As_2O_3:

$$As_2Se_3 \xrightarrow{h\nu} xAs + As_{2-x}Se_3$$

$$0 < x < 2,\ 4As + 3O_2 \xrightarrow{H_2O} 2As_2O_3.$$

При этом предполагалось, что все химические превращения в халькогенидном стекле реализуются в твердой фазе.

В [23] при изучении абляции пленок сульфида мышьяка и As_2Se_3 в режимах сканирования фемтосекундных лазерных импульсов с высокой частотой следования ($\lambda = 1030$ нм), обнаружили образование кристаллов As_2O_3 на абляционных линиях лазерной записи. Для объяснения химических превращений состава пленки в работе [23] также использован твердофазный механизм реакции фоторазложения As_2Se_3, предложенный в работе [22].

Впервые на возможность развития окислительно-восстановительных реакций не в твердой, а в паровой фазе при короткоимпульсной лазерной абляции металлов на воздухе было указано в работе [24]. Позднее в [25] была разработана химико-гидродинамическая модель развития химических реакций в плазменном факеле как основном механизме образования оксидных слоев на поверхности титана. Проведенное в настоящей работе исследование продуктов лазерной абляции и рассчитанные значения температуры позволяют предположить, что окислительно-восстановительные процессы в объемном халькогенидном стекле As_2Se_3 развиваются в паровой фазе при частотах следования фемтосекундных лазерных импульсов порядка 1 МГц.

В настоящей работе в условиях интенсивного накопительного нагрева халькогенидное стекло плавится и испаряется. Над поверхностью абляции образуется плазменный факел. Плазма изначально состоит из паров Se и As. По мере расширения As в паровой фазе вступает во взаимодействие с атмосферным кислородом (O_2) и возвращается на поверхность в виде As_2O_3. Парофазный механизм включает: испарение As_2Se_3, образование паров Se и As, окисление As в паровой фазе при взаимодействии с O_2 и осаждение As_2O_3 и *a*-Se на поверхность образца. Поскольку As является летучим компонентом, он распространяется на большие расстояния от облученной области, что выражается в неравномерном осаждении As_2O_3 на поверхности образца. *a*-Se осаждается преимущественно ближе к кратеру. В самом же кратере образуется модифицированный As_2Se_3 с повышенной концентрацией кольцевых структур Se_8.

Заключение

Экспериментально определен порог одноимпульсной абляции объемного халькогенидного стекла As_2Se_3 при низкой частоте следования фемтосекундных лазерных импульсов ($f = 1$ кГц) на длине волны 515 нм, который составил порядка 0,1 Дж/см^2. Рассчитан коэффициент инкубации $S = 0{,}94$, характеризующий снижение порога абляции с увеличением количества импульсов. Представленные оценки температур нагрева за один импульс хорошо коррелируют с экспериментально определенным порогом одноимпульсной абляции и могут быть использованы в дальнейшем для анализа процессов взаимодействия ультракоротких лазерных импульсов с халькогенидным стеклом As_2Se_3 на длине волны 515 нм. При увеличении частоты следования импульсов до 1 МГц значительным становится накопительный тепловой эффект. Результаты теоретического моделирования продемонстрировали, что при многоимпульсном фемтосекундном лазерном воздействии с частотой следования импульсов 1 МГц накопительный нагрев на поверхности халькогенидного стекла As_2Se_3 может достигать температур, превосходящих температуру кипения As_2Se_3, и способствовать активному развитию парофазных окислительно-восстановительных процессов.

Результаты исследования могут быть использованы для оптимизации режимов фемтосекундного микроструктурирования халькогенидных стекол при высокой частоте следования лазерных импульсов.

Литература

1. Shirshnev P.S., Kang M., Divliansky I., Richardson K.A., Glebov L.B. Engineered refractive and diffractive optical composites via photothermal processes // Optical Materials Express. 2022. V. 12. N 9. P. 3429–3448. https://doi.org/10.1364/ome.465711
2. Kang M., Swisher A.M., Pogrebnyanov A.V., Liu L., Kirk A., Aiken S., et al. Ultralow dispersion multicomponent thin-film chalcogenide glass for broadband gradient-index optics // Advanced Materials. 2018. V. 30. N 39. P. 1803628. https://doi.org/10.1002/adma.201803628
3. Kang M., Triplett B.M., Shalaginov M.Y., Deckoff-Jones S., Blanco C., Truman M., et al. Photochemically engineered large-area arsenic sulfide micro-gratings for hybrid diffractive–refractive infrared platforms // Advanced Photonics Research. 2024. V. 5. N 1. P. 2300241. https://doi.org/10.1002/adpr.202300241
4. Li Z., Shangguan S., Shi W., Wang W., Qi D., Chen F., Zheng H. Femtosecond laser induced periodic subwavelength nanohole arrays structure on As_2Se_3 glass surface // Ceramics International. 2024. V. 51. N 7. P. 8211–8218. https://doi.org/10.1016/j.ceramint.2024.12.302
5. Guk I.V., Kuzmin E.V., Shandybina G.D., Yakovlev E.B., Dyukin R.V., Kulagin V.S. Influence of multi-pulse action on the evolution of silicon microrelief under femtosecond laser irradiation // Journal of Optical Technology. 2017. V. 84. N 7. P. 462–466. https://doi.org/10.1364/JOT.84.000462
6. Fraggelakis F., Mincuzzi G., Lopez J., Manek-Honninger I., Kling R. Texturing metal surface with MHz ultra-short laser pulses // Optics Express. 2017. V. 25. N 15. P. 18131–18139. https://doi.org/10.1364/OE.25.018131
7. Liu J.M. Simple technique for measurements of pulsed Gaussian-beam spot sizes // Optics Letters. 1982. V. 7. N 5. P. 196–198. https://doi.org/10.1364/OL.7.000196
8. Струлёва Е.В., Комаров П.С., Ашитков С.И. Термомеханическая абляция титана при фемтосекундном лазерном воздействии // Теплофизика высоких температур. 2019. Т. 57. № 4. С. 529–533. https://doi.org/10.1134/S0040364419040215
9. Jee Y., Becker M.F., Walser R.M. Laser-induced damage on single-crystal metal surfaces // Journal of the Optical Society of America B. 1988. V. 5. N 3. P. 648–659. https://doi.org/10.1364/JOSAB.5.000648
10. Li Q., Qi D., Wang X., Shen X., Wang R., Tanaka K. Femto- and nano-second laser-induced damages in chalcogenide glasses // Japanese Journal of Applied Physics. 2019. V. 58. N 8. P. 080911. https://doi.org/10.7567/1347-4065/ab3368
11. Kovanda V., Vlček M., Jain H. Structure of As–Se and As–P–Se glasses studied by Raman spectroscopy // Journal of Non-Crystalline Solids. 2003. V. 326-327. P. 88–92. https://doi.org/10.1016/S0022-3093(03)00383-1
12. Shamova A., Polyakov D., Golubkov V., Pankin D., Shimko A., Shandybina G. Processes accompanying ablation in thin-film and bulk chalcogenide glass As_2Se_3 under multipulse femtosecond laser irradiation // Optical Materials. 2025. V. 167. P. 117232. https://doi.org/10.1016/j.optmat.2025.117232
13. Zakery A., Elliott S.R. Optical properties and applications of chalcogenide glasses: a review // Journal of Non-Crystalline Solids. 2003. V. 330. N 1-3. P. 1–12. https://doi.org/10.1016/j.jnoncrysol.2003.08.064
14. Joseph S., Sarkar S., Joseph J. Grating-coupled surface plasmon-polariton sensing at a flat metal–analyte interface in a hybrid-configuration // ACS Applied Materials & Interfaces. 2020. V. 12. N 41. P. 46519–46529. https://doi.org/10.1021/acsami.0c12525
15. Polyakov D., Shandybina G., Shamova A. Analytical 3D modeling of accumulative heating under multipulse laser irradiation of inorganic materials and biological tissues // Thermal Science and Engineering Progress. 2022. V. 31. P. 101284. https://doi.org/10.1016/j.tsep.2022.101284
16. Sobolev V.V. Optical spectra of arsenic chalcogenide glasses over a wide energy range of fundamental absorption // Glass Physics and Chemistry. 2002. V. 28. N 6. P. 399–416. https://doi.org/10.1023/A:1021719017568
17. Easteal A.J., Wilder J.A., Mohr R.K., Moynihan C.T. Heat capacity and structural relaxation of enthalpy in As_2Se_3 glass // Journal of the American Ceramic Society. 1977. V. 60. N 3-4. P. 134–138. https://doi.org/10.1111/j.1151-2916.1977.tb15488.x
18. Knotek P., Kutálek P., Černošková E., Vlček M., Tichý L. The density, nanohardness and some optical properties of As–S and As–Se

References

1. Shirshnev P.S., Kang M., Divliansky I., Richardson K.A., Glebov L.B. Engineered refractive and diffractive optical composites via photothermal processes. *Optical Materials Express*, 2022, vol. 12, no. 9, pp. 3429–3448. https://doi.org/10.1364/ome.465711
2. Kang M., Swisher A.M., Pogrebnyanov A.V., Liu L., Kirk A., Aiken S., et al. Ultralow dispersion multicomponent thin-film chalcogenide glass for broadband gradient-index optics. *Advanced Materials*, 2018, vol. 30, no. 39, pp. 1803628. https://doi.org/10.1002/adma.201803628
3. Kang M., Triplett B.M., Shalaginov M.Y., Deckoff-Jones S., Blanco C., Truman M., et al. Photochemically engineered large-area arsenic sulfide micro-gratings for hybrid diffractive–refractive infrared platforms. *Advanced Photonics Research*, 2024, vol. 5, no. 1, pp. 2300241. https://doi.org/10.1002/adpr.202300241
4. Li Z., Shangguan S., Shi W., Wang W., Qi D., Chen F., Zheng H. Femtosecond laser induced periodic subwavelength nanohole arrays structure on As_2Se_3 glass surface. *Ceramics International*, 2024, vol. 51, no. 7, pp. 8211–8218. https://doi.org/10.1016/j.ceramint.2024.12.302
5. Guk I.V., Kuzmin E.V., Shandybina G.D., Yakovlev E.B., Dyukin R.V., Kulagin V.S. Influence of multi-pulse action on the evolution of silicon microrelief under femtosecond laser irradiation. *Journal of Optical Technology*, 2017, vol. 84, no. 7, pp. 462–466. https://doi.org/10.1364/JOT.84.000462
6. Fraggelakis F., Mincuzzi G., Lopez J., Manek-Honninger I., Kling R. Texturing metal surface with MHz ultra-short laser pulses. *Optics Express*, 2017, vol. 25, no. 15, pp. 18131–18139. https://doi.org/10.1364/OE.25.018131
7. Liu J.M. Simple technique for measurements of pulsed Gaussian-beam spot sizes. *Optics Letters*, 1982, vol. 7, no. 5, pp. 196–198. https://doi.org/10.1364/OL.7.000196
8. Struleva E.V., Komarov P.S., Ashitkov S.I. Thermomechanical ablation of titanium by femtosecond laser irradiation. *High Temperature*, 2019, vol. 57, no. 4, pp. 486–489. https://doi.org/10.1134/s0018151x19040217
9. Jee Y., Becker M.F., Walser R.M. Laser-induced damage on single-crystal metal surfaces. *Journal of the Optical Society of America B*, 1988, vol. 5, no. 3, pp. 648–659. https://doi.org/10.1364/JOSAB.5.000648
10. Li Q., Qi D., Wang X., Shen X., Wang R., Tanaka K. Femto- and nano-second laser-induced damages in chalcogenide glasses. Japanese *Journal of Applied Physics*, 2019, vol. 58, no. 8, pp. 080911. https://doi.org/10.7567/1347-4065/ab3368
11. Kovanda V., Vlček M., Jain H. Structure of As–Se and As–P–Se glasses studied by Raman spectroscopy. *Journal of Non-Crystalline Solids*, 2003, vol. 326-327, pp. 88–92. https://doi.org/10.1016/S0022-3093(03)00383-1
12. Shamova A., Polyakov D., Golubkov V., Pankin D., Shimko A., Shandybina G. Processes accompanying ablation in thin-film and bulk chalcogenide glass As2Se3 under multipulse femtosecond laser irradiation. *Optical Materials*, 2025, vol. 167, pp. 117232. https://doi.org/10.1016/j.optmat.2025.117232
13. Zakery A., Elliott S.R. Optical properties and applications of chalcogenide glasses: a review. *Journal of Non-Crystalline Solids*, 2003, vol. 330, no. 1-3, pp. 1–12. https://doi.org/10.1016/j.jnoncrysol.2003.08.064
14. Joseph S., Sarkar S., Joseph J. Grating-coupled surface plasmon-polariton sensing at a flat metal–analyte interface in a hybrid-configuration. *ACS Applied Materials & Interfaces*, 2020, vol. 12, no. 41, pp. 46519–46529. https://doi.org/10.1021/acsami.0c12525
15. Polyakov D., Shandybina G., Shamova A. Analytical 3D modeling of accumulative heating under multipulse laser irradiation of inorganic materials and biological tissues. *Thermal Science and Engineering Progress*, 2022, vol. 31, pp. 101284. https://doi.org/10.1016/j.tsep.2022.101284
16. Sobolev V.V. Optical spectra of arsenic chalcogenide glasses over a wide energy range of fundamental absorption. *Glass Physics and Chemistry*, 2002, vol. 28, no. 6, pp. 399–416. https://doi.org/10.1023/A:1021719017568
17. Easteal A.J., Wilder J.A., Mohr R.K., Moynihan C.T. Heat capacity and structural relaxation of enthalpy in As_2Se_3 glass. *Journal of the American Ceramic Society*, 1977, vol. 60, no. 3-4, pp. 134–138. https://doi.org/10.1111/j.1151-2916.1977.tb15488.x
18. Knotek P., Kutálek P., Černošková E., Vlček M., Tichý L. The density, nanohardness and some optical properties of As–S and As–Se

chalcogenide bulk glasses and thin films // RSC Advances. 2020. V. 10. N 70. P. 42744–42753. https://doi.org/10.1039/D0RA08939G
19. Flasck R., Rockstad H.K. The thermal conductivity of some chalcogenide glasses // Journal of Non-Crystalline Solids. 1973. V. 12. N 3. P. 353–356. https://doi.org/10.1016/0022-3093(73)90007-0
20. Baker E.H. Arsenic triselenide: boiling-point relation at elevated pressures // Journal of the Chemical Society, Dalton Transactions. 1975. N 15. P. 1589–1591. https://doi.org/10.1039/DT9750001589
21. Shugaev M.V., He M., Levy Y., Mazzi A., Miotello A., Bulgakova N.M., Zhigilei L.V. Laser-induced thermal processes: heat transfer, generation of stresses, melting and solidification, vaporization, and phase explosion // Handbook of Laser Micro- and Nano-Engineering. 2021. P. 83–163. https://doi.org/10.1007/978-3-030-63647-0_11
22. Berkes J.S., Ing S.W., Hillegas W.J. Photodecomposition of amorphous As_2Se_3 and As_2S_3 // Journal of Applied Physics. 1971. V. 42. N 12. P. 4908–4916. https://doi.org/10.1063/1.1659873
23. Paula K.T., Dutta N.S., Almeida J.M., Nolasco L.K., Andrade M.B., Arnold C.B., Mendonça C.R. Femtosecond laser induced damage threshold incubation and oxidation in As_2S_3 and As_2Se_3 thin films // Applied Surface Science. 2024. V. 654. P. 159449. https://doi.org/10.1016/j.apsusc.2024.159449
24. Либенсон М.Н., Шандыбина Г.Д., Шахмин А.Л. Химический анализ продуктов лазерной абляции наносекундного диапазона // Журнал технической физики. 2000. Т. 70. N 9. С. 124–127.
25. Gornushkin I.B., Veiko V.P., Karlagina Yu.Yu., Samokhvalov A.A., Polyakov D.S. Equilibrium model of titanium laser induced plasma in air with reverse deposition of titanium oxides // Spectrochimica Acta Part B: Atomic Spectroscopy. 2022. V. 193. P. 106449. https://doi.org/10.1016/j.sab.2022.106449

chalcogenide bulk glasses and thin films. *RSC Advances*, 2020, vol. 10, no. 70, pp. 42744–42753. https://doi.org/10.1039/D0RA08939G
19. Flasck R., Rockstad H.K. The thermal conductivity of some chalcogenide glasses. *Journal of Non-Crystalline Solids*, 1973, vol. 12, no. 3, pp. 353–356. https://doi.org/10.1016/0022-3093(73)90007-0
20. Baker E.H. Arsenic triselenide: boiling-point relation at elevated pressures. *Journal of the Chemical Society*, Dalton Transactions, 1975, no. 15, pp. 1589–1591. https://doi.org/10.1039/DT9750001589
21. Shugaev M.V., He M., Levy Y., Mazzi A., Miotello A., Bulgakova N.M., Zhigilei L.V. Laser-induced thermal processes: heat transfer, generation of stresses, melting and solidification, vaporization, and phase explosion. *Handbook of Laser Micro- and Nano-Engineering*, 2021, pp. 83–163. https://doi.org/10.1007/978-3-030-63647-0_11
22. Berkes J.S., Ing S.W., Hillegas W.J. Photodecomposition of amorphous As_2Se_3 and As_2S_3. *Journal of Applied Physics*, 1971, vol. 42, no. 12, pp. 4908–4916. https://doi.org/10.1063/1.1659873
23. Paula K.T., Dutta N.S., Almeida J.M., Nolasco L.K., Andrade M.B., Arnold C.B., Mendonça C.R. Femtosecond laser induced damage threshold incubation and oxidation in As_2S_3 and As_2Se_3 thin films. *Applied Surface Science*, 2024, vol. 654, pp. 159449. https://doi.org/10.1016/j.apsusc.2024.159449
24. Libenson M.N., Shandybina G.D., Shakhmin A.L. Chemical analysis of products obtained by nanosecond laser ablation. *Technical Physics*, 2000, vol. 45, no. 9, pp. 1219–1222. https://doi.org/10.1134/1.1318114
25. Gornushkin I.B., Veiko V.P., Karlagina Yu.Yu., Samokhvalov A.A., Polyakov D.S. Equilibrium model of titanium laser induced plasma in air with reverse deposition of titanium oxides. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2022, vol. 193, pp. 106449. https://doi.org/10.1016/j.sab.2022.106449

Авторы

Шамова Александра Андреевна — кандидат физико-математических наук, ассистент, младший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57103706700, https://orcid.org/0000-0001-8728-260X, aashamova@itmo.ru

Поляков Дмитрий Сергеевич — кандидат физико-математических наук, ассистент, научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57212832118, https://orcid.org/0000-0002-6240-4057, polyakovdmitry1988@gmail.com

Долгополов Артур Джуракулович — ассистент, инженер-исследователь, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 58637797100, https://orcid.org/0000-0002-9548-791X, addolgopolov@itmo.ru

Панькин Дмитрий Васильевич — кандидат физико-математических наук, специалист, Санкт-Петербургский государственный университет, Научный парк, Ресурсный центр «Оптические и лазерные методы исследования вещества», Санкт-Петербург, Петергоф, 198504, Российская Федерация, sc 57190487236, https://orcid.org/0000-0003-3114-3410, dmitrii.pankin@spbu.ru

Шандыбина Галина Дмитриевна — кандидат физико-математических наук, доцент, ассистент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 6602435275, https://orcid.org/0000-0003-3501-1925, corchand@gmail.com

Authors

Alexandra A. Shamova — PhD (Physics & Mathematics), Assistant, Junior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57103706700, https://orcid.org/0000-0001-8728-260X, aashamova@itmo.ru

Dmitry S. Polyakov — PhD (Physics & Mathematics), Assistant, Scientific Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57212832118, https://orcid.org/0000-0002-6240-4057, polyakovdmitry1988@gmail.com

Arthur D. Dolgopolov — Assistant, Engineer-Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 58637797100, https://orcid.org/0000-0002-9548-791X, addolgopolov@itmo.ru

Dmitrii V. Pankin — PhD (Physics & Mathematics), Specialist, St. Petersburg State University, Research Park, Center for optical and laser materials research, Saint Petersburg, Peterhof, 198504, Russian Federation, sc 57190487236, https://orcid.org/0000-0003-3114-3410, dmitrii.pankin@spbu.ru

Galina D. Shandybina — PhD (Physics & Mathematics), Associate Professor, Assistant, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 6602435275, https://orcid.org/0000-0003-3501-1925, corchand@gmail.com

Статья поступила в редакцию 31.07.2025
Одобрена после рецензирования 25.09.2025
Принята к печати 12.11.2025

Received 31.07.2025
Approved after reviewing 25.09.2025
Accepted 12.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1014-1023

Modeling and study of FBG interrogator based on a two-dimensional image sensor

Sugumar Venkatesan[1], Subashini Ponnusamy[2], Pandian Chelliah[3]✉

[1,3] Ramakrishna Mission Vivekananda College (Affiliated to University of Madras), Chennai, 600004, India
[2] Government Arts College, Chennai, 600035, India

[1] sugumartamizh@gmail.com, https://orcid.org/0009-0009-2243-1251
[2] shiniponnusamy18@gmail.com, https://orcid.org/0009-0002-0501-7621
[3] cpandian@rkmvc.ac.in✉, https://orcid.org/0000-0002-8847-4520

Abstract
Fiber Bragg Grating (FBG) interrogators contain a movable scattering element that tracks the FBG central wavelength. The movable element of the interrogator limits the interrogation speed. This paper proposes an interrogation method that does not use movable elements. This is achieved by using an Array Waveguide (AWG) to split the FBG reflected spectrum and a Convolutional Neural Network (CNN) for training to determine the central wavelength. Most of the known studies consider the AWG output as a one-dimensional data array for training the neural network. However, CNNs work best with two-dimensional images. This paper proposes to transform the AWG output using a two-dimensional image sensor with a circular configuration. This allows for higher accuracy and improved resolution in predicting the central wavelength. The AWG signal is projected onto a two-dimensional image sensor which has either a grid or a circular configuration. The number of AWG channels used is 32, which corresponds to a distance between channel wavelengths of 0.0625 nm. The circular configuration enables more accurate feature extraction using CNN. A 32-beam passive waveguide array in a circular configuration is used for FBG interrogation. It projects the FBG output signals onto the image sensor, enabling high-resolution Bragg wavelength prediction. Computer simulation of the proposed interrogation device demonstrated a predicted resolution of ±1 pm with 98 % accuracy. It should be noted that the presented values are estimates and are subject to refinement using a hardware prototype. Such devices are relatively easy to manufacture and are readily available to consumers.

Keywords
fiber Bragg grating interrogator, convolutional neural network, arrayed waveguide grating, image sensor

For citation: Venkatesan S., Ponnusamy S., Chelliah P. Modeling and study of FBG interrogator based on a two-dimensional image sensor. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1014–1023. doi: 10.17586/2226-1494-2025-25-6-1014-1023

УДК 681.7.068

Моделирование и исследование ВБР-интеррогатора на основе двумерного датчика изображения

Сугумар Венкатесан[1], Субашини Поннусами[2], Пандиан Челия[3]✉

[1,3] Рамакришна Мишен Вивекананда Колледж (филиал Мадрасского университета), Ченнай, 600004, Индия
[2] Государственный мужской колледж искусств, Ченнай, 600035, Индия

[1] sugumatamizh@gmail.com, https://orcid.org/0009-0009-2243-1251
[2] shiniponnusamy18@gmail.com, https://orcid.org/0009-0002-0501-7621
[3] cpandian@rkmvc.ac.in✉, https://orcid.org/0000-0002-8847-4520

Аннотация
Интеррогаторы на основе волоконной брэгговской решетки (ВБР) содержат подвижный рассеивающий элемент, который отслеживает центральную длину волны решетки. Подвижный элемент интеррогатора ограничивает скорость опроса. Предлагается метод опроса, не использующий подвижные элементы. Это достигается применением массива волноводов (Arrayed Waveguide Grating, AWG) для разделения отраженного спектра ВБР и сверточной нейронной сети для обучения с целью определения центральной длины волны. В большинстве

© Venkatesan S., Ponnusamy S., Chelliah P., 2025

известных исследований выход AWG рассматривается как одномерный массив данных для обучения нейронной сети. Однако сверточная нейронная сеть лучше всего работает с двумерными изображениями. Представлен подход, преобразующий выход AWG с помощью двумерного датчика изображения круговой конфигурации. Метод позволяет повысить точность и улучшить разрешение при прогнозировании центральной длины волны. Сигнал AWG проецируется на двумерный датчик изображения, имеющий форму сетки или круговую конфигурацию. Количество используемых каналов AWG составляет 32, что соответствует расстоянию между длинами волн каналов 0,0625 нм. Круговая конфигурация обеспечивает более точное извлечение признаков с помощью сверточной нейронной сети. Для опроса ВБР используется 32-лучевая пассивная решетка волноводов в круговой конфигурации. Она обеспечивает проецирование выходных сигналов ВБР на датчик изображения, что дает возможность прогнозировать длину волны Брэгга с высоким разрешением. Компьютерное моделирование предложенного устройства опросов продемонстрировало прогнозируемое разрешение ± 1 пм с точностью 98 %. Отметим, что представленные значения являются оценочными и подлежат уточнению на аппаратном прототипе. Такие устройства сравнительно просты в изготовлении и доступны потребителям.

Ключевые слова
интерпретатор на основе волоконной брэгговской решетки, сверточная нейронная сеть, многослойная волноводная решетка, датчик изображения

Ссылка для цитирования: Венкатесан С., Поннусами С., Челия П. Моделирование и исследование ВБР-интеррогатора на основе двумерного датчика изображения // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1014–1023 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-6-1014-1023

Introduction

Fiber Bragg Grating (FBGs) have been used in monitoring civil infrastructures, such as bridges and offshore platforms (for oil explorations) [1], load-bearing structures in the aerospace industry [2], and nuclear power plants [3]. FBGs have also been used as electrochemical sensors [4] and biosensors [5].

FBGs are periodic perturbations of the refractive index in the core of the optical fiber. The periodicity and the effective refractive index of the material of the core of the fiber determine the characteristic central wavelength of the light that is reflected when a broadband light is launched into it [6]. To monitor temperature, strain or any other ambient physical field that affects the FBG, its central Bragg wavelength has to be monitored using an interrogator. A typical interrogator uses a dispersive element such as a monochromator to monitor the wavelengths [7, 8]. These moving type dispersive elements make the interrogators slow and unusable in dynamic environments such as monitoring high-frequency vibration. To realise a fast interrogation system, many novel attempts have been made. Cui et al. [9] made the light from a sensor FBG reflect from a matched reference FBG, and the ratio of the overlapping spectra between the two FBGs to monitor the central wavelength is used. Diaz et al. [10] used an inline Fabry Perot Interferometer (FPI). They convoluted the FBG sensor spectra with the FPI spectra to acquire the data using a photo detector at a much faster rate. Lei et al. [11] used a novel time-stretching method to monitor FBGs. Here, the authors made an ultra-short optical pulse to propagate through a dispersive single-mode fiber. As the pulse undergoes dispersion, the authors mapped time to wavelength. This stretched pulse was allowed to propagate through the FBG. The pulse propagating through the FBG was modulated by it. This modulated pulse was made to pass through another dispersive single-mode fiber. Thus, by monitoring the time stretch between a sensing FBG and a reference FBG, the authors were able to interrogate the FBGs. The problem with all these above techniques is that they need a reference FBG. These reference FBGs are difficult to characterise and the reference wavelength of these FBGs often changes with changing ambient environment.

Arrayed Waveguide Gratings (AWG) are a good alternative to the conventional dispersive elements used in the interrogation system since they do not have any moving parts. Marrazzo et al. [12] coupled the output of each AWG channel to a photodiode in their design of the FBG interrogator. Using an analytical model, the authors showed that the wavelength shift of the FBG was a nonlinear function of the outputs of the AWGs. Niewczas et al. [13] did similar work but employed curve fitting to measure the wavelength shift from the AWG outputs. Marrazzo et al. [14] interrogated the FBGs using AWGs by employing a complete analogue system without using ADC or DAC. They did this by using Programmable logic controllers [13]. Trita et al. [15] designed and developed a compact FBG interrogator using AWGs. These interrogators are now being superseded by using an artificial neural network. Barino et al. [16] used an array of FBGs instead of AWG to split the spectra from a long-period grating into various wavelength divisions. The optical output from each of the FBG was used to train a neural network. Chen et al. [17] designed an FBG interrogator by feeding the output of AWG to a one-dimensional Convolutional Neural Network (CNN). The authors used a power meter to measure the output of each of the AWGs channels. Ren et al. [18] used nine channel AWG to interrogate the FBGs, by exploiting the overlapping nature of the spectra from the AWG channels and a neural network. They had an accuracy of + 68 pm. Most work in literature uses a linear array of Charge-Coupled Device sensors or linear detectors along with an AWG for their interrogation design. There is very limited work on FBG interrogators with 2D image sensors and AWGs. Tan et al. [19] used a 2D image sensor coupled with a virtually imaged phased array and an infrared camera. Jiang et al. [20] designed a fiber spectrum analyser, allowing a planar waveguide array to focus on a camera sensor without using any focusing lens. Ding et al. [21] improved upon this design of image sensor and AWG, and used a deep neural network to design an FBG interrogator

which has no moving parts — its speed limited only by the speed of image capture and computation processing. It has to be noted that: the studies listed above use the FBG spectra on an image sensor in one dimension (even where 2D sensors are used, only a rectangular section similar to 1D image is used); the studies do not predict the central wavelength of the FBG per se. They monitor the shift in central wavelength with temperature or strain. For example, in the [20] study, the authors monitored wavelength shift with temperature, and the sensitivity was 6 pm/°C for such a measurement.

In this study, we would like to take a simulation approach similar to this FBG interrogator, wherein the AWG output falls on an image sensor. In all these studies, even though an image sensor is used, the authors have not exploited the two dimensions of the image sensor. Using a true 2D image sensor, we demonstrate that a CNN can be used to design an FBG interrogator. We demonstrate a circle configuration of arranging the output of the AWG on an Image sensor and implement a CNN for measuring the central wavelength with ±1 pm accuracy. The results of the paper are based on simulation studies alone without a hardware prototype.

Description of the proposed method

FBG Interrogator Setup — Simulation

Consider a broadband laser source incident on an FBG through an optical isolator as shown in Fig. 1. The reflected light from the FBG passes through the optical isolator output channel and enters the AWG. AWG demultiplexes the wavelengths into *n* channels. Each channel has a narrow wavelength range. In the literature, the AWG output is allowed to fall linearly on an image sensor. Here, it is proposed that the AWG channels are arranged in a grid, and the output of each channel falls on a "single pixel" or a "group of non-overlapping pixels" of the image sensor, as shown in Fig. 1. In simulation studies, we can always assume that each AWG output falls on a "single pixel" or a "group of non-overlapping pixels" which can be reduced to single pixels through kernel multiplication. The AWG channels are arranged in two different patterns — grid and circle (explained in detail in the next section). The image sensor records the AWG output and sends it to a computer for processing. A CNN is designed to map the image sensor output to the central wavelength of the FBG.

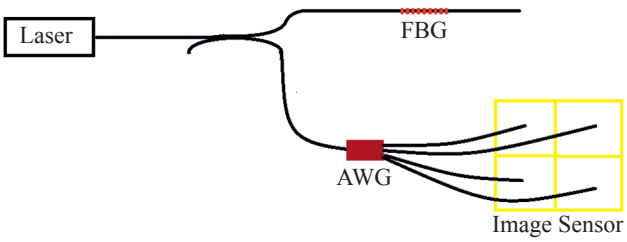

Fig. 1. Design of an FBG interrogator. The light from the laser is incident on the FBG through an optical isolator. The reflected light from the FBG through the output arm of the isolator is split into multiple channels through AWG. The AWG channels are incident on an image sensor

FBG dataset

The reflected spectra of the temperature dependent FBGs are simulated using transfer matrix [22, 23] in Python. The parameters used for the simulation are shown in Table 1. The reflection spectra of FBG at different temperatures — from 30 °C to 100 °C in steps of 5 °C — were simulated in the wavelength range of 1,558 nm to 1,560 nm. The wavelength range 1,558 to 1,560 nm was divided into 32 wavelength steps. Or, in other words, the spacing between two points in the FBG spectra is designed as 2 nm/32 ≈ 0.062 nm. Ismail et al. [24] have demonstrated AWGs with much smaller wavelength spacing of 0.01 nm. Hence, the design parameter is realistic. Additive random Gaussian noise was added with a standard deviation of 0.001. Examples of FBG spectra for temperature at 30 °C to 90 °C in steps of 20 °C are shown in Fig. 2. 4,000 data were generated for each of the 15 temperatures. 60,000 data points were generated and used for training the CNN.

AWG output configuration

The FBG reflection spectra in the range 1,558 nm to 1,560 nm are divided into 32 wavelength divisions. The FBG spectra passing through the AWG is de-multiplexed into 32 channels. If we use an array of photo detectors, or a linear array, the output of the first channel will go into the first detector, the second to the second, and so on. However, the output of the AWG goes to a 2D image sensor. To convert a 1D array to a 2D array, we can arrange the AWG channels into a grid or a circle. The two configurations are discussed below.

Grid configuration

For elucidation, consider an AWG with 36 channels. The 36 channels can be fed to a 2D image sensor whose size is 6 × 6. The first six consecutive channels of the AWG

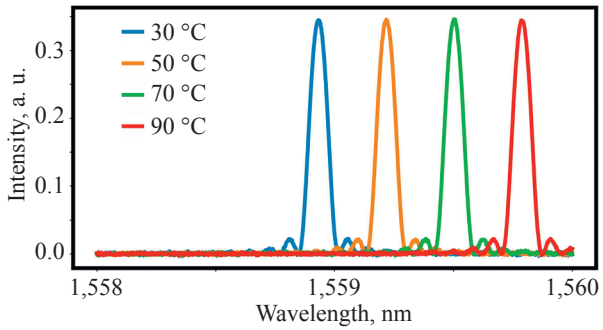

Fig. 2. Simulated Reflection spectra of FBG at different temperatures. Simulation was carried out in Python

Table 1. Parameters used for FBG Simulation

Parameter	Value
Refractive index of core n_1	1.46
Refractive index of clad n_2	1.45
Length of grating *L*, mm	10
Grating period at 30 °C λ, nm	534
Linear expansion temperature coefficient α, 1/K	0.55×10^{-6}
Thermo optic coefficient ζ, 1/K	8.6×10^{-6}

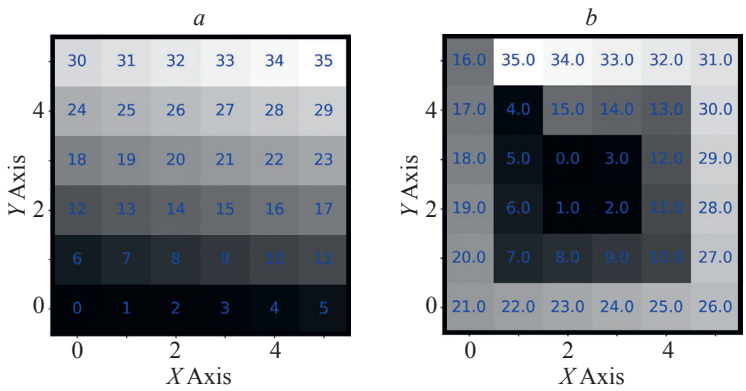

Fig. 3. AWG outputs can be coupled to an image sensor in two different configurations: grid (*a*) and circle (*b*)

go to the first row of the image sensor; the next six channels of the AWG go to the second row of the image sensor and so on. This can be seen in Fig. 3, *a*. This configuration is referred to as the grid configuration.

Circle configuration

A large image sensor circumvents interference between the output of two consecutive channels of the AWGs. By keeping the image sensor size constant and modifying only the AWG arrangement on the sensor, interference between consecutive channels can be reduced. It is proposed that instead of arranging the AWG outputs linearly in a grid, the outputs be placed in concentric circles. For an even number of pixels in the image sensor, the central four pixels are occupied by four consecutive AWG channels. The next neighbouring pixels are then filled in a circle. Fig. 3, *b* elucidates the circle configuration for a 6 × 6 image sensor coupled to 36 channels of the AWG.

Grid and circle configuration with FBG dataset

The FBG reflected spectra are divided into 32 wavelength divisions by AWG. The output of the AWG is converted into a 32 × 32 pixels image when the AWG output falls on the image sensor. The AWG can be arranged in a grid or a circle configuration. To understand the configurations, we first divide the FBG spectra (at 32 °C) into 1,024 divisions and project them onto the image sensor. The first output goes to the grid point (1, 1), the second to the grid point (1, 2), etc., on the image sensor Fig. 4, *a*. This configuration leads to clustering of all the data points along the central axis of the image sensor. If 1,024 points are projected on the image sensor in a circular pattern as discussed in the previous section, the peaks are distanced from each other Fig. 4, *b*. There is de-clustering of peaks. FBG spectra in both configurations (grid and circular) can be used to train a CNN to predict the central wavelength, as they contain all the information of the peak. However, it is not practical to design an AWG with 1,024 channels.

The FBG spectra can be divided into 128 divisions. Instead of having exactly 128 divisions, we use every 8th data point from the previous dataset where the FBG spectra was divided into 1,024 data points. This is done to maintain the same size of the image sensor. In this case, every point of the FBG spectra is projected onto the eighth point of the image sensor grid. The first output is projected on the point (1, 8), the second output on the point (1, 16), etc., on the image sensor Fig. 4, *c*. The image retains most of the points needed to train an ANN successfully. The clustering is reduced slightly, mainly due to the reduced number of points. However, if the same 128 points are projected on the grid in a circular pattern, Fig. 4, *d*, the clustering is further reduced, and a square-like pattern appears.

If we divide the FBG spectra into 32 divisions (as before, we take every 32 data points from the original dataset with 1,024 divisions), then every output of the AWG channel is projected at the thirty second point on the grid. For example, the first output is projected on the point (1, 32) in the grid, the second on (2, 32), and so on (Fig. 4, *e*). The visible points of the output are packed together. The remaining projected points are not visible because of their low intensity. However, in a circle configuration, the data points are farther away as can be seen in Fig. 4, *f*. A CNN can now be trained to extract the important features from this image sensor to predict the central wavelength of the FBG.

Grid and circle configuration with 32 spectral divisions

The aim of the paper is to design an FBG interrogator with an image sensor instead of a monochromator, using only 32 spectral divisions using the AWG, each separated by 0.0625 nm. In the following section, we show that a circular configuration with the image sensor is able to predict the central wavelength with 1 pm accuracy.

The temperature of the FBG is increased from 30 °C to 100 °C in steps of 10 °C, and the image is projected by the 32-channel AWG on the image sensor in a grid pattern. To understand the nature of the image, three representative temperatures, 30, 70 and 100 °C are shown in Fig. 5. It can be seen that in the grid pattern the peak data are clustered together and with an increase in temperature, the cluster of data moves vertically. For the same temperatures, when the AWG output is projected in a circle configuration, the radius of the circle appears to increase, and the position of the peaks is seen to move in different directions. The effect of a change in temperature is more readily visible in the image when a circle configuration is used.

Since the wavelength range of the FBG spectra is 1,558 to 1,560 nm, and there are 32 AWG channels, the wavelength spacing between each AWG channel is therefore 0.0625 nm. Such AWGs are not very difficult to fabricate. Another important effect of using sparse data is that it reduces interference from neighbouring AWG

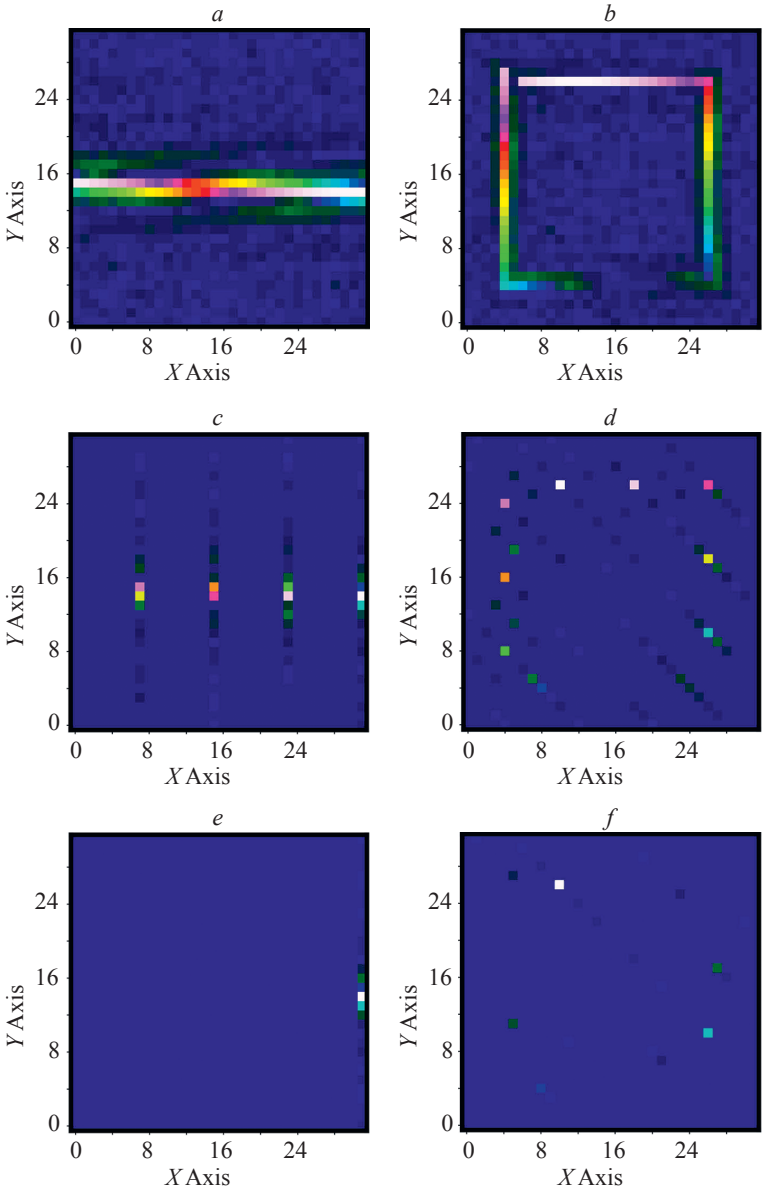

Fig. 4. Output of the image sensor when output from FBG is coupled in grid (*a*, *c*, *e*) and circle (*b*, *d*, *f*) configurations with 1,024 (*a*, *b*), 128 (*c*, *d*), and 32 AWG (*e*, *f*) channels at 30 °C

channels. The images obtained from the image sensor are trained using CNN.

CNN Architecture

The generated data set was converted to a grid and a circle configuration. A CNN model was developed using Tensor flow with Keras. 50 % of the dataset, 30,000 data, were used for training. 25 % of the dataset, 15,000 data, were used for testing and the remaining 25 %, 15,000 data, were used for validation. The input data are the output of the image sensor (Fig. 4 and 5) of size 32 × 32. The input layer was designed as a CNN layer, with a kernel size of 3 × 3 and depth 32. This layer was fed to a max-pooling layer of filter size 2 × 2. A convolution layer with depth 64 and the same kernel size and Max-pooling layer was repeated. A further convolution layer was added to increase the depth to 32. The output of this layer was flattened to a 1-Dimensional layer with 2,048 features. The output of the flattened layer was connected to a dense layer with 32 inputs. A dropout layer with ratio 0.4 was introduced. The output of this layer was fed into a dense layer with 64 inputs. The output of this layer is the final output corresponding to the wavelength peak of the FBG sensor. All CNN layers except the last output layer had ReLU as the activation function. The output layer was designed with softmax as an activation function. A simplified schematic of the CNN architecture used is shown in Fig. 6. Since CNN is being used for a regression problem, the mean square error is used as a loss function. To calculate the accuracy of the Artificial Neural Network, it was desirable to define our own accuracy function. This is because in a CNN the accuracy generally refers to identifying classes and not predicting value. So we developed our own metric. Firstly, we converted the FBG central wavelength to a range — 0 to 100. Thus, a 2 nm wavelength range was divided into 100 parts. In the converted system, it corresponds to 0.02 nm resolution. We defined a metric, wherein the

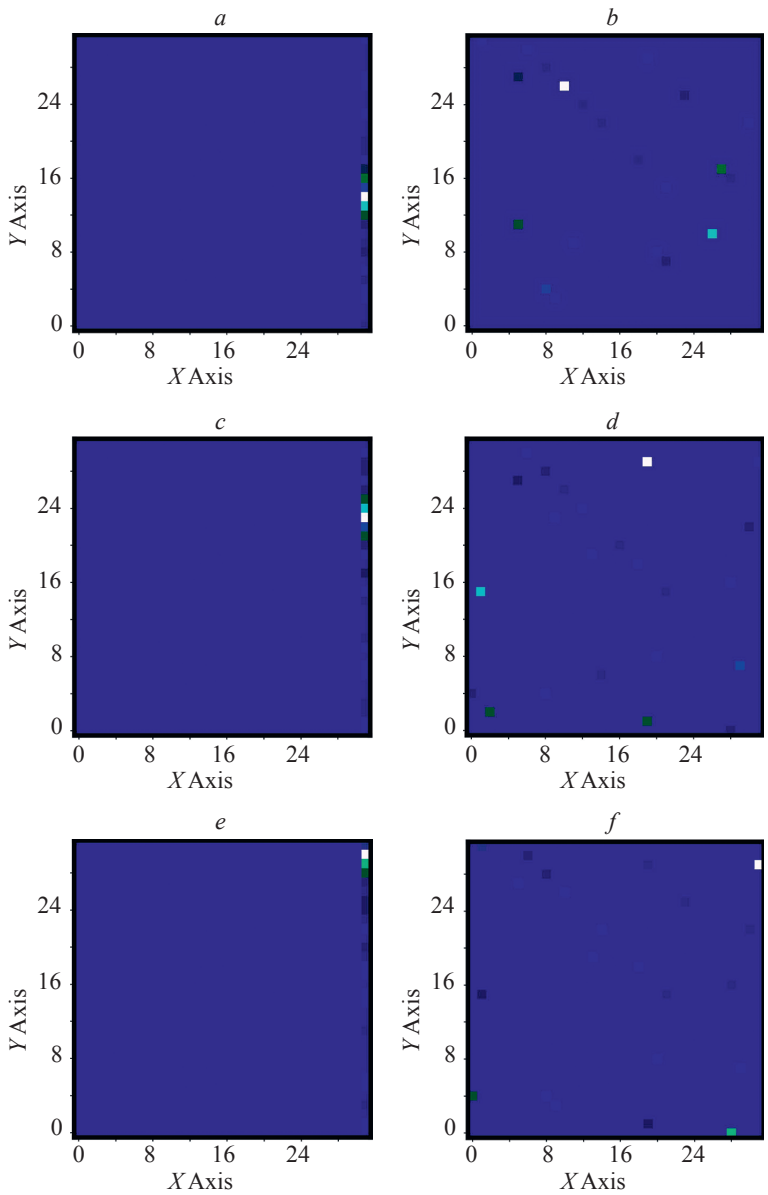

Fig. 5. Output of the image sensor when output from FBG is coupled in grid (*a*, *c*, *e*) and circle (*b*, *d*, *f*) configurations with 32 AWG channels at temperatures: 30 °C (*a*, *b*), 70 °C (*c*, *d*), and 100 °C (*e*, *f*)

predicted value was counted as accurate if the predicted value was within 0.001 nm of the actual output value. Thus, the CNN was designed to predict the output with an accuracy of ± 1 pm.

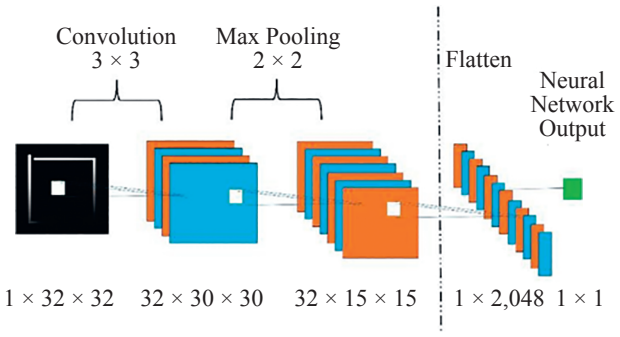

Fig. 6. Schematic of the CNN architecture

Testing of the proposed method

The CNN architecture was designed using Tensor flow with Keras was used to predict the central wavelength peak. When the FBG data with 32 AWG channels arranged in a grid was trained to 200 epochs (as shown in Fig. 7). Its loss value reached 953.382, its accuracy reached 0.0625 (6.25 %), as shown in Table 2. This was expected since all the AWG outputs on the image sensor were clustered and there was no distinction between two FBGs at different temperatures except for a translation along the vertical axis. The same CNN architecture was also used for training 32 AWG outputs arranged in a circle configuration. The loss value reached 0.2367 and the accuracy of the measurement reached 0.9868 (98.68 %), Fig. 7, *b*, Table 2. This is expected since, arranging the AWG channels in a circle makes the features distinct and easy to extract with only 32 channels, leading to an accurate prediction of peaks.

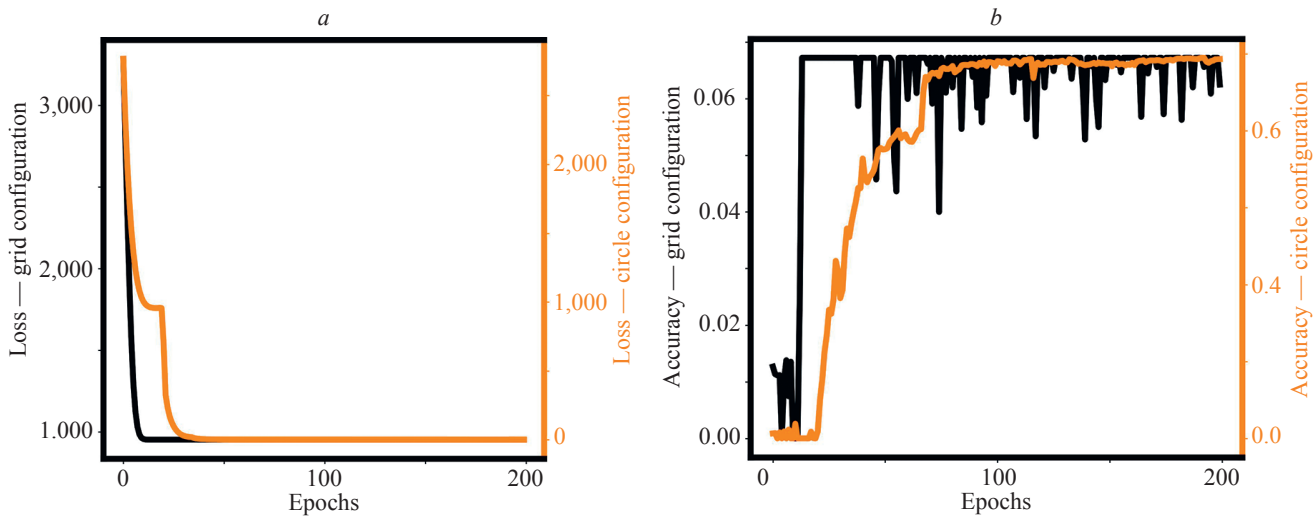

Fig. 7. CNN performance: loss (*a*) and accuracy (*b*) as a function of epochs for different configurations

To test the success of the CNN, another 15,000 untested data was used. The actual central FBG wavelength of the FBG and the predicted output using CNN for all four configurations are shown in Fig. 8 as a function of different temperatures. It can be seen that there is a complete overlap between the actual and the predicted data when the AWG output is arranged in circles. In the case of the grid, the prediction was abysmal, as could be inferred from the accuracy values as well.

A circle configuration of the AWG channels on the image sensor was used to measure the central wavelength of the FBG with 98.68 % accuracy with 1 pm resolution.

Table 2. Parameters used for FBG Simulation

Configure	Loss	Accuracy
Grid	953.3820	0.0625
Circle	0.2367	0.9868

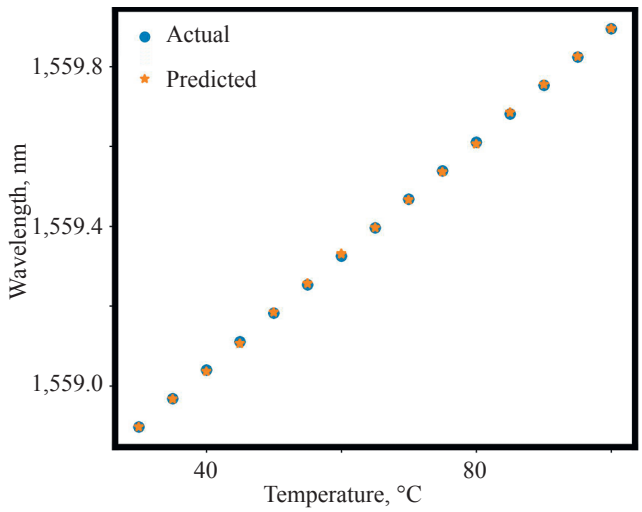

Fig. 8. CNN performance: comparing the CNN prediction with the actual data for different temperatures for circle configuration with 32 AWG channels

This compares well with the commercial FBG interrogators available. However, the limitation of the commercial interrogator is the speed of the interrogation, the limitation being imposed by the moving monochromators.

Discussion

Challenges on Practical Implementation

The results of the FBG interrogator proposed are based on simulation studies alone. If a practical implementation of the same is desired, the following challenges may be faced.

Noises

In the design of the interrogator, there are many possible sources of noises; some of them are listed here.

Laser noise: intensity (amplitude) noise and phase noise. **AWG noise:** Crosstalk, Scattering Loss and Waveguide Propagation Loss: Light loses energy during propagation due to material absorption and geometric factors.

Sensor Noise

Internal Noise: Thermal (Johnson) Noise, Shot Noise, Flicker (1/f) Noise: Low-frequency fluctuations in sensor sensitivity or gain, Readout Noise, Dark Current Noise and Fixed Pattern Noise: Constant noise from manufacturing variations between pixels/components. External Noise: Electromagnetic Interference and Environmental Noise.

We can broadly classify these noises into two types – temperature dependent or independent. Each of this type can be classified as wavelength dependent and wavelength independent.

The strength of the CNN based interrogator is that the prediction of the central wavelength depends on the training set. A large training set consisting of wavelength independent noise would be able to predict the central wavelength with the same accuracy. However, if the noise is wavelength dependent, then it could lead to a reduction in the accuracy of the measurement.

Errors (sources of wavelength Error)

Since the wavelength prediction depends on the CNN model, which in turn depends on a large training

dataset, it is possible that if the dataset does not include measurements with noises, and varied temperature range, the CNN would not be able to predict the wavelength with the same accuracy.

Projection of AWG output to Sensor

To project AWG output to the sensor, two possible experimental setups can be devised. The output fiber channels of the AWG can be tapered, so that the fiber is butt fixed to a "single pixel" or a "group of non-overlapping pixels". The second option is to make the output of the AWG fall on a transparent glass kept at a finite distance such that the projections of the output of the AWGs do not overlap on the transparent glass. An infrared camera operating in this wavelength range can be used to record the output of these AWGs. Such a set up would reduce the speed and compromise on the compactness of the interrogator system. It is also possible to design a plug and play image sensor made with avalanche photodiodes. The AWG channels are coupled to these photodiodes. Such an image sensor would increase the cost of the interrogator.

Calibration

The CNN based FBG interrogator depends on a large dataset for accurately predicting the central wavelength of the FBG. The major source of error could arise from AWG projection on the sensor. Hence two step calibrations could be implemented in a practical device.

Step 1: At ambient conditions, the image sensor output can be compared with the calibration images stored in the system, and it can be adjusted to match the same.

Step 2: A large number of ambient condition measurements are to be taken to the central wavelength of the sensor as an additive error.

Comparison

The AWG based FBG interrogation method studied in [11] offers a resonance speed ranging from 100 Hz to 500 MHz, a wavelength resolution of up to 500 pm, and supports a measurable range of 500 MHz. AWG based FBG interrogation studied in [12] operates at 100 GHz resonance speed with a wavelength ranging from 0.7 to 1 nm in Full Width at Half Maximum, and non-linearity is approximately 1.2 pm. Fully Analog Fiber Optic Monitoring System using AWG, operating at 100 kHz resonance speed, with measurement resolution range up to 800 pm, accuracy with root mean square error of 0.99 and wavelength coverage up to 50 nm was studied in [13]. FBG interrogator utilizes an AWG spectrometer with 1 pm resolution as demonstrated in [14]. AWG-based FBG wavelength interrogation using a cascaded neural network, achieving about 5.672 pm resolution with 97.7 % accuracy was discussed in [16]. Since our results are based on simulation alone, we expect the speed of the interrogation to be limited by the sensor readout speed alone, which can be improved by using avalanche photo detectors-based image sensors. Thus the speed of the interrogation will be comparable with that discussed in the literature. However, our accuracy with 1pm resolution is better than the results discussed in the literature.

Conclusion

As fiber optic sensors are becoming an integral part of structural health monitoring, it is important to design faster interrogation systems. In this paper, we have proposed and demonstrated the use of a passive AWG and an image sensor for monitoring the central wavelength of the FBG. We have also demonstrated a method to reduce the number of AWG output channels and also discussed a circle configuration that lends itself to CNN based image processing for high resolution measurement. This FBG interrogator can be designed and implemented in real-time monitoring with a high resolution.

References

1. Zhou Z., Ou J. Development of FBG sensors for structural health monitoring in civil infrastructures. *Sensing Issues in Civil Structural Health Monitoring*, 2005, pp. 197–207. https://doi.org/10.1007/1-4020-3661-2_20
2. Kahandawa G.C., Epaarachchi J., Wang H., Lau K.T. Use of FBG sensors for SHM in aerospace structures. *Photonic Sensors*, 2012, vol. 2, no. 3, pp. 203–214. https://doi.org/10.1007/s13320-012-0065-4
3. Lee J.R., Chong S.Y., Yun C.Y., Sohn H. Design of Fiber Bragg Grating acoustic sensor for structural health monitoring of nuclear power plant. *Advanced Materials Research*, 2010, vol. 123-125, pp. 859–862. https://doi.org/10.4028/www.scientific.net/amr.123-125.859
4. Riza M.A., Go Y.I., Harun S.W., Maier R.R.J. FBG sensors for environmental and biochemical applications review. *IEEE Sensors Journal*, 2020, vol. 20, no. 14, pp. 7614–7627. https://doi.org/10.1109/jsen.2020.2982446
5. Presti D.L., Massaroni C., Leitao C.S.J., Domingues M.D., Sypabekova M., Barrera D., et al. Fiber Bragg Gratings for medical applications and future challenges: a review. *IEEE Access*, 2020, vol. 8, pp. 156863–156888. https://doi.org/10.1109/ACCESS.2020.3019138
6. Kashyap R. *Fiber Bragg Gratings*. Academic press, 2009, 458 p.
7. Sengupta D. Fiber Bragg Grating sensors and interrogation systems. *Optical Fiber Sensors Advanced Techniques and Applications*, 2015, pp. 207–256.

Литература

1. Zhou Z., Ou J. Development of FBG sensors for structural health monitoring in civil infrastructures // Sensing Issues in Civil Structural Health Monitoring. 2005. P. 197–207. https://doi.org/10.1007/1-4020-3661-2_20
2. Kahandawa G.C., Epaarachchi J., Wang H., Lau K.T. Use of FBG sensors for SHM in aerospace structures // Photonic Sensors. 2012. V. 2. N 3. P. 203–214. https://doi.org/10.1007/s13320-012-0065-4
3. Lee J.R., Chong S.Y., Yun C.Y., Sohn H. Design of Fiber Bragg Grating acoustic sensor for structural health monitoring of nuclear power plant // Advanced Materials Research. 2010. V. 123-125. P. 859–862. https://doi.org/10.4028/www.scientific.net/amr.123-125.859
4. Riza M.A., Go Y.I., Harun S.W., Maier R.R.J. FBG sensors for environmental and biochemical applications review // IEEE Sensors Journal. 2020. V. 20. N 14. P. 7614–7627. https://doi.org/10.1109/jsen.2020.2982446
5. Presti D.L., Massaroni C., Leitao C.S.J., Domingues M.D., Sypabekova M., Barrera D., et al. Fiber Bragg Gratings for medical applications and future challenges: a review // IEEE Access. 2020. V. 8. P. 156863–156888. https://doi.org/10.1109/ACCESS.2020.3019138
6. Kashyap R. Fiber Bragg Gratings. Academic press, 2009. 458 p.
7. Sengupta D. Fiber Bragg Grating sensors and interrogation systems // Optical Fiber Sensors Advanced Techniques and Applications. 2015. P. 207–256.

8. Santos J.L., Ferreira L.A., Araujo F.M. Fiber Bragg Grating interrogation systems. *Fiber Bragg Grating Sensors: Recent Advancements, Industrial Applications and Market Exploitation*, 2011, pp. 78–98.
9. Cui J., Hu Y., Feng K., Li J., Tan J. FBG interrogation method with high resolution and response speed based on a reflective-matched FBG scheme. *Sensors*, 2015, vol. 15, no. 7, pp. 16516–16535. https://doi.org/10.3390/s150716516
10. Diaz C.A., Leitão C., Marques C.A., Domingues M., Alberto N., Pontes M., et al. Low-cost interrogation technique for dynamic measurements with FBG-based devices. *Sensors*, 2017, vol. 17, no. 10, pp. 2414. https://doi.org/10.3390/s17102414
11. Lei M., Zou W., Li X., Chen J. Ultrafast FBG interrogator based on time-stretch method. *IEEE Photonics Technology Letters*, 2016, vol. 28, no. 7, pp. 778-781. https://doi.org/10.1109/LPT.2015.2513903
12. Marrazzo V.R., Fienga F., Riccio M., Irace A., Breglio G. Multichannel approach for arrayed waveguide grating-based FBG interrogation systems. *Sensors*, 2021, vol. 21, no. 18, pp. 6214. https://doi.org/10.3390/s21186214
13. Niewczas P., Willshire A.J., Dziuda L., McDonald J.R. Performance analysis of the Fiber Bragg Grating interrogation system based on an arrayed waveguide grating. *IEEE Transactions on Instrumentation and Measurement*, 2004, vol. 53, no. 4, pp. 1192–1196. https://doi.org/10.1109/tim.2004.830780
14. Marrazzo V.R., Fienga F., Laezza D., Riccio M., Irace A., Buontempo S., Breglio G. Full analog fiber optic monitoring system based on arrayed waveguide grating. *Journal of Lightwave Technology*, 2021, vol. 39, no. 15, pp. 4990–4996. https://doi.org/10.1109/jlt.2021.3083061
15. Trita A., Vickers G., Mayordomo I., van Thourhout D., Vermeiren J. Design, integration, and testing of a compact FBG interrogator, based on an AWG spectrometer. *Proceedings of SPIE*, 2014, vol. 9133, pp. 91330D. https://doi.org/10.1117/12.2058107
16. Barino F.O., dos Santos A.B. LPG interrogator based on FBG array and artificial neural network. *IEEE Sensors Journal*, 2020, vol. 20, no. 23, pp. 14187–14194. https://doi.org/10.1109/JSEN.2020.3007957
17. Chen S., Yao F., Ren S., Wang G., Huang M. Cost-effective improvement of the performance of AWG-based FBG wavelength interrogation via a cascaded neural network. *Optics Express*, 2022, vol. 30, no. 5, pp. 7647–7663. https://doi.org/10.1364/oe.449004
18. Ren S., Chen S., Yang J., Wang J., Yang Q., Xue C., et al. High-efficiency FBG array sensor interrogation system via a neural network working with sparse data. *Optics Express*, 2023, vol. 31, no. 5, pp. 8937–8952. https://doi.org/10.1364/oe.479708
19. Tan Z., Ren W., Liu Z, Feng S., Chen Z. Fiber Bragg Grating sensor interrogator based on 2D imaging system. *Applied Optics*, 2014, vol. 53, no. 23, pp. 5259–5263. https://doi.org/10.1364/ao.53.005259
20. Jiang X., Yang Z., Wu L., Dang Z., Ding Z., Liu Z., et al. Fiber spectrum analyzer based on planar waveguide array aligned to a camera without lens. *Optics and Lasers in Engineering*, 2022, vol. 159, pp. 107226. https://doi.org/10.1016/j.optlaseng.2022.107226
21. Ding Z., Chang Q., Deng Z., Ke S., Jiang X., Zhang Z. FBG interrogator using a dispersive waveguide chip and a CMOS camera. *Micromachines*, 2024, vol. 15, no. 10, pp. 1206. https://doi.org/10.3390/mi15101206
22. Phing H.S., Ali J., Rahman R.A., Tahir B.A. Fiber Bragg Grating modeling, simulation and characteristics with different grating lengths. *Malaysian Journal of Fundamental and Applied Sciences*, 2007, vol. 3, no. 2, pp. 167–175. https://doi.org/10.11113/mjfas.v3n2.26
23. Ikhlef A., Hedara R., Chikh-Bled M. Uniform Fiber Bragg Grating modeling and simulation used matrix transfer method. *International Journal of Computer Science Issues*, 2012, vol. 9, no. 1, pp. 368–374.
24. Ismail N., Sun F., Sengo G., Wörhoff K., Driessen A., de Ridder R.M., Pollnau M. Improved arrayed-waveguide-grating layout avoiding systematic phase errors. *Optics Express*, 2011, vol. 19, no. 9, pp. 8781–8794. https://doi.org/10.1364/oe.19.008781

8. Santos J.L., Ferreira L.A., Araujo F.M. Fiber Bragg Grating interrogation systems // Fiber Bragg Grating Sensors: Recent Advancements, Industrial Applications and Market Exploitation. 2011. P. 78–98.
9. Cui J., Hu Y., Feng K., Li J., Tan J. FBG interrogation method with high resolution and response speed based on a reflective-matched FBG scheme // Sensors. 2015. V. 15. N 7. P. 16516–16535. https://doi.org/10.3390/s150716516
10. Diaz C.A., Leitão C., Marques C.A., Domingues M., Alberto N., Pontes M., et al. Low-cost interrogation technique for dynamic measurements with FBG-based devices // Sensors. 2017. V. 17. N 10. P. 2414. https://doi.org/10.3390/s17102414
11. Lei M., Zou W., Li X., Chen J. Ultrafast FBG interrogator based on time-stretch method // IEEE Photonics Technology Letters. 2016. V. 28. N 7. P. 778-781. https://doi.org/10.1109/LPT.2015.2513903
12. Marrazzo V.R., Fienga F., Riccio M., Irace A., Breglio G. Multichannel approach for arrayed waveguide grating-based FBG interrogation systems // Sensors. 2021. V. 21. N 18. P. 6214. https://doi.org/10.3390/s21186214
13. Niewczas P., Willshire A.J., Dziuda L., McDonald J.R. Performance analysis of the Fiber Bragg Grating interrogation system based on an arrayed waveguide grating // IEEE Transactions on Instrumentation and Measurement. 2004. V. 53. N 4. P. 1192–1196. https://doi.org/10.1109/tim.2004.830780
14. Marrazzo V.R., Fienga F., Laezza D., Riccio M., Irace A., Buontempo S., Breglio G. Full analog fiber optic monitoring system based on arrayed waveguide grating // Journal of Lightwave Technology. 2021. V. 39. N 15. P. 4990–4996. https://doi.org/10.1109/jlt.2021.3083061
15. Trita A., Vickers G., Mayordomo I., van Thourhout D., Vermeiren J. Design, integration, and testing of a compact FBG interrogator, based on an AWG spectrometer // Proceedings of SPIE. 2014. V. 9133. P. 91330D. https://doi.org/10.1117/12.2058107
16. Barino F.O., dos Santos A.B. LPG interrogator based on FBG array and artificial neural network // IEEE Sensors Journal. 2020. V. 20. N 23. P. 14187–14194. https://doi.org/10.1109/JSEN.2020.3007957
17. Chen S., Yao F., Ren S., Wang G., Huang M. Cost-effective improvement of the performance of AWG-based FBG wavelength interrogation via a cascaded neural network // Optics Express. 2022. V. 30. N 5. P. 7647–7663. https://doi.org/10.1364/oe.449004
18. Ren S., Chen S., Yang J., Wang J., Yang Q., Xue C., et al. High-efficiency FBG array sensor interrogation system via a neural network working with sparse data // Optics Express. 2023. V. 31. N 5. P. 8937–8952. https://doi.org/10.1364/oe.479708
19. Tan Z., Ren W., Liu Z, Feng S., Chen Z. Fiber Bragg Grating sensor interrogator based on 2D imaging system // Applied Optics. 2014. V. 53. N 23. P. 5259–5263. https://doi.org/10.1364/ao.53.005259
20. Jiang X., Yang Z., Wu L., Dang Z., Ding Z., Liu Z., et al. Fiber spectrum analyzer based on planar waveguide array aligned to a camera without lens // Optics and Lasers in Engineering. 2022. V. 159. P. 107226. https://doi.org/10.1016/j.optlaseng.2022.107226
21. Ding Z., Chang Q., Deng Z., Ke S., Jiang X., Zhang Z. FBG interrogator using a dispersive waveguide chip and a CMOS camera // Micromachines. 2024. V. 15. N 10. P. 1206. https://doi.org/10.3390/mi15101206
22. Phing H.S., Ali J., Rahman R.A., Tahir B.A. Fiber Bragg Grating modeling, simulation and characteristics with different grating lengths // Malaysian Journal of Fundamental and Applied Sciences. 2007. V. 3. N 2. P. 167–175. https://doi.org/10.11113/mjfas.v3n2.26
23. Ikhlef A., Hedara R., Chikh-Bled M. Uniform Fiber Bragg Grating modeling and simulation used matrix transfer method // International Journal of Computer Science Issues. 2012. V. 9. N 1. P. 368–374.
24. Ismail N., Sun F., Sengo G., Wörhoff K., Driessen A., de Ridder R.M., Pollnau M. Improved arrayed-waveguide-grating layout avoiding systematic phase errors // Optics Express. 2011. V. 19. N 9. P. 8781–8794. https://doi.org/10.1364/oe.19.008781

Authors

Sugumar Venkatesan — M.Sc., (PhD), Researcher, Ramakrishna Mission Vivekananda College (Affiliated to University of Madras), Chennai, 600004, India, sc 58959894800, https://orcid.org/0009-0009-2243-1251, sugumartamizh@gmail.com

Авторы

Венкатесан Сугумар — магистр, (PhD), исследователь, Рамакришна Мишен Вивекананда Колледж (филиал Мадрасского университета), Ченнай, 600004, Индия, sc 58959894800, https://orcid.org/0009-0009-2243-1251, sugumartamizh@gmail.com

Subashini Ponnusamy — MCA, (PhD), Researcher, Government Arts College, Chennai, 600035, India, https://orcid.org/0009-0002-0501-7621, shiniponnusamy18@gmail.com

Pandian Chelliah — PhD, Associate Professor, Ramakrishna Mission Vivekananda College (Affiliated to University of Madras), Chennai, 600004, India, SC 36439116600, https://orcid.org/0000-0002-8847-4520, cpandian@rkmvc.ac.in

Поннусами Субашини — магистр компьютерных наук, (PhD), исследователь, Государственный мужской колледж искусств, Ченнай, 600035, Индия, https://orcid.org/0009-0002-0501-7621, shiniponnusamy18@gmail.com

Челия Пандиан — PhD, доцент, Рамакришна Мишен Вивекананда Колледж (филиал Мадрасского университета), Ченнай, 600004, Индия, SC 36439116600, https://orcid.org/0000-0002-8847-4520, cpandian@rkmvc.ac.in

Received 31.07.2025
Approved after reviewing 21.10.2025
Accepted 22.11.2025

Статья поступила в редакцию 31.07.2025
Одобрена после рецензирования 21.10.2025
Принята к печати 22.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1024-1032
УДК 681.78; 621.3; 621.396

Новый подход к измерениям и расчетам спектральных характеристик селективных элементов оптико-электронных систем

Александр Владимирович Ильинский[1], Надежда Константиновна Мальцева[2]✉, Евгений Олегович Раскин[3]

[1,2,3] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] avilinskii@itmo.ru, https://orcid.org/0009-0005-6413-6411
[2] nkmaltceva@itmo.ru✉, https://orcid.org/0009-0004-4515-0261
[3] raskin@itmo.ru, https://orcid.org/0000-0001-7830-6841

Аннотация
Введение. Исследованы вопросы энергетической калибровки и точных измерений спектральных характеристик элементов оптико-электронных систем, предназначенных для анализа излучения удаленных объектов. Современные методы калибровки требуют учета переменных параметров окружающей среды, особенно в инфракрасном диапазоне спектра. Это приводит к значительным погрешностям измерений и создает затруднения при проведении наземных испытаний. Сложность компенсации атмосферных эффектов, низкая точность и трудоемкость их учета при использовании классических методик приводят к значительным ошибкам и снижают качество полученных данных. Предложенный подход базируется на использовании узкоспектральных излучателей, расположенных непосредственно перед входным иллюминатором прибора, и исключает необходимость учета спектрального пропускания атмосферы. **Метод.** Вместо традиционных методов, основанных на использовании эталонных источников типа «черное тело» или фотоприемников, требующих учета пропускания воздушного тракта, предлагается применять в качестве калибровочного излучателя ряд узкоспектральных (спектрозональных) потоков излучения, воздействующих на испытуемую оптико-электронную систему непосредственно в плоскости ее входного иллюминатора, что позволяет получать спектральные характеристики исследуемых элементов напрямую, минуя этап определения пропускания воздушного промежутка. Такой подход уменьшает неопределенность измерений и позволяет проводить калибровку без привлечения сложных компенсирующих измерений. **Основные результаты.** Проведенные эксперименты подтвердили, что предложенный метод снижает неопределенность измерений минимум на два порядка по сравнению с традиционными подходами. Эффективность метода продемонстрирована на конкретных примерах, которые показали преимущества метода при исследовании источников света и измерениях спектральной чувствительности приборов. **Обсуждение.** Новизна предложенного подхода состоит в устранении основного источника неопределенности (учета спектрального пропускания атмосферы), что кардинально улучшает метрологические показатели калибровки. Предложенный метод эффективен практически во всех прикладных ситуациях и обеспечивает существенное повышение точности измерений. По сравнению с классическими решениями, метод проще реализуется технически и дает значительно лучшие результаты в области дистанционной оптической разведки, медицины, сельского хозяйства и экологии.

Ключевые слова
спектрофотометрия, оптико-электронная система, спектральные характеристики, энергетическая калибровка, инфракрасный диапазон, калибровка, излучение удаленных объектов, черное тело, системы обнаружения, спектральный анализ

Ссылка для цитирования: Ильинский А.В., Мальцева Н.К., Раскин Е.О. Новый подход к измерениям и расчетам спектральных характеристик селективных элементов оптико-электронных систем // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1024–1032. doi: 10.17586/2226-1494-2025-25-6-1024-1032

© Ильинский А.В., Мальцева Н.К., Раскин Е.О., 2025

Innovative method for measuring and calculating spectral characteristics of selective elements in optoelectronic systems

Alexander V. Ilinskii[1], Nadezhda K. Maltseva[2]✉, Eugene O. Raskin[3]

[1,2,3] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] avilinskii@itmo.ru, https://orcid.org/0009-0005-6413-6411
[2] nkmaltceva@itmo.ru✉, https://orcid.org/0009-0004-4515-0261
[3] raskin@itmo.ru, https://orcid.org/0000-0001-7830-6841

Abstract
The paper deals with the issues of energy calibration and precise measurements of the spectral characteristics of optoelectronic systems elements designed for the analysis of radiation from remote objects. Modern calibration methods require the consideration of environmental variables, particularly in the infrared spectrum range. This leads to significant measurement errors and creates difficulties during ground-based tests. The complication of compensating for atmospheric effects, the low accuracy, and the labor-intensive nature of their consideration when using traditional techniques result in significant errors and reduce the quality of the obtained data. The proposed approach is based on the use of narrow-spectrum emitters located directly in front of the instrument input port, eliminating the need to account for atmospheric spectral transmission. Instead of traditional methods based on the use of standard sources, such as "black bodies" or photodetectors that have to take into account the transmission of the air gap, it is proposed to use a series of narrow-spectrum (spectral-zone) radiation fluxes as a calibration emitter that affects the testing optical-electronic system directly in its input window plane, which allows to provide the direct measurement of the spectral characteristics of the elements under study, bypassing the step of determining the transmission of the air gap. This approach reduces the measurements uncertainty and allows carrying out calibration without the need for complex compensating measurements. The conducted experiments confirmed that the proposed method reduces the measurement uncertainty by at least two orders of magnitude compared to traditional approaches. The effectiveness of the method is demonstrated through specific examples that show the advantages of the method in the study of light sources and measurements of the spectral sensitivity of instruments. The novelty of the proposed approach is the elimination of the main source of uncertainty (accounting for the spectral transmission of the atmosphere), which significantly improves the metrological performance of calibration. The proposed method is effective in almost all application situations and provides a significant increase in measurement accuracy. Compared to classical solutions, this method is easier to implement technically and it provides significantly better results in the fields of remote optical reconnaissance, medicine, agriculture, and ecology.

Keywords
spectrophotometry, optoelectronic system, spectral characteristics, energy calibration, infrared range, remote object radiation, black body, detection systems, spectral analysis

For citation: Ilinskii A.V., Maltseva N.K., Raskin E.O. Innovative method for measuring and calculating spectral characteristics of selective elements in optoelectronic systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1024–1032 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1024-1032

Введение

Энергетическая калибровка и точные измерения спектральных характеристик элементов оптико-электронных систем (ОЭС) являются фундаментальной задачей спектральной фотометрии и радиометрии [1–6]. Точность этих процедур напрямую определяет достоверность данных, получаемых ОЭС самого различного назначения: от дистанционного зондирования Земли (ДЗЗ) и мониторинга природных катастроф [7–9] до систем промышленного контроля [10, 11], медицинской диагностики [12], сельскохозяйственного и экологического мониторинга. Особенно критичны требования к калибровке в инфракрасном (ИК) диапазоне, где присутствуют полосы сильного поглощения излучения атмосферными газами [13, 14], что существенно усложняет как наземные испытания аппаратуры, предназначенной для работы в космосе или атмосфере, так и интерпретацию получаемых данных.

Существующие методы энергетической калибровки и измерения спектральных характеристик селективных элементов (фотоприёмных устройств, фильтров, детекторов) ОЭС [15, 16] часто требуют учета или строгого контроля параметров окружающей среды, в первую очередь — спектрального пропускания воздушного тракта между эталонным источником излучения [17] (таким как модель абсолютно черного тела [18]) и калибруемым прибором [4, 5, 8, 14, 19]. Эта необходимость обусловлена тем, что традиционные методики предполагают измерение сигнала прибора в ответ на известное падающее на его вход излучение [2, 3, 5, 8, 13]. В ИК диапазоне, особенно в областях сильного поглощения (например, полосы водяного пара и углекислого газа), учет пропускания атмосферы становится сложной, трудоемкой задачей и зачастую основным источником неопределенности метрологической задачи [8, 13, 14]. Точное измерение или расчет спектрального пропускания реального воздушного тракта требуют сложного дополнительного оборудования (высокоточные спектрометры, контроль параметров атмосферы) и вносят значительные погрешности [5, 19]. Кроме того, существующие методы калибровки многоканальных и сканирующих систем [4, 9, 10] могут быть трудоемкими и требующими сложных эталонных установок, обеспечивающих равномерное поле освещения в широком спектральном и динамическом диапазонах [4, 5, 8, 17, 18].

Однако, несмотря на достигнутый прогресс [1, 7, 9], проблема необходимости учета спектрального пропу-

скания атмосферы в тракте между эталонным источником и калибруемым прибором при высокоточных измерениях в ИК диапазоне, особенно в полосах поглощения, остается актуальной [13, 14]. Существующие методики либо требуют сложных компенсационных измерений, либо вынужденно ограничиваются спектральными диапазонами с высокой прозрачностью атмосферы, либо вносят значительную неопределенность. В настоящей работе предлагается подход, позволяющий упростить процедуру калибровки ОЭС и повысить точность спектральных измерений за счет устранении необходимости учитывать спектральное пропускание атмосферы, что традиционно создает значительную неопределенность и сложность в процессе калибровки, особенно в ИК диапазоне.

Предлагаемое решение

Предлагаемый принципиально новый подход основан на использовании в качестве калибровочного излучателя ряда узкоспектральных (спектрозональных) потоков излучения, воздействующих на испытуемую ОЭС непосредственно в плоскости ее входного иллюминатора. Это позволяет получать спектральные характеристики исследуемых элементов напрямую, минуя этап определения пропускания воздушного промежутка.

Под спектральным составом оптического излучения понимается ряд потоков излучения в соответствующем ряде спектральных зон. Установка для получения набора узкоспектральных зон описана в работе [14]. Эти зоны могут быть заданы посредством системы узких фильтрующих функций, разделяющих весь спектр на составляющие части. При этом одновременно можно достичь двух целей: с одной стороны калибруется исследуемая ОЭС, с другой — определяются спектральные характеристики и спектрозональные параметры пропускания оптики самого спектроанализирующего прибора (САП).

Формальная постановка задачи. Рассмотрены методы, расчеты и измерения спектральных характеристик элементов ОЭС, в основе которых лежит следующее выражение [14]:

$$U = \int_{\lambda_н}^{\lambda_в} \Phi_\lambda(\lambda)\tau(\lambda)S(\lambda)d\lambda, \qquad (1)$$

где $\Phi_\lambda(\lambda)$ — спектральная плотность потока излучения источника оптического сигнала в пределах выделяемого диапазона длин волн; U — сигнал приемника оптического излучения, обладающего спектральной чувствительностью $S(\lambda)$ в пределах этого же диапазона длин волн; $\tau(\lambda)$ — спектральный коэффициент пропускания оптической трассы между источником и приемником; $(\lambda_н, \lambda_в)$ — минимальный спектральный промежуток между длинами волн $\lambda_н$ и $\lambda_в$, за пределами которого подынтегральное выражение $\Phi_\lambda(\lambda)\tau(\lambda)S(\lambda)$ равно нулю.

Базовые соотношения. Выражение (1) описывает сигнал на выходе некоторой ОЭС, на вход которой поступает оптическое излучение источника со спектральной плотностью $\Phi_\lambda(\lambda)$. Если величина $\Phi_\lambda(\lambda)$ в пределах спектрального промежутка $(\lambda_н, \lambda_в)$ постоянна и равна спектральной плотности потока излучения на длине волны $\lambda_{ср}$ $\Phi_\lambda(\lambda_{ср})$, при этом $\lambda_{ср} = \frac{\lambda_н + \lambda_в}{2}$, то появляется возможность корректного измерения. В этом случае соотношение (1) может быть преобразовано в следующий вид:

$$\Phi_\lambda(\lambda_{ср}) = \frac{U}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda}. \qquad (2)$$

Здесь полезно заметить, что интеграл, стоящий в знаменателе дроби (2), по сути, представляет собой некоторую чувствительность данной ОЭС именно к спектральной плотности потока излучения на своем входе, которую можно откалибровать. Для этого на вход рассматриваемой ОЭС следует подать излучение опорного источника оптического излучения с постоянной и известной спектральной плотностью $\Phi_\lambda^0(\lambda_{ср})$ в пределах спектрального промежутка $(\lambda_н, \lambda_в)$. Соответствующий сигнал U^0 определится формулой (1) и составит:

$$U^0 = \Phi_\lambda^0(\lambda_{ср}) \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda. \qquad (3)$$

Из (2) и (3) получим формулу для расчета искомой спектральной плотности падающего на вход ОЭС потока излучения:

$$\Phi_\lambda(\lambda_{ср}) = \Phi_\lambda^0(\lambda_{ср}) \frac{U}{U^0}. \qquad (4)$$

Классический подход и его ограничения. Выражения (1)–(4) описывают известный метод измерения спектральной характеристики потоков оптического излучения с использованием черного тела [18]. Однако требование постоянства спектральной характеристики источника излучения в пределах спектрального промежутка $(\lambda_н, \lambda_в)$ не позволяет его расширить, например, с целью увеличения отношения сигнала к шуму при подобных измерениях. Это связано с тем, что непостоянство (флуктуации) спектральных характеристик источников излучения как исследуемого, так и опорного, в пределах рабочего промежутка $(\lambda_н, \lambda_в)$ влечет за собой появление соответствующих методических неопределенностей результата измерения, которыми невозможно пренебречь.

С целью исключения указанных неопределенностей предлагается новый подход к данным измерениям. Для его описания обратимся к выражению (1). Допустим, что измеряемая спектральная характеристика $\Phi_\lambda(\lambda)$ на промежутке $(\lambda_н, \lambda_в)$ — переменная, но при этом она изменяется линейно:

$$\Phi_\lambda(\lambda) = a\lambda + b, \text{ где } \lambda_н < \lambda < \lambda_в, \qquad (5)$$

где $a = \frac{\Phi_\lambda(\lambda_в) - \Phi_\lambda(\lambda_н)}{\lambda_в - \lambda_н}$; $b = \frac{\Phi_\lambda(\lambda_н)\lambda_в - \Phi_\lambda(\lambda_в)\lambda_н}{\lambda_в - \lambda_н}$.

Подставив (5) в (1), получим:

$$U = \int_{\lambda_н}^{\lambda_в} (a\lambda + b)\tau(\lambda)S(\lambda)d\lambda = a\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)S(\lambda)d\lambda + b\int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda = \tag{6}$$

$$= \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda \left[a\frac{\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)\Phi_\lambda(\lambda)d\lambda}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda} + b \right] = S\Phi_\lambda(\lambda_1),$$

где $S = \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda$; $\lambda_1 = \frac{\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)\Phi_\lambda(\lambda)d\lambda}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda}$.

Предлагаемый подход: учет линейной зависимости. Предположим, что спектральная характеристика опорного источника $\Phi_\lambda^0(\lambda)$ на промежутке ($\lambda_н$, $\lambda_в$) изменяется линейно. Исходя из этого, применим формулу (6) для вычисления сигнала опорного источника:

$$U^0 = S\Phi_\lambda^0(\lambda_1). \tag{7}$$

Из (6) и (7) получим формулу для расчета искомой спектральной плотности:

$$\Phi_\lambda(\lambda_1) = \Phi_\lambda^0(\lambda_1)\frac{U}{U^0},$$

где

$$\lambda_1 = \frac{\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)\Phi_\lambda(\lambda)d\lambda}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda}, \tag{8}$$

аналогично выражению (6).

В результате требование постоянства спектральной характеристики источников в пределах спектрального промежутка ($\lambda_н$, $\lambda_в$) снято. Теперь появилась необходимость расчетного или экспериментального определения величины λ_1, которую называют спектральной координатой [14]. В данном случае функция $\tau(\lambda)S(\lambda)$ является спектральной характеристикой чувствительности ОЭС, используемой для данных измерений.

Покажем принадлежность λ_1 спектральному промежутку ($\lambda_н$, $\lambda_в$). Для этого обратимся к интегралу (1) и применим к нему теорему о среднем [20]:

$$U = \int_{\lambda_н}^{\lambda_в} \Phi_\lambda(\lambda)\tau(\lambda)S(\lambda)d\lambda = S\Phi_\lambda(\vartheta), \tag{9}$$

где ϑ принадлежит промежутку ($\lambda_н$, $\lambda_в$) и $S = \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda$.

Формула (9) справедлива в самом общем случае.

Сопоставляя (6) и (9), нетрудно видеть, что $\Phi_\lambda(\vartheta) = \Phi_\lambda(\lambda_1)$. Для линейной функции $\Phi_\lambda(\lambda)$ из равенства ее значений следует равенство соответствующих значений аргумента, т. е. $\lambda_1 = \vartheta$ и принадлежит промежутку ($\lambda_н$, $\lambda_в$), что и требовалось показать.

Отметим, что λ_1 является координатой центра масс плоской фигуры, ограниченной сверху графиком функции $\tau(\lambda)S(\lambda)$, снизу — осью абсцисс на промежутке ($\lambda_н$, $\lambda_в$). Если график данной функции симметричен относительно середины данного промежутка, то:

$$\lambda_1 = \frac{1}{2}(\lambda_н + \lambda_в). \tag{10}$$

Применив формулы (8) и (4), можно провести измерение спектральной плотности потока излучения в двух вариантах без методических неопределенностей при соблюдении соответствующих условий. Для первого варианта это постоянство спектральных характеристик измеряемого и опорного источников в рабочем спектральном промежутке, а для второго — это их линейность на том же промежутке.

Однако эти условия могут не соблюдаться. Потому необходимо провести анализ влияния данных факторов на методическую неопределенность результата измерения для обоих вариантов.

Пусть для первого варианта относительные изменения спектральных характеристик измеряемого и опорного источников не превышают следующих величин соответственно:

$$\frac{\Delta\Phi_\lambda(\lambda)}{\Phi_\lambda(\lambda_{ср})} \le \delta; \frac{\Delta\Phi_\lambda^0(\lambda)}{\Phi_\lambda^0(\lambda_{ср})} \le \delta^0, \tag{11}$$

где λ принадлежит промежутку ($\lambda_н$, $\lambda_в$).

Уровень сигнала ОЭС при воздействии на нее опорного источника получим аналогично (9). Для этого обратимся к интегралу (1) и применим к нему теорему о среднем:

$$U^0 = \int_{\lambda_н}^{\lambda_в} \Phi_\lambda^0(\lambda)\tau(\lambda)S(\lambda)d\lambda = S\Phi_\lambda^0(\vartheta^0), \tag{12}$$

где ϑ^0 принадлежит промежутку ($\lambda_н$, $\lambda_в$) и $S = \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda$.

Из (9) и (12) получим следующее соотношение:

$$\Phi_\lambda(\vartheta) = \Phi_\lambda^0(\vartheta^0)\frac{U}{U^0}, \tag{13}$$

где ϑ и ϑ^0 принадлежат промежутку ($\lambda_н$, $\lambda_в$).

Полученная формула (13) является точной, но непригодной для расчета по причине неизвестных ϑ и ϑ^0. Для вычислений применим формулу (4) с оценкой возникших неопределенностей, для которых получим необходимые соотношения. Для этого введем следующие обозначения:

$$\Delta\Phi = \Phi_\lambda(\vartheta) - \Phi_\lambda(\lambda_{ср}); \Delta\Phi^0 = \Phi_\lambda^0(\vartheta^0) - \Phi_\lambda^0(\lambda_{ср}). \tag{14}$$

Используя обозначения (14), придадим соотношению (13) следующий вид:

$$\Phi_\lambda(\lambda_{ср}) = \Phi_\lambda^0(\lambda_{ср})\frac{U}{U^0} + \Delta\Phi^0\frac{U}{U^0} - \Delta\Phi. \tag{15}$$

Второе и третье слагаемые в правой части равенства (15) являются фактическими отклонениями измеряемой величины $\Phi_\lambda(\lambda_{ср})$ от результата измерения $\Phi_\lambda^0(\lambda_{ср})\frac{U}{U^0}$, вызванными непостоянством спектральной характеристики опорного и измеряемого источников.

Из (11) следует, что:

$$\frac{|\Delta\Phi|}{\Phi_\lambda(\lambda_{ср})} \le \delta, \qquad (16)$$

$$\frac{|\Delta\Phi^0|}{\Phi_\lambda^0(\lambda_{ср})} \le \delta^0. \qquad (17)$$

Умножив числитель и знаменатель дроби в левой части (17) на положительную величину $\frac{U}{U^0}$, получим:

$$\frac{|\Delta\Phi^0|\frac{U}{U^0}}{\Phi_\lambda^0(\lambda_{ср})\frac{U}{U^0}} \le \delta^0. \qquad (18)$$

Из (15) следует, что $\Phi_\lambda(\lambda_{ср}) \cong \Phi_\lambda^0(\lambda_{ср})\frac{U}{U^0}$, поэтому выражение (18) приобретает окончательный вид:

$$\frac{|\Delta\Phi^0|\frac{U}{U^0}}{\Phi_\lambda(\lambda_{ср})} \le \delta^0. \qquad (19)$$

Полученные формулы (16) и (19) показывают, что относительные неопределенности результата измерения спектральной плотности $\Phi_\lambda(\lambda_{ср})$ не превышают величин δ и δ^0, являющихся, в соответствии с (11), максимальными относительными изменениями спектральных характеристик измеряемого и опорного источников, на промежутке $(\lambda_н, \lambda_в)$.

В результате, на основании предложенного подхода, получены соотношения для оценки методических неопределенностей по первому варианту измерений.

Анализ неопределенностей. Для оценки методических неопределенностей по второму варианту необходимо определиться с оценкой нелинейности спектральных характеристик обоих источников на промежутке $(\lambda_н, \lambda_в)$.

На рис. 1 показана мера нелинейности спектральной характеристики потока излучения $\Phi_\lambda(\lambda)$ на промежутке $(\lambda_н, \lambda_в)$, за которую принята величина h, представляющая собой расстояние в направлении оси ординат между двумя параллельными прямыми, зажимающими график спектральной характеристики $\Phi_\lambda(\lambda)$ на промежутке $(\lambda_н, \lambda_в)$. Прямой линии, находящейся сверху от графика $\Phi_\lambda(\lambda)$, соответствует функция $\Phi^В(\lambda)$, а линии, находящейся снизу от $\Phi_\lambda(\lambda)$ соответствует линейная функция $\Phi^Н(\lambda)$.

Из рис. 1 следует двойное неравенство:

$$\Phi^Н(\lambda) \le \Phi_\lambda(\lambda) \le \Phi^В(\lambda), \qquad (20)$$

которое справедливо для всех λ, принадлежащих промежутку $(\lambda_н, \lambda_в)$.

Из выражения (20) следует двойное интегральное неравенство:

$$\int_{\lambda_н}^{\lambda_в} \Phi^Н(\lambda)\tau(\lambda)S(\lambda)d\lambda \le \int_{\lambda_н}^{\lambda_в} \Phi_\lambda(\lambda)\tau(\lambda)S(\lambda)d\lambda \le \le \int_{\lambda_н}^{\lambda_в} \Phi^В(\lambda)\tau(\lambda)S(\lambda)d\lambda. \qquad (21)$$

Функции $\Phi^Н(\lambda)$ и $\Phi^В(\lambda)$ — линейные, поэтому применим формулу (6) к крайним интегралам неравенства (21). К среднему интегралу неравенства (21) применим теорему о среднем (9) и разделим его на величину:

$$S = \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda.$$

Тогда получим:

$$\Phi^Н(\lambda_1) \le \Phi_\lambda(\vartheta) \le \Phi^В(\lambda_1), \qquad (22)$$

где $\lambda_1 = \frac{\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)\Phi_\lambda(\lambda)d\lambda}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda}$ и ϑ принадлежат промежутку $(\lambda_н, \lambda_в)$.

Из неравенства (20) следует:

$$\Phi^Н(\lambda_1) \le \Phi_\lambda(\lambda_1) \le \Phi^В(\lambda_1). \qquad (23)$$

Умножив неравенство (23) на минус единицу и сложив полученный результат с (22), запишем:

$$|\Phi_\lambda(\vartheta) - \Phi_\lambda(\lambda_1)| \le \Phi^В(\lambda_1) - \Phi^Н(\lambda_1) = h. \qquad (24)$$

На рис. 2 показано графическое пояснение к выводу формулы (24).

Точки B и F ограничивают отрезок, содержащий измеряемую величину $\Phi_\lambda(\lambda_1)$ и результат измерения $\Phi_\lambda(\vartheta)$. Точка B имеет координаты λ_1 и $\Phi^В(\lambda_1)$, а точка F — координаты λ_1 и $\Phi^Н(\lambda_1)$.

Поясним важность соотношения (24) на следующем примере.

Пусть имеем ОЭС с откалиброванными параметрами:

$$S = \int_{\lambda_н}^{\lambda_в} \tau(\lambda)S(\lambda)d\lambda;\ \lambda_1 = \frac{\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)\Phi_\lambda(\lambda)d\lambda}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda}.$$

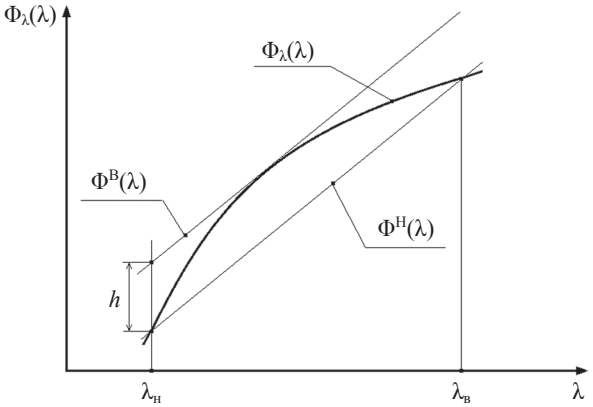

Рис. 1. Графическое построение, поясняющее принятую в настоящей работе меру h нелинейности характеристики $\Phi_\lambda(\lambda)$ на промежутке $(\lambda_н, \lambda_в)$

Fig. 1. Graphical illustration to describe the nonlinearity h of the characteristics $\Phi_\lambda(\lambda)$ at the interval $(\lambda_н, \lambda_в)$

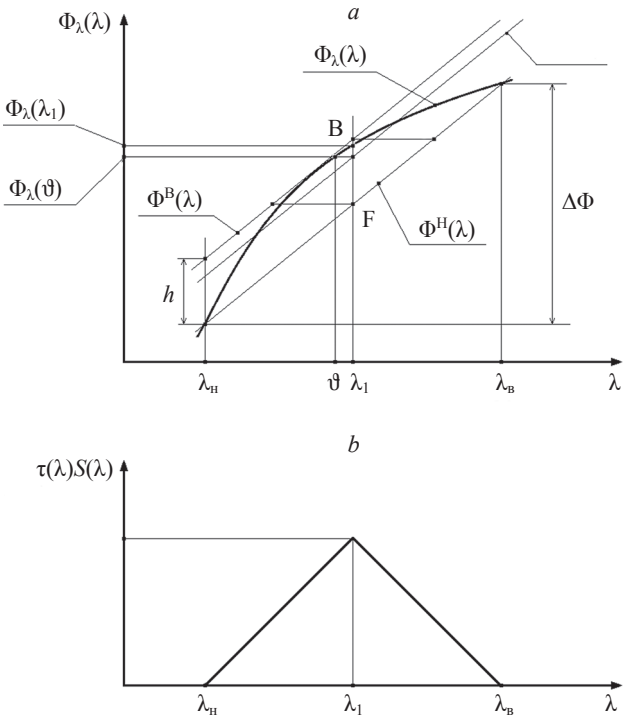

Рис. 2. Графическое пояснение к выводу формулы (24) на примере измерения спектральной плотности потока излучения $\Phi_\lambda(\lambda_1)$ на длине волны λ_1 в случае плавно нелинейно изменяющейся спектральной характеристики потока излучения: график спектрального распределения потока излучения (*a*), график спектральной характеристики чувствительности оптико-электронной системы (*b*)

Fig. 2. Graphical illustration of formula (24) derivation based on the example of optical flux spectral density $\Phi_\lambda(\lambda_1)$ measurement at the wavelength λ_1 in case of continuous curve nonlinearly spectral characteristic $\Phi_\lambda(\lambda)$ (*a*) and certain optics-electronic system spectral sensitivity characteristic $\tau(\lambda)S(\lambda)$ (*b*)

Спектральная характеристика чувствительности данной ОЭС имеет форму равнобедренного треугольника и приведена на рис. 2, *b*. Такие ОЭС в работах [6, 21] называют спектрометрами, с помощью которых можно измерить спектральную плотность $\Phi_\lambda(\lambda_1)$ потока излучения некоторого источника. В этом случае λ_1 рассчитывается в соответствии с формулой (10).

Если спектральная характеристика источника на $\Phi_\lambda(\lambda)$ на промежутке $(\lambda_н, \lambda_в)$ не линейна, тогда в соответствии с формулой (9) имеем $\Phi_\lambda(\vartheta) = \dfrac{U}{S}$ — результат измерения величины $\Phi_\lambda(\lambda_1)$.

При этом формула (24) гарантирует, что $|\Phi_\lambda(\vartheta) - \Phi_\lambda(\lambda_1)| \le h$.

Таким образом, методическая неопределенность результата измерения не превышает меру нелинейности измеряемой спектральной характеристики на $(\lambda_н, \lambda_в)$.

На рис. 2, *a* представлен пример плавной, монотонно возрастающей характеристики $\Phi_\lambda(\lambda)$, для которой разница $|\Phi_\lambda(\vartheta) - \Phi_\lambda(\lambda_1)|$ гораздо меньше h.

Изменим характеристику $\Phi_\lambda(\lambda)$ в пределах отрезка BF (рис. 3). Уровень $\Phi_\lambda(\vartheta)$ практически не изменился. Однако разность $\Phi_\lambda(\vartheta) - \Phi_\lambda(\lambda_1)$ поменяла знак и значительно увеличилась по модулю. При этом ϑ оказалась больше λ_1. Видно, что, прижимая основной график

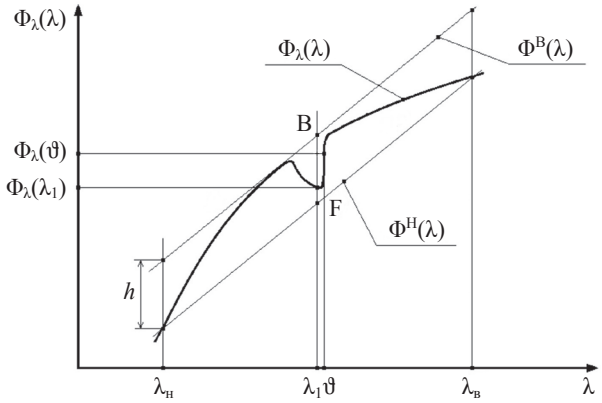

Рис. 3. Графическое пояснение вывода формулы (24) в случае измерения спектральной характеристики потока излучения, сильно изменяющейся в окрестности длины волны λ_1

Fig. 3. Graphical illustration of formula (24) derivation based on example of measurement of optical flux highly changeable spectral characteristic at the wavelength λ_1

функции $\Phi_\lambda(\lambda)$ к прямой $\Phi^B(\lambda)$, уровень $\Phi_\lambda(\vartheta)$ приближается к точке B. С другой стороны, одновременно сужая и вытягивая средний выступ графика $\Phi_\lambda(\lambda)$ к точке F, уровень $\Phi_\lambda(\lambda_1)$ также приближается к точке F. В результате этих действий появляется вариант измеряемой характеристики, для которой отклонение результата измерения от измеряемой величины положительно и достигает максимального значения равного h.

Аналогичным образом можно показать, что отклонение результата измерения от измеряемой величины может быть отрицательным и достигать по модулю максимального значения близкого к h. Соответствующий вариант измеряемой характеристики на промежутке $(\lambda_н, \lambda_в)$ будет иметь график, близкий к $\Phi^H(\lambda_1)$ и в районе отрезка BF узкий выступ в сторону точки B.

Формула (24) справедлива также для описания потока излучения опорного источника и имеет вид:

$$|\Phi^0_\lambda(\vartheta) - \Phi^0_\lambda(\lambda_1)| \le h^0, \qquad (25)$$

где h^0 — нелинейность спектральной характеристики излучения опорного источника на промежутке $(\lambda_н, \lambda_в)$.

Получив ключевые формулы (24) и (25), найдем связь неопределенности результата измерения спектральной плотности $\Phi_\lambda(\lambda_1)$ с нелинейностями h и h^0 спектральных характеристик измеряемого и опорного источников для второго варианта.

Для этого введем следующие обозначения аналогичные обозначениям (14):

$$\Delta\Phi_1 = \Phi_\lambda(\vartheta) - \Phi_\lambda(\lambda_1);\ \Delta\Phi^0_1 = \Phi^0_\lambda(\vartheta^0) - \Phi^0_\lambda(\lambda_1). \qquad (26)$$

Отсюда получим:

$$\Phi_\lambda(\vartheta) = \Phi_\lambda(\lambda_1) + \Delta\Phi_1;\ \Phi^0_\lambda(\vartheta^0) = \Phi^0_\lambda(\lambda_1) + \Delta\Phi^0_1. \qquad (27)$$

Формула (13) справедлива для обоих вариантов, поэтому, подставив в нее (27), получим, после несложных преобразований, соотношение, аналогичное (15) для второго варианта,

$$\Phi_\lambda(\lambda_1) = \Phi_\lambda^0(\lambda_1)\frac{U}{U^0} + \Delta\Phi_1^0\frac{U}{U^0} - \Delta\Phi_1. \qquad (28)$$

Второе и третье слагаемые в правой части равенства (28) являются фактическими отклонениями измеряемой величины $\Phi_\lambda(\lambda_1)$ от результата измерения $\Phi_\lambda^0(\lambda_1)\frac{U}{U^0}$, вызванными нелинейностью спектральной характеристики опорного и измеряемого источников.

Модуль отношения третьего слагаемого (28) к измеряемой величине $\Phi_\lambda(\lambda_1)$ в соответствии с (24) и (26) удовлетворяет неравенству:

$$\left|\frac{\Delta\Phi_1}{\Phi_\lambda(\lambda_1)}\right| \le \frac{h}{\Phi_\lambda(\lambda_1)} = \gamma, \qquad (29)$$

где γ — относительная нелинейность спектральной характеристики измеряемого источника.

Модуль отношения второго слагаемого (28) к измеряемой величине $\Phi_\lambda(\lambda_1)$ в соответствии с (25) и (26) удовлетворяет следующему неравенству:

$$\left|\frac{\Delta\Phi_1^0\frac{U}{U^0}}{\Phi_1(\lambda_1)}\right| \cong \left|\frac{\Delta\Phi_1^0\frac{U}{U^0}}{\Phi_\lambda^0(\lambda_1)\frac{U}{U^0}}\right| = \left|\frac{\Delta\Phi_1^0}{\Phi_\lambda^0(\lambda_1)}\right| \le \frac{h^0}{\Phi_\lambda^0(\lambda_1)} = \gamma^0, \qquad (30)$$

где γ^0 — относительная нелинейность спектральной характеристики опорного источника.

Полученные формулы (29) и (30) показывают, что относительные неопределенности результата измерения спектральной плотности $\Phi_\lambda(\lambda_{ср})$ по второму варианту не превышают соответствующих величин γ и γ^0.

На практике относительные изменения спектральных характеристик на спектральных промежутках, как правило, гораздо больше относительных нелинейностей, поэтому в большинстве случаев измерения по второму варианту гораздо точнее, чем по первому.

Основные результаты

Приведем конкретные примеры измерения спектральных характеристик ОЭС.

Пример 1. Измерение спектральной плотности энергетической яркости такого источника оптического излучения как глобар (стержень из порошкообразного карбида кремния, использующийся при нагревании как источник ИК излучения) [22] в нормальном к излучающей поверхности направлении, на длине волны 3,7 мкм, при температуре глобара 1400 °С. В качестве опорного источника оптического излучения используем черное тело при температуре 800 °С. Положим, что спектральная характеристика чувствительности измерительной ОЭС имеет треугольную форму с максимумом на длине волны 3,7 мкм и протяженность — 0,1 мкм.

Рассчитали относительные изменения и нелинейности спектральных характеристик источников, используя формулу Планка для излучения абсолютно черного тела [21]. В результате этих вычислений получены следующие значения:

— для глобара $\delta = 6{,}6$ % и $\gamma = 0{,}06$ %:
— для черного тела $\delta^0 = 3{,}4$ % и $\gamma^0 = 0{,}01$ %.

Эти значения, подставленные в формулы (16), (19), (29) и (30), показывают, что методические неопределенности результата данного измерения при новом подходе оказываются, по крайней мере, на два порядка меньше, чем при традиционных методах измерения.

Пример 2. Измерение спектральной чувствительности ОЭС на экспериментальной установке, подробно описанной в [14]. Сигнал ОЭС описывается интегралом (1) при условии, что функция $S(\lambda)$ считается линейной в пределах спектрального промежутка ($\lambda_н$, $\lambda_в$), а произведение $\Phi_\lambda(\lambda)\tau(\lambda)$ зависит от длины волны. Сложность расчета связана с тем, что $\tau(\lambda)$ включает в себя пропускание атмосферной трассы длиной около 10 м. В соответствии с предлагаемой теорией соотношение (1) можно представить в виде:

$$U = \Phi S(\lambda_1), \qquad (31)$$

где $\Phi = \int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda$; $\lambda_1 = \frac{\int_{\lambda_н}^{\lambda_в} \lambda\tau(\lambda)\Phi_\lambda(\lambda)d\lambda}{\int_{\lambda_н}^{\lambda_в} \tau(\lambda)\Phi_\lambda(\lambda)d\lambda}$.

Из формулы (31) следует, что для определения $S(\lambda_1)$ необходимо измерить три величины: U — сигнал ОЭС; Φ — спектрозональный поток, непосредственно воздействующий на ОЭС; λ_1 — спектральную координату потока Φ.

Спектрозональный поток в данном случае [14] измеряется посредством опорного фотоприемного устройства (ОФУ). Для измерения λ_1 служит откалиброванный САП. Испытуемая ОЭС, ОФУ и САП работают поочередно и устанавливаются так, чтобы длина атмосферного хода лучей от источника до приемника, во всех трех случаях, была одинакова.

Косвенное измерение λ_1 основано на спектрограмме потока Φ, вырабатываемой САП. При этом используются методы численного интегрирования данной спектрограммы для определения ее спектральной координаты. Отметим, что спектральные координаты потока и соответствующей спектрограммы оказываются равны с достаточной для практики точностью при соблюдении некоторых условий. Важно, чтобы аппаратная функция САП была симметричной и неизменной в пределах спектрального промежутка ($\lambda_н$, $\lambda_в$). При несоблюдении этих условий возникает методическая неопределенность измерения λ_1, которую, впрочем, можно оценить.

Пример 3. Измерение спектрального пропускания. По аналогии с примером 2, можно описать процедуру измерения спектральной характеристики пропускания $\tau(\lambda)$ оптической системы, входящей в подынтегральное выражение (1).

Обсуждение

Главное преимущество предлагаемого метода — исключение необходимости учета спектрального пропускания атмосферы в измерительном тракте, что является основным источником неопределенности в

традиционных методах, особенно в ИК диапазоне. Это кардинально упрощает и ускоряет процедуру калибровки, а также повышает ее точность.

Метод позволяет представить спектрозональные фотометрические величины посредством двух параметров: уровня потока и его спектральной координаты. Это существенно сокращает объем необходимых данных и упрощает их обработку.

Эффективность метода зависит от отношения относительной нелинейности измеряемой характеристики к ее относительному изменению. В большинстве практических случаев нелинейность значительно меньше изменения, что обеспечивает высокую эффективность метода. В окрестностях экстремумов, где эти величины могут быть сравнимы, метод остается применимым, но точность снижается.

По сравнению с существующими решениями, требующими сложных компенсационных измерений или ограничивающимися диапазонами высокой прозрачности атмосферы, предлагаемый подход является более универсальным и метрологически обоснованным.

Заключение

Разработан новый подход к измерениям и расчетам спектральных характеристик селективных элементов оптико-электронных систем, основанный на использовании спектрозональных потоков излучения и их спектральных координат. Подход позволяет исключить влияние переменных параметров атмосферы на результаты измерений, что особенно актуально для инфракрасного диапазона. Диапазон применения данного метода весьма широк. Он применим для измерения спектральной плотности излучения источников, спектральной чувствительности приемников и спектрального пропускания оптических систем. При этом эффективность подхода зависит от отношения относительной нелинейности измеряемой характеристики к ее относительному изменению. В большинстве случаев первая величина гораздо меньше второй и эффективность высокая. В некоторых случаях возможно равенство этих величин (например, в окрестностях экстремальных точек), тогда метод все равно применим, но снижение методической неопределенности не будет происходить.

Теоретически обоснована и подтверждена на примерах более высокая точность метода по сравнению с традиционными подходами.

Главным положительным моментом метода является возможность представления сложных спектральных распределений двумя интегральными величинами — уровнем потока излучения и его спектральной координатой, что открывает новые возможности для метрологического обеспечения широкого спектра оптико-электронных систем в промышленности, науке и экологии.

Литература

1. Григорьев А.В., Демин А.В., Сечак Е.Н. Многоспектральная оптико-электронная система // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25. № 1. С. 1–8. https://doi.org/10.17586/2226-1494-2025-25-1-1-8
2. Manakov A. Evaluation of computational radiometric and spectral sensor calibration techniques // Proceedings of SPIE. 2016. V. 9896. P. 98960O. https://doi.org/10.1117/12.2228071
3. Pospelov G.V., Savin S.V. Optimization of the operating spectral band of optoelectronic devices for detecting point objects against the background of outer space // Journal of Optical Technology. 2018. V. 85. N 7. P. 416–418. https://doi.org/10.1364/JOT.85.000416
4. Vasil'ev V.N., Gridin A.S., Dmitriev I.Yu., Sinel'shchikov V.V., Tomeev K.A. Device for the functional monitoring of the through path of a scanning optoelectronic device with a large-format multi-element radiation detector // Journal of Optical Technology. 2019. V. 86. N 8. P. 466–470. https://doi.org/10.1364/jot.86.000466
5. Eckart M., Adams J., Boyce K., Brown G., Chiao M., Fujimoto R., et al. Ground calibration of the Astro-H (Hitomi) soft x-ray spectrometer // Proceedings of SPIE. 2016. V. 9905. P. 99053W. https://doi.org/10.1117/12.2233053
6. Кукушкин Д.Е., Сазоненко Д.А., Бахолдин А.В., Красавцев В.М. Моделирование спектральных систем. Часть первая. СПб: Университет ИТМО, 2021. С. 5.
7. Горбунов Г.Г., Демин А.В., Никифоров В.О., Савицкий А.М., Скворцов Ю.С., Сокольский М.Н., Трегуб В.П. Гиперспектральная аппаратура для дистанционного зондирования Земли // Оптический журнал. 2009. Т. 76. № 10. С. 75–82.
8. Гектин Ю.М., Зорин С.М., Трофимов Д.О. Способ измерения спектральных характеристик в видимом и инфракрасном спектральных диапазонах и установка, реализующая этот способ. Патент RU 2710382C1. Опубликовано: 26.12.2019 Бюл. № 36.
9. Feng F., Huo D., Zhang Z., Lou Y., Wang S., Gu Z., et al. Symbiotic evolution of photonics and artificial intelligence: a comprehensive review // Advanced Photonics. 2025. V. 7. N 2. P. 024001. https://doi.org/10.1117/1.ap.7.2.024001
10. Markushin G.N., Korotaev V.V., Koshelev A.V., Samokhina I.A., Vasilev A.S., Timofeev A.N., Vasileva A.V., Yaryshev S.N. Dual-band

References

1. Grigoriev A.V., Demin A.V., Sechak E.N. Multispectral optoelectronic system. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 1, pp. 1–8. (in Russian). https://doi.org/10.17586/2226-1494-2025-25-1-1-8
2. Manakov A. Evaluation of computational radiometric and spectral sensor calibration techniques. *Proceedings of SPIE*, 2016, vol. 9896, pp. 98960O. https://doi.org/10.1117/12.2228071
3. Pospelov G.V., Savin S.V. Optimization of the operating spectral band of optoelectronic devices for detecting point objects against the background of outer space. *Journal of Optical Technology*, 2018, vol. 85, no. 7, pp. 416–418. https://doi.org/10.1364/JOT.85.000416
4. Vasil'ev V.N., Gridin A.S., Dmitriev I.Yu., Sinel'shchikov V.V., Tomeev K.A. Device for the functional monitoring of the through path of a scanning optoelectronic device with a large-format multi-element radiation detector. *Journal of Optical Technology*, 2019, vol. 86, no. 8, pp. 466–470. https://doi.org/10.1364/jot.86.000466
5. Eckart M., Adams J., Boyce K., Brown G., Chiao M., Fujimoto R., et al. Ground calibration of the Astro-H (Hitomi) soft x-ray spectrometer. *Proceedings of SPIE*, 2016, vol. 9905, pp. 99053W. https://doi.org/10.1117/12.2233053
6. Kukushkin D.E., Sazonenko D.A., Bakholdin A.V., Krasavtsev V.M. *Spectral Systems Modelling*. Part I. St. Petersburg, ITMO University, 2021, pp. 5. (in Russian)
7. Gorbunov G.G., Demin A.V., Nikiforov V.O., Savitskii A.M., Skvortsov Y.S., Sokol'skii M.N., Tregub V.P. Hyperspectral apparatus for remote probing of the earth. *Journal of Optical Technology*, 2009, vol. 76, no. 10, pp. 651–656. https://doi.org/10.1364/JOT.76.000651
8. Gektin Y.M., Zorin S.M., Trofimov D.O. Method of measuring spectral characteristics in the visible and infrared spectral ranges and an apparatus which realizes said method. Patent RU 2710382C1, 2019. (in Russian)
9. Feng F., Huo D., Zhang Z., Lou Y., Wang S., Gu Z., et al. Symbiotic evolution of photonics and artificial intelligence: a comprehensive review. *Advanced Photonics*, 2025, vol. 7, no. 2, pp. 024001. https://doi.org/10.1117/1.ap.7.2.024001
10. Markushin G.N., Korotaev V.V., Koshelev A.V., Samokhina I.A., Vasilev A.S., Timofeev A.N., Vasileva A.V., Yaryshev S.N. Dual-band

optoelectronic poaching detection systems // Journal of Optical Technology. 2022. V. 89. N 9. P. 528–536. https://doi.org/10.1364/jot.89.000528

11. Iureva R.A., Timko A.S., Maltseva N.K., Raskin E.O., Chernaya A.S. Optoelectronic methods and tools for pipeline's internal surface diagnosis // Proceedings of SPIE. 2018. V. 10690. P. 106901R. https://doi.org/10.1117/12.2306567
12. Sun Y., Liang R. High-resolution, Full-freedom spectral tunable light source for skin discrimination // Proc. of the 14th International Conference on Optics-photonics Design & Fabrication. 2024.
13. Osipov V.M., Borisova N.F. Taking absorbed infrared radiation into account during the testing and certification of optoelectronic equipment // Journal of Optical Technology. 2019. V. 86. N 8. P. 471–475. https://doi.org/10.1364/jot.86.000471
14. Ильинский А.В., Мальцева Н.К. Метод многоспектральной имитации излучения точечных объектов // Оптический журнал. 2010. Т. 77. № 2. С. 74–78.
15. Maltseva N.K., Ilinskii A.V. Providing the uniform field of Illumination in wide spectral and dynamic ranges // Proceedings of SPIE. 2020. V. 11483. P. 114830B. https://doi.org/10.1117/12.2575618
16. Butler J.J., Brown S.W., Saunders R.D., Johnson B.C., Biggar S.F., Zalewski E.F., et al. Radiometric measurement comparison on the integrating sphere source used to calibrate the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Landsat 7 Enhanced Thematic Mapper plus (ETM+) // Journal of Research of the National Institute of Standards and Technology. 2003. V. 108. N 3. P. 199–228. https://doi.org/10.6028/jres.108.020
17. Scharpf D., Holt J., Durell C.N., Mascia P., Dann M., Dobrowski L., Rabade S. A new solar spectrum source for absolute radiometric calibration of advanced multispectral sensors and hyperspectral imagers // Proceedings of SPIE. 2024. V. 13143. P. 1314307. https://doi.org/10.1117/12.3026762
18. Scheiding S., Driescher H., Walter I., Hanbuch K., Paul M., Hartmann M., Scheiding M. Compact blackbody calibration sources for in-flight calibration of spaceborne infrared instruments // Proceedings of SPIE. 2017. V. 10563. P. 105635P. https://doi.org/10.1117/12.2304132
19. Трофимов Д.О., Гектин Ю.М., Зорин С.М., Зайцев А.А. Метрологические и методические аспекты спектрально-энергетических калибровок оптико-электронной аппаратуры ДЗЗ // Ракетно-космическое приборостроение и информационные системы. 2018. Т. 5. № 2. С. 26–33. https://doi.org/10.30894/issn2409-0239.2018.5.2.26.33
20. Смирнов В.И. Курс высшей математики. Т. 1. М.: Государственное издательство технико-теоретической литературы, 1954. С. 224–225.
21. Мирошников М.М. Теоретические основы оптико-электронных приборов. СПб: Лань, 2022. С. 258–264.
22. Носенко Т.Н., Ситникова В.Е., Стрельникова И.Е., Фокина М.И. Практикум по колебательной спектроскопии: учебное пособие. СПб: Университет ИТМО, 2021. С. 16.

optoelectronic poaching detection systems. *Journal of Optical Technology*, 2022, vol. 89, no. 9, pp. 528–536. https://doi.org/10.1364/jot.89.000528

11. Iureva R.A., Timko A.S., Maltseva N.K., Raskin E.O., Chernaya A.S. Optoelectronic methods and tools for pipeline's internal surface diagnosis. *Proceedings of SPIE*, 2018, vol. 10690, pp. 106901R. https://doi.org/10.1117/12.2306567
12. Sun Y., Liang R. High-resolution, Full-freedom spectral tunable light source for skin discrimination. *Proc. of the 14th International Conference on Optics-photonics Design & Fabrication*, 2024.
13. Osipov V.M., Borisova N.F. Taking absorbed infrared radiation into account during the testing and certification of optoelectronic equipment. *Journal of Optical Technology*, 2019, vol. 86, no. 8, pp. 471–475. https://doi.org/10.1364/jot.86.000471
14. Il'inskiĭ A.V., Mal'tseva N.K. Method of multispectral simulation of the radiation of point objects. *Journal of Optical Technology*, 2010, vol. 77, no. 2, pp. 137–140. https://doi.org/10.1364/jot.77.000137
15. Maltseva N.K., Ilinskii A.V. Providing the uniform field of Illumination in wide spectral and dynamic ranges. *Proceedings of SPIE*, 2020, vol. 11483, pp. 114830B. https://doi.org/10.1117/12.2575618
16. Butler J.J., Brown S.W., Saunders R.D., Johnson B.C., Biggar S.F., Zalewski E.F., et al. Radiometric measurement comparison on the integrating sphere source used to calibrate the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Landsat 7 Enhanced Thematic Mapper plus (ETM+). *Journal of Research of the National Institute of Standards and Technology*, 2003, vol. 108, no. 3, pp. 199–228. https://doi.org/10.6028/jres.108.020
17. Scharpf D., Holt J., Durell C.N., Mascia P., Dann M., Dobrowski L., Rabade S. A new solar spectrum source for absolute radiometric calibration of advanced multispectral sensors and hyperspectral imagers. *Proceedings of SPIE*, 2024, vol. 13143, pp. 1314307. https://doi.org/10.1117/12.3026762
18. Scheiding S., Driescher H., Walter I., Hanbuch K., Paul M., Hartmann M., Scheiding M. Compact blackbody calibration sources for in-flight calibration of spaceborne infrared instruments. *Proceedings of SPIE*, 2017, vol. 10563, pp. 105635P. https://doi.org/10.1117/12.2304132
19. Trofimov D.O., Gektin Yu.M., Zorin S.M., Zaitsev A.A. Metrological and methodical aspects of spectral-energetic calibrations of optoelectronic ERS equipment. *Rocket-Space Device Engineering and Information Systems*, 2018. vol. 5, no. 2, pp. 26–33. (in Russian). https://doi.org/10.30894/issn2409-0239.2018.5.2.26.33
20. Smirnov V.I. *High Mathematics Course*. Vol. 1. Moscow, Gosudarstvennoe izdatel'stvo tehniko-teoreticheskoj literatury, 1954, pp. 224–225. (in Russian)
21. Miroshnikov M.M. *Optical-Electronic Devices. Theoretical Aspects*. St. Petersburg, Lanbook, 2022, pp. 258–264. (in Russian)
22. Nosenko T.N., Sitnikova V.E., Strel'nikova I.E., Fokina M.I. *Practicum on Vibrational Spectroscopy*. St. Petersburg, ITMO University, 2021, pp. 16. (in Russian)

Авторы

Ильинский Александр Владимирович — заместитель директора научно-образовательного центра «Музей истории Университета ИТМО», Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0005-6413-6411, avilinskii@itmo.ru

Мальцева Надежда Константиновна — кандидат технических наук, директор научно-образовательного центра «Музей истории Университета ИТМО», Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 51564440800, https://orcid.org/0009-0004-4515-0261, nkmaltceva@itmo.ru

Раскин Евгений Олегович — директор по молодежной политике, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57190605879, https://orcid.org/0000-0001-7830-6841, raskin@itmo.ru

Authors

Alexander V. Ilinskii — Vice-Director, Scientific Educational Center "ITMO University History Museum", ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0005-6413-6411, avilinskii@itmo.ru

Nadezhda K. Maltseva — PhD, Director, Scientific Educational Center "ITMO University History Museum", ITMO University, Saint Petersburg, 197101, Russian Federation, sc 51564440800, https://orcid.org/0009-0004-4515-0261, nkmaltceva@itmo.ru

Eugene O. Raskin — Director for Youth Policy, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57190605879, https://orcid.org/0000-0001-7830-6841, raskin@itmo.ru

Статья поступила в редакцию 25.07.2025
Одобрена после рецензирования 10.10.2025
Принята к печати 25.11.2025

Received 25.07.2025
Approved after reviewing 10.10.2025
Accepted 25.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2025 Том 25 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2025 Vol. 25 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

ФОТОНИКА И ОПТОИНФОРМАТИКА

PHOTONICS AND OPTOINFORMATICS

doi: 10.17586/2226-1494-2025-25-6-1033-1046
УДК 535.329

Плазмонный резонанс и аномальная дисперсия диэлектрической проницаемости и показателя преломления пористой лазерно-структурированной поверхности анодированного титана

Алёна Андреевна Кострина[1]✉, Анна Владимировна Цибульникова[2], Василий Анатольевич Слежкин[3], Иван Игоревич Лятун[4], Максим Владимирович Царьков[5], Дмитрий Александрович Артамонов[6], Михаил Сергеевич Курицкий[7], Илья Геннадьевич Самусев[8], Валерий Вениаминович Брюханов[9]

[1,2,3,4,5,6,7,8,9] Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация

[1] aakostrina@gmail.com✉, https://orcid.org/0009-0002-0943-9290
[2] memorgold@mail.ru, https://orcid.org/0000-0001-8578-0701
[3] vasiliy.slezhkin@klgtu.ru, https://orcid.org/0000-0002-2801-7029
[4] ILyatun@kantiana.ru, https://orcid.org/0000-0002-4988-8077
[5] max.offreal@gmail.com, https://orcid.org/0009-0000-5345-4223
[6] euroset2016ig98@icloud.com, https://orcid.org/0000-0003-2141-8900
[7] Mix-Mux@yandex.ru, https://orcid.org/0000-0002-4125-4678
[8] ISamusev@kantiana.ru, https://orcid.org/0000-0001-5026-7510
[9] Bryukhanov_v.v@mail.ru, https://orcid.org/0000-0003-4689-7207

Аннотация

Введение. Лазерно-структурированные поверхности диоксида титана (TiO_2) привлекают внимание благодаря сочетанию высокой удельной площади, биосовместимости и уникальных оптических свойств, открывающих перспективы для фотоники, сенсорики и энергетики. Особый интерес представляет изучение оптических характеристик пористых пленок Ti/TiO_2, полученных методом лазерного структурирования, с возможным проявлением плазмонных резонансов и аномальной дисперсии. **Метод.** Для создания образцов использовалась титановая лента, подвергнутая анодированию в растворе гидроксида калия с последующим наносекундным лазерным структурированием при длине волны 1064 нм и плотности энергии $(3{,}2 \pm 0{,}2)\cdot 10^3$ Дж/см^2. Морфология поверхностей исследовалась методами сканирующей электронной микроскопии и оптической профилометрии, а оптические характеристики – методами спектрофотометрии и эллипсометрии. Для интерпретации спектральных данных применена модифицированная модель Адачи–Фороухи в дипольном приближении, позволяющая количественно описать вклад межзонных переходов и плазмонных мод. **Основные результаты.** Полученные в результате лазерного структурирования поверхности образцов характеризуются выраженной пористостью (линейный размер пор составляет 300–1100 нм, глубина — около 200 нм), наличием субмикронных трещин и наночастиц лазерно-структурированного материала. В спектрах отражения выявлены минимумы, соответствующие возбуждению поверхностных плазмонов и интерференционных мод. В спектрах диэлектрической проницаемости обнаружена область аномальной дисперсии и локализации поля в области длины волны 625 нм. Рассчитаны толщина скин-слоя, коэффициент Парселла для нанопоры, длина затухания плазмонных колебаний на поверхности образца, длина распространения поверхностных плазмонов и критическое значение усиления поляризуемости в области локализации плазмонного резонанса. Согласно полученным результатам моделирования, ширина запрещенной зоны сужается до 1,016 эВ. Определены значения влияния насыщения межзонного поглощения, а также воздействий, обусловленных изменением зонной структуры и от свободных носителей в полупроводниковую компоненту диэлектрической проницаемости. Решающее значение имеет влияние, вызванное сужением запрещенной зоны, однако наибольший вклад в экспериментально наблюдаемое поведение диэлектрической проницаемости вносит генерация резонансных плазмонных мод.

© Кострина А.А., Цибульникова А.В., Слежкин В.А., Лятун И.И., Царьков М.В., Артамонов Д.А., Курицкий М.С., Самусев И.Г., Брюханов В.В., 2025

Обсуждение. Установлено, что ключевым механизмом оптического отклика является резонансная локализация электромагнитного поля в нанопорах, что подтверждает проявление гиперболического метаматериального поведения. Полученный материал характеризуется значительным сужением запрещенной зоны вследствие наносекундного лазерного воздействия. Результаты исследования демонстрируют перспективность пористых лазерно-структурированных поверхностей анодированного Ti для фотонных и сенсорных устройств, а также в качестве волноводных структур.

Ключевые слова
лазерное структурирование, пористая поверхность, оксид титана, диэлектрическая проницаемость, плазмонный резонанс, эффект Парселла

Благодарности
Исследование выполнено в рамках Федерального задания Министерства науки и высшего образования Российской Федерации (проект № FZWM-2024-0010).

Ссылка для цитирования: Кострина А.А., Цибульникова А.В., Слежкин В.А., Лятун И.И., Царьков М.В., Артамонов Д.А., Курицкий М.С., Самусев И.Г., Брюханов В.В. Плазмонный резонанс и аномальная дисперсия диэлектрической проницаемости и показателя преломления пористой лазерно-структурированной поверхности анодированного титана // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1033–1046. doi: 10.17586/2226-1494-2025-25-6-1033-1046

Plasmon resonance and anomalous dispersion of the dielectric permittivity and refractive index of the porous laser-structured surface of anodized titanium

Alena A. Kostrina[1]✉, Anna V. Tcibulnikova[2], Vasily A. Slezhkin[3], Ivan I. Lyatun[4], Maksim V. Tsarkov[5], Dmitry A. Artamonov[6], Michael S. Kuritskij[7], Ilia G. Samusev[8], Valery V. Bryukhanov[9]

[1,2,3,4,5,6,7,8,9] Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation

[1] aakostrina@gmail.com✉, https://orcid.org/0009-0002-0943-9290
[2] memorgold@mail.ru, https://orcid.org/0000-0001-8578-0701
[3] vasiliy.slezhkin@klgtu.ru, https://orcid.org/0000-0002-2801-7029
[4] ILyatun@kantiana.ru, https://orcid.org/0000-0002-4988-8077
[5] max.offreal@gmail.com, https://orcid.org/0009-0000-5345-4223
[6] euroset2016ig98@icloud.com, https://orcid.org/0000-0003-2141-8900
[7] Mix-Mux@yandex.ru, https://orcid.org/0000-0002-4125-4678
[8] ISamusev@kantiana.ru, https://orcid.org/0000-0001-5026-7510
[9] Bryukhanov_v.v@mail.ru, https://orcid.org/0000-0003-4689-7207

Abstract
In recent years, laser-structured titanium dioxide (TiO_2) surfaces have attracted considerable attention due to their combination of high specific surface area, biocompatibility, and unique optical properties, offering promising opportunities for photonics, sensing, and energy applications. Of particular interest is the study of the optical manifestations of porous Ti/TiO_2 films fabricated via laser structuring, with potential evidence of plasmonic resonances and anomalous dispersion. The samples were prepared from titanium foil subjected to anodization in potassium hydroxide solution, followed by nanosecond laser structuring at the wavelength of 1064 nm and an energy density of $(3.2 \pm 0.2)\cdot 10^3$ J/cm^2. Surface morphology was analyzed using scanning electron microscopy and optical profilometry, while optical characteristics were investigated by spectrophotometry and ellipsometry. To interpret the spectral data, a modified Adachi-Forouhi model within the dipole approximation was applied, enabling quantitative description of the contributions of interband transitions and plasmonic modes. The surfaces produced by laser structuring exhibited pronounced porosity (pore sizes of 300–1100 nm, depth ~200 nm), submicron cracks, and nanoparticles of the laser-structured material. Reflection spectra revealed minima corresponding to the excitation of surface plasmons and interference modes. Dielectric permittivity spectra displayed a region of anomalous dispersion and field localization at a wavelength of 625 nm. Calculated parameters included the skin layer thickness, Purcell factor for a nanopore, damping length of plasmon oscillations on the surface, propagation length of surface plasmons, and the critical value of polarizability enhancement in the plasmon resonance localization region. Modeling indicated a narrowing of the bandgap to 1.016 eV. Contributions to the dielectric permittivity of the semiconductor component from interband absorption saturation, changes in band structure, and free carriers were determined. While the bandgap narrowing played a decisive role, the dominant contribution to the experimentally observed dielectric behavior arose from the generation of resonant plasmonic modes. It was established that the key mechanism of the optical response is the resonant localization of the electromagnetic field within the nanopores, confirming the manifestation of hyperbolic metamaterial behavior. The material obtained exhibited significant bandgap narrowing due to the nanosecond laser treatment. The results highlight the potential of porous laser-structured anodized titanium surfaces for photonic and sensing devices, as well as for use in waveguiding structures.

Keywords
laser structuring, porous surface, titanium dioxide, dielectric function, plasmon resonance, Purcell effect

Acknowledgements
This research was carried out within the framework of the Federal Program of the Ministry of Science and Higher Education of the Russian Federation (project No. FZWM-2024-0010).

For citation: Kostrina A.A., Tcibulnikova A.V., Slezhkin V.A., Lyatun I.I., Tsarkov M.V., Artamonov D.A., Kuritskij M.S., Samusev I.G., Bryukhanov V.V. Plasmon resonance and anomalous dispersion of the dielectric permittivity and refractive index of the porous laser-structured surface of anodized titanium. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1033–1046 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1033-1046

Введение

В последнее десятилетие исследование пористых поверхностей диоксида титана (TiO_2) привело к возможности рассматривать его новые прикладные назначения в различных сферах наукоемкого производства. Высокая удельная площадь поверхности, прочность, биосовместимость, химическая активность, высокая способность к адгезии с органическими веществами и клетками, а также относительная дешевизна — эти свойства обуславливают значительную востребованность пористых поверхностей TiO_2 в фотокатализе [1], сенсорах [2], литий-ионных аккумуляторах [3], сенсибилизированных красителем солнечных элементах [4], биомедицине, оптике, электронике и др.

Фотокаталитические свойства пористых поверхностей TiO_2 активно используются в процессах разложения органических загрязнителей и летучих органических соединений [5]. Способность пористой поверхности TiO_2 образовывать реакционноспособные радикалы в результате поглощения ультрафиолетового (УФ) излучения позволяет использовать его в устройствах для очистки воды и воздуха и антибактериальных покрытиях [6].

В области возобновляемой энергетики такие поверхности находят применение в солнечных элементах, самоочищающихся покрытиях и технологиях фотокаталитического расщепления воды для получения водорода. Дополнительно, модифицированные поверхности TiO_2 способствуют повышению чувствительности и эффективности сенсоров, оптических устройств, диодов и транзисторов, а также росту долговечности покрытий [7]. Кроме того, в биомедицине пористые поверхности Ti/TiO_2, поглощающие УФ-излучение, применяются в качестве каркасов для роста костной ткани, в системах доставки лекарств и для обработки медицинских устройств [8, 9]. Создание материалов с пористой иерархической структурой различной морфологии может улучшать функциональные свойства материала, включая взаимодействие с биомолекулами и клетками [10].

Кроме того, пористые поверхности TiO_2 применяют в области оптики и фотоники ввиду плазмонных свойств и способности проявлять свойства метаматериалов с отрицательным показателем преломления [11]. Такие материалы, обладая отрицательной диэлектрической и магнитной проницаемостями, позволяют реализовать эффекты сверхвысокого разрешения и отрицательного преломления света, что открывает перспективы для создания сверхплоских линз, высокоэффективных сенсоров и фотонных устройств, адаптивных оптических систем и стелс-технологий [11, 12].

Методы модификации поверхности TiO_2 имеют решающее значение в управлении его пористостью и другими свойствами. В частности, методы фемто- и наносекундного лазерного скрайбирования позволяют синтезировать поверхности с контролируемой микро- и наноразмерной морфологиями, что в дальнейшем определяет ее оптические и фотофизические свойства [13–15]. Исходя из этого, в настоящей работе был использован метод наносекундного лазерного воздействия с заданными энергетическими и геометрическими параметрами скрайбирования.

Отметим, что оптические характеристики материалов из TiO_2 хорошо изучены в основном в УФ-области, соответствующей ширине запрещенной зоны TiO_2 (3,0–3,6 эВ), а рассмотрению проблем взаимодействия излучения на пористой поверхности уделено достаточно мало внимания. Задачи распространения излучения на таких поверхностях сложны из-за физических особенностей материала и требуют дополнительных экспериментальных исследований.

Заметим, что именно модификация поверхности короткими лазерными импульсами позволяет создавать нанопористые образования с контролируемыми геометрическими параметрами, способные дискретно поглощать электромагнитное излучение определенных длин волн. При этом влияние размерности таких образований на оптические процессы в тонкой пленке TiO_2 изучено недостаточно. Научная новизна настоящей работы заключается в экспериментальном выявлении и моделировании резонансного поведения диэлектрической проницаемости и показателя преломления лазерно-структурированного анодированного Ti с пористой поверхностью размером, сопоставимым с длиной волны падающего излучения.

Таким образом, целью работы является исследование оптических процессов, связанных с распределением падающего излучения в тонкой пленке TiO_2 с расположенными на ее поверхности нанопорами, сформированными наносекундным лазерным воздействием в условиях контроля энергетических параметров излучения и морфологических свойств поверхности.

Методика проведения экспериментов и подготовки образцов

Методика подготовки образцов анодированной поверхности Ti. В качестве исходного материала для синтезирования поверхностей использовалась титановая лента (ВТ1-0, 0,1 × 100 мм, ООО «Фортис Металл и Дизайн», Россия), которая тщательно обрабатывалась перед началом лазерного структурирования. Нарезанные пластины площадью S = 25 см^2 подвергались снятию окислов и загрязнений в 40 % азотной кислоте в течение 10 мин, после чего промывались в дистиллированной воде и были проанодированы в 3 % растворе гидроксида калия, что позволяло увеличить механическую устойчивость и улучшить оптические свойства поверхности металла. Схема процесса анодирования включала в себя программируемый источник

питания MAISHENG WSD-60H15 (Китай), резервуар с электролитом, а также две титановых пластины (катод и анод), к которым были присоединены провода для подключения к источнику питания. Провода подсоединялись к пластинам с помощью щупов с зажимом типа «крокодил». Механизм образования и роста оксидного покрытия в значительной степени определяется химическим составом и кристаллическим строением исходного материала, типом применяемого электролита и выбранным технологическим режимом анодирования. В связи с этим были предварительно подобраны оптимальные параметры источника питания и время выдержки образца в электролите для достижения равномерного покрытия Ti пленкой фиолетового цвета толщиной от 70 до 200 нм: 1,4–1,7 А в течение 3 мин [16]. О формировании пленки TiO_2 свидетельствовал характерный цвет поверхности образцов (фиолетовый), коррелирующий с ее толщиной. Далее образцы промывались в дистиллированной воде и сушились при температуре 20 °C. Всего было изготовлено более 600 образцов, часть из которых впоследствии подвергалась лазерному структурированию.

Методика лазерного структурирования образцов анодированного Ti. Лазерное структурирование образцов анодированного Ti проводилось с использованием твердотельного импульсного Nd:YAG-лазера LQ 929A (Solar Laser Systems, Беларусь) с выходной энергией до 1,5 Дж согласно структурной схеме (рис. 1). Лазер генерировал излучение основной гармоники на длине волны λ = 1064 нм и второй гармоники — на λ = 532 нм, причем непосредственно для лазерного структурирования поверхностей анодированного Ti использовалась основная. Длительность импульсов составляла 10 нс при частоте следования 10 Гц.

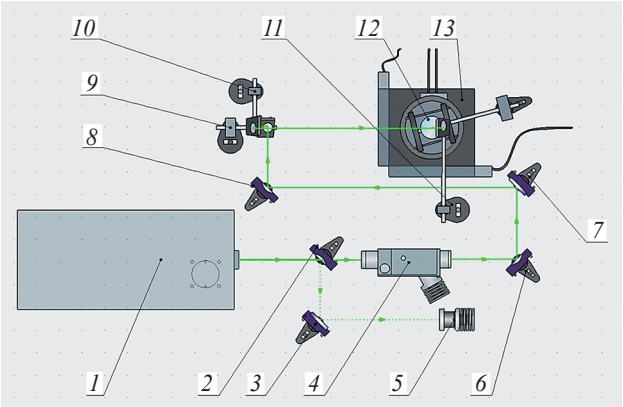

Рис. 1. Структурная схема: *1* — лазерный источник; *2* — светоделитель; *3*, *6–11* — зеркала; *4* — ослабитель лазерного излучения; *5* — заглушка; *12* — фокусирующая линза; *13* — моторизированный линейный транслятор (Standa 8MTF) с поворотной платформой (Standa 8MRL120-15-MEn)

Fig. 1. Optical setup: *1* — laser source; *2* — beam splitter; *3*, *6–11* — mirrors; *4* — laser beam attenuator; *5* — beam stop; *12* — focusing lens; *13* — motorized linear translator (Standa 8MTF) with a rotation stage installed on it (Standa 8MRL120-15-MEn)

В начале структурной схемы установлен светоделитель *2*, предназначенный для отражения излучения с λ = 532 нм и пропускания излучения с λ = 1064 нм, которое направляется на зеркало *3* и заглушку *5* соответственно. Таким образом, в экспериментах использовалось только излучение с λ = 1064 нм и энергией пучка 0,5 Дж, проходящее через ослабитель лазерного излучения *4* (ослабление излучения до 0,36 Дж), а затем — через систему зеркал *6–11*. Далее луч попадал на лепестковую диафрагму (d = 1 мм) и на фокусирующую линзу *12* (f = 10 см), установленную на высоте 10 см от поверхности закрепленного на поворотной платформе анодированного Ti. Пятно лазерного луча направлялось строго в центр линзы под углом 90° к ее плоской поверхности (выпуклая поверхность линзы обращена вниз к образцу), а его размер и морфология контролировались при помощи визуализатора излучения (ЗАО «НПФ «Люминофор», Россия). Диаметр пятна на образце анодированного Ti составил 120 ± 5 мкм. Энергия импульса лазерного излучения, падающего на поверхность Ti после прохождения структурной схемы, составила 0,36 ± 0,01 Дж. Отклонения формы пятна от круга были недопустимы, что могло привести к неравномерному распределению энергии по поверхности материала. В результате могли возникнуть локальные дефекты, неоднородное плавление или испарения и, как следствие, формирование нерегулярных или искаженных наноструктур, что может негативно сказаться на результатах эксперимента.

Перед каждым процессом лазерного структурирования проводилось совмещение оси лазерного луча с осью вращения поворотной платформы, синхронизированной с моторизированным линейным транслятором *13*. Образец анодированного Ti был закреплен в центре поверхности поворотной платформы при помощи специальных креплений типа «лапки» и липкого слоя, после чего производился выбор геометрии структуры. Управление моторизированной частью схемы *13* осуществлялось с помощью персонального компьютера с установленной программной средой XILab, которая содержала скрипт программного кода с возможностью задания параметров перемещения элементов моторизированной части схемы. В программе XILab были заданы скорость вращения поворотной платформы (постоянное значение скорости или ее изменение для получения определенного заранее заданного количества точек лазерного воздействия на каждое кольцо лазерного трека), начальный радиус первого внутреннего кольца лазерного трека, радиальное расстояние между соседними кольцами и их количество, количество проходов лазерного луча по выбранной траектории и возможность изменения направления движения поворотной платформы при лазерном структурировании.

Плотность энергии излучения, падающего на поверхность Ti, составляла $(3{,}2 \pm 0{,}2)\cdot 10^3$ Дж/см^2. После окончания лазерного структурирования готовая поверхность подвергалась ультразвуковой чистке в ультразвуковой ванне Elmasonic P (Elma, Германия) в течение 15 мин в дистиллированной воде для удаления осевших во время лазерного структурирования наночастиц TiO_2.

Очищенные образцы были высушены при комнатной температуре.

По методике лазерного структурирования для оценки воспроизводимости результатов была изготовлена серия из 10 образцов. Контроль повторяемости результатов был осуществлен методом сканирующей электронной микроскопии. Учитывая однотипность выявленных структурных характеристик, рассмотрение каждого из образцов представляется нецелесообразным, в связи с чем дальнейшее описание выполнено на примере одного типичного образца анодированного Ti, лазерно-структурированного согласно методике.

Методика исследования образцов. Для характеризации морфологии поверхности оксидной пленки Ti после лазерного воздействия были получены ее изображения методом сканирующей электронной микроскопии (Zeiss Cross Beam-540 (FIB-SEM, Германия)). Спектры отражения поверхности в видимой и ближней инфракрасной (БИК) областях были зарегистрированы на спектрофотометре UV-2600 (Shimadzu, Япония), а спектры диэлектрической проницаемости и показателя преломления измерены с помощью эллипсометра AUTO SE (HORIBA and Jobin Yvon, Франция). Результаты профилометрии поверхности образцов получены и обработаны с использованием оптического 3D профилометра NanoView-2400 (Nanosystem, КНДР).

Обсуждение

В результате импульсного лазерного воздействия в приповерхностном слое металла развиваются сложные неравновесные процессы, обусловленные фазовыми переходами, локальным испарением, абляцией и действием термокапиллярных сил [17, 18]. Переход материала в жидкое состояние сопровождается нестабильностями расплава, вызванными градиентами температуры и поверхностного натяжения, а также неравномерным тепловым расширением [19]. По завершении воздействия лазерного импульса происходит сверхбыстрое охлаждение и затвердевание, сопровождающееся образованием новой микроструктуры, зерен с близкой кристаллографической ориентацией, термострикционными напряжениями и общим изменением рельефа поверхности. Совокупность вышеперечисленных явлений определяет морфологию формирующихся наноструктур, а наличие, характер и вклад этих явлений, в свою очередь, определяются условиями лазерного структурирования [17–19].

Параметры лазерного структурирования исследуемого образца: плотность энергии — $(3{,}2 \pm 0{,}2)\cdot 10^3$ Дж/см^2; расстояние между соседними кольцами — 50 мкм; скорость вращения поворотной платформы — 24 °/с (при радиусе кольца $r < 5$ мм) и 12 °/с ($r > 5$ мм); количество точек лазерного воздействия на кольцо — 150 ($r < 5$ мм) и 300 ($r > 5$ мм).

С учетом известных настроек лазера и задаваемой угловой скорости вращения поворотной платформы было определено количество точек лазерного воздействия в кольце:

$$n = \frac{L}{\omega(T+\tau)} = \frac{2\pi}{\omega(T+\tau)},$$

где L — длина окружности кольца; ω — угловая скорость вращения поворотной платформы; T — период между лазерными импульсами, равный 0,1 с; $\tau << T$ — длительность лазерных импульсов, составляющая 10 нс.

Рассмотрим морфологию полученной структуры центральной области поверхности образца (в радиусе до 5 мм от центра колец) (рис. 2).

На рис. 2 можно заметить «гребни» трека лазерного луча и промежутки между ними, а также наличие различных наноструктур. Наблюдаемые структуры могут оказывать значительное влияние на оптические и физико-химические свойства поверхности. Плавление материала свидетельствует о локальном нагреве поверхности свыше 1840 °С.

Снижение скорости вращения поворотной платформы в периферийной области структуры (с 24 до 12 °/с) позволило увеличить частоту точек лазерного воздействия и обеспечить равномерное волнообразное распределение материала с периодичностью около 10 мкм с характерным локальным оплавлением пленки на всей площади структуры (рис. 2, *a*). В центре полученной структуры наблюдается значительное растрескивание оксидной пленки. Было замечено, что на поверхности образца также наблюдается переосаждение субмикронных частиц и наночастиц TiO_2 диаметром от 110 до 1000 нм (рис. 2, *b*, *c*). Наблюдается интенсивная пористость поверхности (рис. 2, *c*) в центральной области образца (радиусом до 5 мм от центра колец). Ширина пор варьируется от значений субмикронных до 1,1 мкм и демонстрирует квазипериодичность, что является признаком процессов самоорганизации структуры.

Для периферийной области структуры (радиусом от 5 мм от центра колец) характерно образование растрескиваний материала и субмикронных ям, однако пор и их массивов не наблюдается (рис. 2, *d*). Напротив, центральная область (радиусом до 5 мм от центра колец) характеризуется набором различных микро- и наноструктур, образующих массивы, вследствие чего именно она была выбрана для дальнейшего исследования.

Рассмотрим результаты исследования шероховатости поверхности образца методом оптической профилометрии (рис. 3, таблица).

При наблюдаемых параметрах шероховатости поверхности размеры рельефа сопоставимы с длинами волн видимого и БИК диапазонов. Это создает условия для проявления оптических поверхностных явлений, таких как генерация плазмонных колебаний, генерация плазмон-фононных колебаний и нелинейные эффекты в спектральном составе оптических функций [11, 12]. Отрицательное значение коэффициента асимметрии (Ssk = –0,501) свидетельствует о преобладании впадин в рельефе. Максимальные параметры рельефа (Sp = 1,66 мкм, Sv = 3,16 мкм, Sz = 4,82 мкм) подтверждают наличие как выступающих участков, так и углублений, сформировавшихся в результате перераспределения материала и, возможно, его локального удаления.

Отметим постепенное увеличение высоты поверхности образца по мере удаления от его центра, что, вероятно, связано с вынесением материала в периферийные зоны и его частичным переосаждением в виде

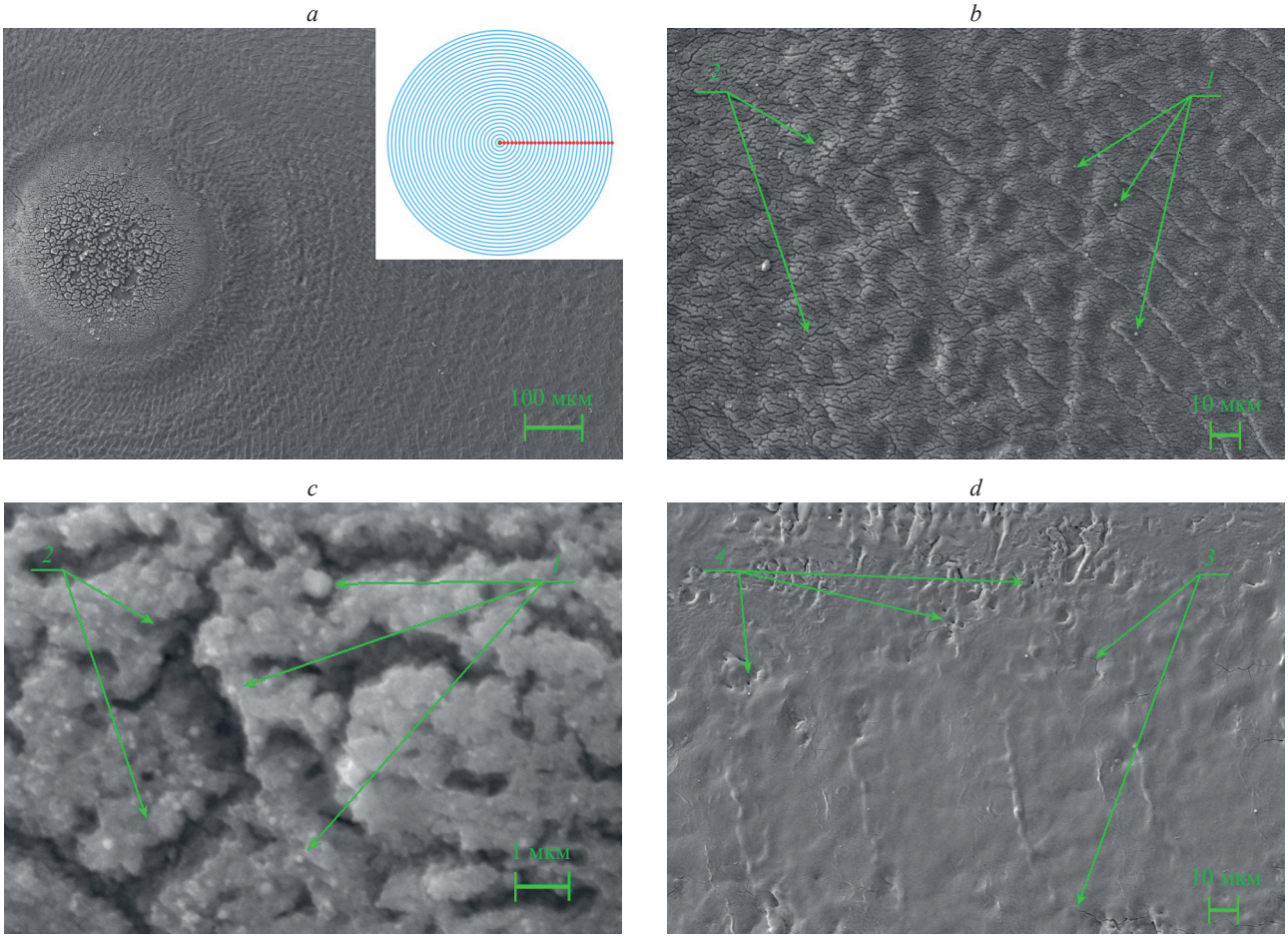

Рис. 2. Результаты сканирующей электронной микроскопии центральной (в радиусе до 5 мм от центра колец) (*a–c*) и периферийной (*d*) областей образца.

1 — субмикронные частицы и наночастицы; *2* — пористые образования и их массивы; *3* — растрескивания; *4* — субмикронные ямы. На вставке (рис. 2, *a*) представлено схематичное изображение процесса формирования лазерно-структурированных поверхностей с кольцевой геометрией

Fig. 2. Scanning electron microscopy results of the central region of the sample (within a radius of up to 5 mm from the ring center) (*a–c*) and the peripheral region (*d*)

1 — submicron particles and nanoparticles; *2* — porous structures and their arrays; *3* — cracking; *4* — submicron holes. The inset (Fig. 2, *a*) shows a schematic representation of the process of forming laser-structured surfaces with a ring-shaped geometry

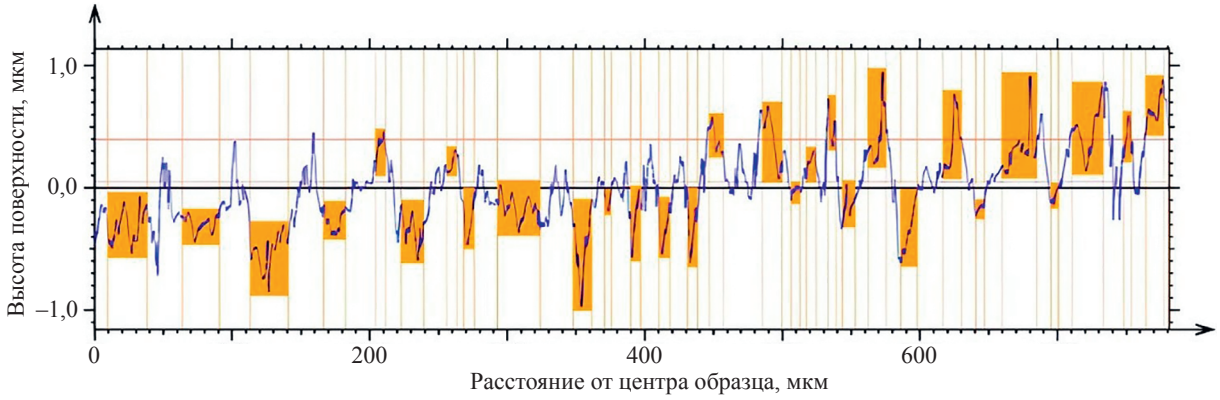

Рис. 3. Результаты профилометрического анализа центральной области образца.

Оранжевым цветом отмечены локальные элементы поверхности образца

Fig. 3. Profilometric analysis results of the central region of the sample.

Local elements of the sample surface are marked in orange

Таблица. Параметры лазерно-структурированной части поверхности образца, полученные методом профилометрического анализа

Table. Parameters of the laser-structured region of the sample surface obtained from profilometric analysis

Параметр	Величина
Среднее квадратичное отклонение высот *a*, мкм	0,402
Коэффициент асимметрии *Ssk*	–0,501
Коэффициент эксцесса *Sku*	4,31
Максимальное отклонение по высоте *Sp*, мкм	1,66
Максимальная глубина впадин *Sv*, мкм	3,16
Максимальная высота профиля *Sz*, мкм	4,82
Средний размер шероховатости $\bar{\xi}^2$, мкм	0,312

субмикронных частиц, наночастиц. Диапазон ширины гребней структуры варьируется от 4,88 до 30,9 мкм, что свидетельствует о значительной гетерогенности (неоднородности) микроструктуры в области, где происходило снижение скорости вращения поворотной платформы (в радиусе 5 мм от центра колец). Средняя глубина пор составляет около 200 нм.

Рассмотрим оптические свойства поверхности образца в видимом диапазоне длин волн (рис. 4). В дальнейшем используются следующие обозначения: *Rs* — коэффициент отражения *s*-поляризованного излучения; *Rp* — коэффициент отражения *p*-поляризованного излучения; *R* — коэффициент отражения неполяризованного излучения; Re ε — действительная компонента функции диэлектрической проницаемости (ДП); Im ε — мнимая компонента функции ДП; *n* — показатель преломления.

Аномальная дисперсия, наблюдаемая в спектрах ДП и показателя преломления, может быть обусловлена анизотропией поверхности и неравномерным распределением нанопор [20]. Рассмотрим распределение падающего излучения с длиной волны видимого и ИК диапазона (450–1000 нм) в области локализации нанопор поверхности. Ширина пор составляет от 300 до 1000 нм, расстояние между порами — от 500 до 2000 нм. Отметим, что длина волны поглощения составляет 610 нм, что лежит в диапазоне размера пор. Исходя из этого, электромагнитное излучение длин волн от 450 до 600 нм, поглощается с большей интенсивностью, чем излучение более длинноволновой части спектра ($\lambda > 600$ нм) (рис. 4, *a*). Заметим, что интенсивность отражения *p*-излучения в указанном диапазоне меньше, чем для *s*-компоненты на 20 %, что в большей степени обусловлено плазмонной природой поглощения падающего излучения наряду с эванесцентными волнами и потерями в металле [21, 22]. Также спектры поляризованного излучения имеют минимумы на длине волны 725 нм: $Rs = 0{,}968$, $Rp = 0{,}781$, что может указывать на наличие дополнительной моды возбуждения, связанной с интерференционными эффектами в пористых структурах [23]. Максимумы в областях 625 и 825 нм свидетельствуют о возбуждении, соответственно, поверхностного плазмонного резонанса ($\lambda \approx 625$ нм) и интерференционного резонанса между дном пористых образований и поверхностью ($\lambda \approx 825$ нм), что связано с субмикронной морфологией и глубиной структур около 200 нм [24].

В спектре отражения на рис. 4, *b* наблюдается минимум на длине волны 560 нм ($R = 3$ %), указывающий на возбуждение плазмонных мод на границе Ti/TiO_2, при котором падающая энергия концентрируется внутри нанопор. Незначительный скачок в области 825 нм согласуется со вторым максимумом отражения на рис. 4, *a*, который может указывать на возбуждение дополнительной моды резонансного типа, обусловленной, например, интерференционными эффектами.

Спектр ДП (рис. 4, *c*) демонстрирует наличие резонансных процессов в области длины волны 625 нм. Действительная и мнимая части функции в данной области изменяются скачкообразно, что указывает на эффективное поглощение энергии падающего излучения и возбуждение локализованных резонансных мод [22, 23]. За пределами резонансной области ($\lambda < 450$ нм и $\lambda > 720$ нм) обе компоненты стремятся к нулю, что может указывать на реализацию «epsilon-near-zero» (ENZ) состояния [25, 26]. Для лазерно-структурированных пористых сред такое поведение обусловлено взаимной компенсацией влияния свободных носителей, зонных переходов и резонансных эффектов, что сопровождается снижением эффективной поляризуемости и плотности оптических состояний. Отметим, что для поверхности анодированного Ti скачкообразных эффектов не наблюдается (вставка, рис. 4, *c*).

На спектре показателя преломления (рис. 4, *d*) виден скачок скорости распространения электромагнитных волн в области длины волны 625 нм ($n = 67$), что указывает на аномальную дисперсию, связанную с резонансным взаимодействием падающего излучения с плазмонными модами пористой структуры, а отрицательное значение в минимуме на 650 нм ($n = -4$) предполагает наличие обратного эффекта, характерного для метаматериалов и резонансных структур [27, 28]. Далее значение *n* стремится к нулю вплоть до 1000 нм.

Учитывая высокую степень анизотропности материала, вклад экситонных или фононных мод не оказывает определяющего влияния на оптические характеристики образца. Это может быть обусловлено изменением электронной структуры материала, образованием дефектов, субоксидов, аморфных и частично проводящих фаз, а также особенностями морфологии поверхности в результате ее лазерной модификации.

Рассмотрим данные процессы более подробно.

На рис. 5 представлена трехмерная модель элемента пористой поверхности.

Выполним анализ взаимодействия излучения видимого диапазона длин волн с наноразмерной порой в рамках формализма Парселла [29]. Поскольку размер пор сравним с длиной волны падающего излучения, представляется возможным использовать классический подход к описанию преобразования падающей на шероховатую поверхность электромагнитной волны. Коэффициент Парселла *Fp* характеризует скорость взаимодействия дипольного излучателя с формой ре-

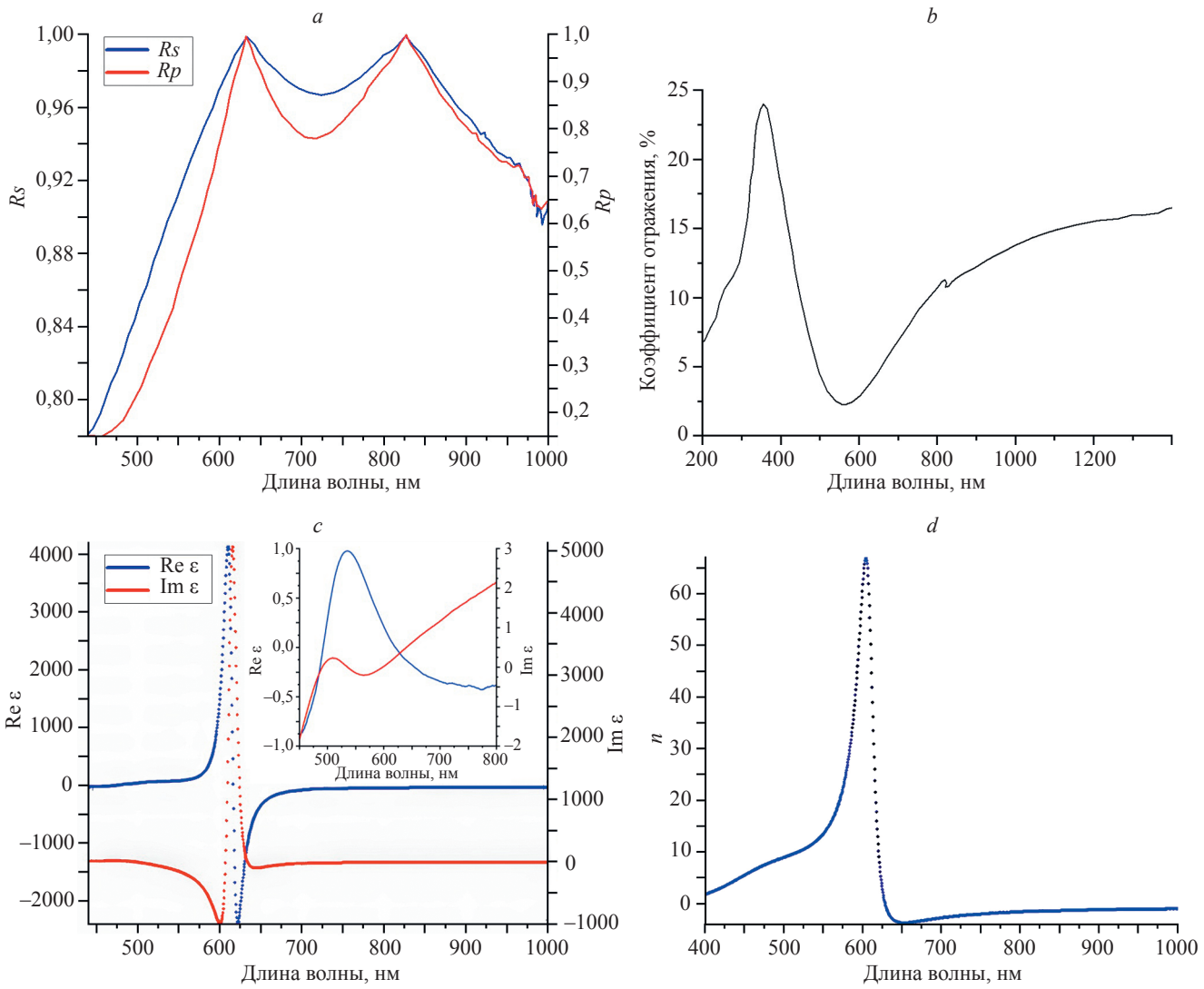

Рис. 4. Оптические характеристики лазерно-структурированной поверхности.

Спектры отражения поляризованного излучения (*a*); отражения неполяризованного излучения (*b*); функций диэлектрической проницаемости (*c*); показателя преломления (*d*).

Вставка (рис. 4, *c*) — спектры диэлектрической проницаемости поверхности анодированного титана

Fig. 4. Optical characteristics of the laser-structured surface.

Reflection spectra of polarized radiation (*a*); reflection spectrum of unpolarized radiation (*b*); dielectric permittivity functions spectra (*c*); refractive index spectrum (*d*).

The inset (Fig. 4, *c*) shows the dielectric permittivity spectra of the anodized titanium surface

зонатора: геометрические и морфологические свойства резонатора определяют его возможность усиливать взаимодействие излучения с материалом резонатора посредством генерации плазмонных мод различного спектрального состава [30, 31]. Плазмонный отклик в спектрах диэлектрических функций определяет как интенсивность и скорость затухания генерируемых колебательных мод, так и особенности процессов поглощения, отражения и рассеяния на пористой шероховатой поверхности. Спектральная форма резонанса функций ДП демонстрирует свойства гиперболических метаматериалов, для которых характерно различие знаков у действительной и мнимой компонент [20, 29, 32]. Резонансное поведение ДП и показателя преломления в области 625 нм соответствует асимптоте гиперболической функции.

Вычислим фактор Парселла, устанавливающий фактор усиления интенсивности колебания плазмонов в нанопоре [21]:

$$Fp = \frac{3}{4\pi^2}\lambda^3\frac{Q}{V}, \qquad (1)$$

где λ — длина волны в среде; $Q = \alpha/\delta$ — добротность (α — линейный размер резонансной полости, принятый за 900 нм, δ — глубина скин-слоя при данной частоте падающей электромагнитной волны); V — объем резонатора. Таким образом, для вычисления фактора Парселла необходимо найти значение глубины скин-слоя полученного материала. Представленные резонансные полости (нанопоры) можно рассматривать как соединенные между собой составляющие пористых

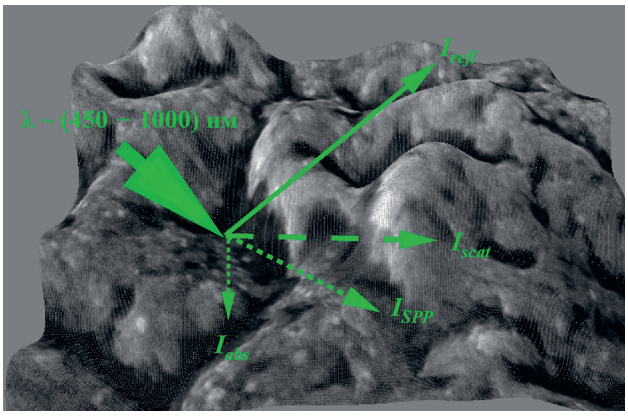

Рис. 5. Трехмерное изображение элемента пористой поверхности образца: I_{refl}, I_{scat} и I_{abs} — интенсивности отраженного, рассеянного и поглощенного излучений; I_{SPP} — интенсивность излучения, генерируемого поверхностными плазмонными колебаниями

Fig. 5. Three-dimensional image of a porous surface element: I_{refl}, I_{scat}, I_{abs} — intensities of reflected, scattered, and absorbed radiation; I_{SPP} — intensity of radiation generated by surface plasmon oscillations

массивов, в связи с чем в качестве линейного размера резонатора (нанопоры) была выбрана его ширина.

В результате лазерного структурирования пленки TiO_2 происходят сложные химические и физические процессы, приводящие не только к изменению морфологии, но и к изменению фазового состояния материала. Высокая температура и энергия могут приводить к разрушению кристаллической структуры TiO_2, взаимодействию материала с присутствующими в атмосфере газами, к изменению кристаллической фазы, аморфизации, а также к редуцированию до субоксидов при определенных локальных условиях (например, вследствие пониженного содержания кислорода) [33]. Таким образом, материал внутри пор представляет собой, вероятно, смесь TiO_2, его редуцированных форм и проводящих включений (наночастиц Ti). Это позволяет рассматривать его как комбинированную среду. Таким образом, величину скин-слоя для полученного материала можно определить по следующим формулам [34]:

$$\delta = \frac{1}{k_i} = \frac{\lambda_0}{2\pi\sqrt{\mathrm{Re}\,\varepsilon}\,\mathrm{Im}[(1 + i\tan\zeta)^{1/2}]},$$
$$k = \frac{2\pi}{\lambda}\sqrt{\mathrm{Re}\,\varepsilon - i\,\mathrm{Im}\,\varepsilon}, \qquad (2)$$

где k — комплексное волновое число; λ_0 — длина волны в вакууме; $\tan\zeta$ — тангенс угла потерь. С учетом полученных ранее в работе результатов и анализов спектров, по формулам (2) было вычислено значение $\delta = 2{,}78$ нм. При данной величине скин-слоя поглощение энергии достаточно интенсивное, и основная часть поля затухает вблизи поверхности, что подтверждается спектрами отражения. Таким образом, с использованием рассчитанных и найденных экспериментально величин, по формуле (1) было получено значение: $Fp = 40$.

Определим длины затухания плазмонных колебаний на шероховатой поверхности и распространения поверхностных плазмонов [35]:

$$l_{sp}^{(m)} = \left[2\frac{\omega}{c}\,\mathrm{Im}\left(\frac{\varepsilon}{\varepsilon + 1}\right)^{1/2}\right]^{-1},$$
$$l_{sp}^{(sp)} \approx \left[\frac{3}{2}\pi\bar{\xi}^2 a^2 \frac{\omega^5}{c^5}\frac{|\varepsilon|^{7/2}}{|\varepsilon + 1|^{9/2}}\right]^{-1}, \qquad (3)$$

где ω — частота плазмонного резонанса; c — скорость света в вакууме; ε — комплексная ДП среды; экспериментально полученные величины: $\bar{\xi}^2$ — средний размер наблюдаемой шероховатости; a — известное среднее квадратичное отклонение высот или длина корреляции (таблица).

Так, для резонансной длины волны плазмонного поглощения 625 нм, с учетом известных параметров шероховатости, значения $l_{sp}^{(m)}$ и $l_{sp}^{(sp)}$ составляют 500 и 20 нм соответственно. Полученные по формулам (3) значения свидетельствуют о сильной локализации электромагнитного поля в нанопорах материала.

Критическое значение усиления поляризуемости в области локализации плазмонного резонанса может быть определено по формуле [23]

$$\alpha_c = \gamma\frac{2\mathrm{Re}\,\varepsilon(\omega_{spp}) + 1}{2c\sqrt{\mathrm{Re}\,\varepsilon\,(\omega_{spp})}}, \qquad (4)$$

где γ — фактор затухания; ω_{spp} — частота возбуждения поверхностных плазмонов.

С учетом фактора затухания для титановой поверхности $\gamma = 72\cdot10^{12}$ с$^{-1}$ [36], согласно формуле (4), α_c составляет $3{,}024\cdot10^3$ см$^{-1}$, что согласуется по порядку величины с данными для серебра и золота (α_c около 10^3 см$^{-1}$ [36]).

Выполним моделирование метаповерхности методом эллипсометрии. Для аппроксимации функций ДП была использована модифицированная модель Адачи–Фороухи [37] с дипольным приближением. Приведем основные уравнения для аппроксимации экспериментально полученных функций комплексной диэлектрической проницаемости $\varepsilon(E)$:

$$\varepsilon(E) = \varepsilon_0(E) + \varepsilon_1(E),$$
$$\varepsilon_0(E) = A_0\left\{\frac{f(\chi_0)}{E_0^{3/2}} + \frac{f(\chi_{S0})}{2(E_0 + \Delta_0)^{3/2}}\right\},$$
$$\varepsilon_1(E) = (n_1(E) + ik_1(E))^2,$$

где $\varepsilon_0(E)$ — компонента, описывающая вклад межзонных электронных переходов; $\varepsilon_1(E)$ — компонента, отражающая вклад осцилляторных процессов; A_0 — параметр, пропорциональный величине поглощения в максимуме и квадратичному матричному элементу перехода; E — энергия фотонов; E_0 — центральная энергия пика поглощения; $f(\chi)$ — функция, описывающая спектральное распределение вероятности переходов; χ_0 — комплексный параметр, учитывающий уширение спектра; χ_{S0} — комплексный параметр, учитывающий сдвиг за счет ширины запрещенной зоны;

Δ_0 — значение ширины запрещенной зоны; $n_1(E)$ — действительная часть показателя преломления, описывающая скорость распространения электромагнитной волны в материале и связанная с дисперсией; $\varepsilon_1(E) = (n_1(E) + ik_1(E))^2$ — мнимая часть показателя преломления, описывающая поглощение энергии волны материалом и связанная с затуханием.

Коэффициент согласия Пирсона χ^2 для моделирования составляет 0,023 и свидетельствует о возможности применения данной модели для расчета экспериментальных функций.

Результат аппроксимации экспериментальных функций примененной моделью Адачи–Фороухи представлен на рис. 6.

С использованием модифицированной модели Адачи–Фороухи были получены такие параметры как фактор затухания (Γ_0 = 8,76), ширина запрещенной зоны (Δ_0 = 1,016 эВ), величина размытия в запрещенную зону (0,021 эВ), высокочастотная ДП (ε_∞ = 6000), впоследствии использованные для вычислений полупроводниковых вкладов в величину ДП. Отметим, что теоретические значения ширины запрещенной зоны для фазы рутила и анатаза составляют 3,0 и 3,2 эВ соответственно, для брукита — от 2,3 до 3,3 эВ [38], а для диэлектрических пленок TiO_2 значения могут достигать 3,9 эВ [39]. Данные результаты значительно отличаются от полученного значения в настоящей работе — 1,016 эВ. Такое сужение ширины запрещенной зоны указывает на существенное изменение электронных свойств материала и может быть обусловлено образованием дефектных уровней, субоксидов, аморфных фаз, металлических включений и фазовых преобразований.

Так, исходя из диаграммы состояний для оксидов титана, можно предположить, что при лазерной модификации происходят локальный нагрев и восстановление до таких форм как, например, Ti_2O_3 и Ti_3O_5, для которых Δ_0 составляет от 0,1 до 0,3 эВ [40]. Появление включений указанных и иных оксидов приводит к переходу от диэлектрической среды к узкозонным или проводящим средам, что подтверждается экспериментальными данными.

Поскольку поверхность представляет собой шероховатый рельеф ($\bar{\xi}^2$ = 0,312 мкм, a = 0,402 мкм), а сам материал является сложной комбинированной средой, процессы трансформации излучения также сложны и функции отклика ДП таких поверхностей представляют собой суммарный отклик физических процессов, связанных в том числе с проводящими и полупроводниковыми свойствами системы [41]:

$$\varepsilon_m = \varepsilon_g + \Delta\varepsilon_{pop} + \Delta\varepsilon_{bgs} + \Delta\varepsilon_{fcr},$$

где ε_g — ДП материала; влияния: $\Delta\varepsilon_{pop}$ — насыщения межзонного поглощения; $\Delta\varepsilon_{bgs}$ — изменения зонной структуры (сжатия запрещенной зоны); $\Delta\varepsilon_{fcr}$ — свободных носителей. Определим воздействие каждой компоненты согласно следующим выражениям:

$$\Delta\varepsilon_{pop} = -(\varepsilon_g - 1)\frac{N_{e-h}}{N_0}, \quad (5)$$

$$\Delta\varepsilon_{bgs} = \varepsilon(\hbar\omega + \Delta E_{gap}) - \varepsilon(\hbar\omega), \quad (6)$$

$$\Delta\varepsilon_{fcr} = -\frac{N_{e-h}e^2}{\varepsilon_0 m^*_{opt} m_e \omega^2}\frac{1}{1 + i\dfrac{1}{\omega\tau_D}} = -\left(\frac{\omega_P}{\omega}\right)^2\frac{1}{1 + i\dfrac{1}{\omega\tau_D}}, \quad (7)$$

где N_{e-h} — плотность возбужденных электрон-дырочных пар; N_0 — плотность электронных состояний в валентной зоне; $\Delta E_{gap} \sim \sqrt[3]{N_{e-h}}$ — сдвиг зоны запрещенной энергии из-за плотности носителей; m^*_{opt} — оптическая эффективная масса носителей; τ_D — время затухания электронного газа Друде; $\omega_P = \sqrt{N_{e-h}e^2/\varepsilon_0 m^*_{opt} m_e}$ — плазменная частота; e — заряд электрона; ε_0 — диэлектрическая постоянная; m_e — масса электрона.

Плотности электронов и возбужденных электрон-дырочных пар были оценены на основе экспериментально полученных спектров ДП и сопоставления с данными научных публикаций для близких по составу и морфологии пленок. Значение $N_0 = 1{,}87 \cdot 10^{20}$ см$^{-3}$, что близко к значению, соответствующему пленке TiO_2 [42]. Величина $N_{e-h} \cong 10^{20}$ см$^{-3}$ — совпадает с верхней границей диапазона от 10^{19} до 10^{20} см$^{-3}$ [43], вычисленное значение частоты $\omega = 3{,}02 \cdot 10^{15}$ рад/с, а значение m^*_{opt} с учетом данных работы [44] составляет

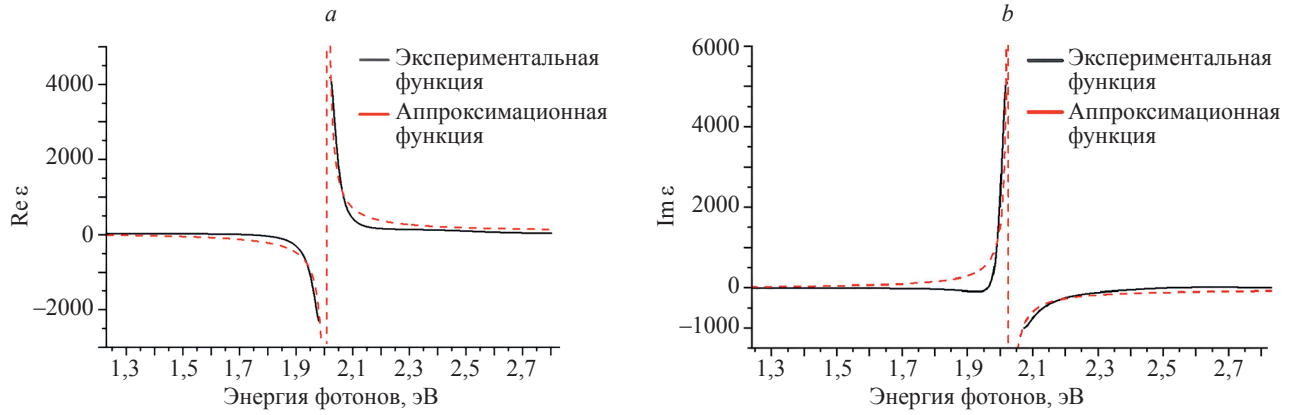

Рис. 6. Спектры действительной (*a*) и мнимой (*b*) компонент функции диэлектрической проницаемости

Fig. 6. Spectra of the real (*a*) and imaginary (*b*) components of the dielectric permittivity function

0,55. Значение ДП $\varepsilon_g \approx -0{,}1$ было найдено из экспериментального спектра поверхности фиолетового анодированного Ti (вставка, рис. 4, *c*). Полученное время затухания плазмонных колебаний (свободные электроны модели Друде) составило 1 пс, что согласуется с длительностью затухания поверхностных плазмонов на металлических шероховатых поверхностях [23].

Полученные по формулам (5)–(7) значения составили: $\Delta\varepsilon_{pop} \approx 2{,}22$; $\Delta\varepsilon_{bgs} \approx 2138 + 642i$; $\Delta\varepsilon_{fcr} \approx -0{,}107 + 0{,}04i$. Таким образом, было установлено, что основной вклад в изменение полупроводниковой компоненты ДП вносят процессы изменения зонной структуры материала (значение действительной части — 2138) в тонкой пленке TiO_2.

Заключение

Анализ оптических спектров выявил резонансные эффекты и аномальную дисперсию, характерные для гиперболических метаматериалов, а также сильную локализацию электромагнитного поля в нанопорах размера, сопоставимого с длиной волны падающего излучения. Аналитическая модель Адачи–Фороухи позволила выделить физические параметры, характеризующие полупроводниковые свойства и резонансные процессы в нанопорах. В рамках модели установлено, что ключевой вклад в изменение полупроводниковой компоненты функции диэлектрической проницаемости вносит изменение зонной структуры вследствие сужения запрещенной зоны. Решающее значение для наблюдаемого поведения диэлектрической проницаемости имеет генерация плазмонных мод.

Полученные результаты открывают перспективы использования структурированных наносекундным лазерным воздействием пористых поверхностей титана в фотонных и плазмонных приложениях, в частности в качестве волноводных поверхностей с удерживающей модой поверхностных плазмонных колебаний на длине волны 625 нм. Дальнейшие исследования, включая моделирование и структурный анализ, позволят углубить понимание взаимосвязи между морфологией и оптическими свойствами наблюдаемых наноструктур.

Литература

1. Macak J.M., Zlamal M., Krysa J., Schmuki P. Self-organized TiO_2 nanotube layers as highly efficient photocatalysts // Small. 2007. V. 3. N 2. P. 300–304. https://doi.org/10.1002/smll.200600426
2. Arafat M.M., Dinan B., Akbar S.A., Haseeb A.S.M.A. Gas sensors based on one dimensional nanostructured metal-oxides: a review // Sensors. 2012. V. 12. N 6. P. 7207–7258. https://doi.org/10.3390/s120607207
3. Ma D., Li K., Pan J.H. Ultraviolet-induced interfacial crystallization of uniform nanoporous biphasic TiO_2 spheres for durable lithium-ion battery // ACS Applied Energy Materials. 2020. V. 3. N 5. P. 4186–4192. https://doi.org/10.1021/acsaem.0c00816
4. Ji Y., Zhang M., Cui J., Lin K., Zheng H., Zhu J., Samia A.C.S. Highly-ordered TiO_2 nanotube arrays with double-walled and bamboo-type structures in dye-sensitized solar cells // Nano Energy. 2012. V. 1. N 6. P. 796–804. https://doi.org/10.1016/j.nanoen.2012.08.006
5. Yeung K.L., Yau S.T., Maira A.J., Coronado J.M., Soria J., Yue P.L. The influence of surface properties on the photocatalytic activity of nanostructured TiO_2 // Journal of Catalysis. 2003. V. 219. N 1. P. 107–116. https://doi.org/10.1016/s0021-9517(03)00187-8
6. Jafari S., Mahyad B., Hashemzadeh H., Janfaza S., Gholikhani T., Tayebi L. Biomedical applications of TiO_2 nanostructures: recent advances // International Journal of Nanomedicine. 2020. V. 15. P. 3447–3470. https://doi.org/10.2147/ijn.s249441
7. Hoshian S., Jokinen V., Hjort K., Ras R.H.A., Franssila S. Amplified and localized photoswitching of TiO_2 by micro- and nanostructuring // ACS Applied Materials & Interfaces. 2015. V. 7. N 28. P. 15593–15599. https://doi.org/10.1021/acsami.5b04309
8. Yavari S.A., van der Stok J., Chai Y.C., Wauthle R., Birgani T.Z., Habibovic P., et al. Bone regeneration performance of surface-treated porous titanium // Biomaterials. 2014. V. 35. N 24. P. 6172–6181. https://doi.org/10.1016/j.biomaterials.2014.04.054
9. Yavari S.A., Wauthle R., Böttger A.J., Schrooten J., Weinans H., Zadpoor A.A. Crystal structure and nanotopographical features on the surface of heat-treated and anodized porous titanium biomaterials produced using selective laser melting // Applied Surface Science. 2014. V. 290. P. 287–294. https://doi.org/10.1016/j.apsusc.2013.11.069
10. Reghunath S., Pinheiro D., Devi K.R.S. A review of hierarchical nanostructures of TiO_2: Advances and applications // Applied Surface Science Advances. 2021. V. 3. P. 100063. https://doi.org/10.1016/j.apsadv.2021.100063
11. Kaushik V. Negative Index Materials: Metamaterials. Research Review International Journal of Multidisciplinary. 2019. V. 4. Issue 4 [Электронный ресурс]. URL: http://www.rrjournals.com. (дата обращения: 25.12.2024)

References

1. Macak J.M., Zlamal M., Krysa J., Schmuki P. Self-organized TiO_2 nanotube layers as highly efficient photocatalysts. *Small*, 2007, vol. 3, no. 2, pp. 300–304. https://doi.org/10.1002/smll.200600426
2. Arafat M.M., Dinan B., Akbar S.A., Haseeb A.S.M.A. Gas sensors based on one dimensional nanostructured metal-oxides: a review. *Sensors*, 2012, vol. 12, no. 6, pp. 7207–7258. https://doi.org/10.3390/s120607207
3. Ma D., Li K., Pan J.H. Ultraviolet-induced interfacial crystallization of uniform nanoporous biphasic TiO_2 spheres for durable lithium-ion battery. *ACS Applied Energy Materials*, 2020, vol. 3, no. 5, pp. 4186–4192. https://doi.org/10.1021/acsaem.0c00816
4. Ji Y., Zhang M., Cui J., Lin K., Zheng H., Zhu J., Samia A.C.S. Highly-ordered TiO_2 nanotube arrays with double-walled and bamboo-type structures in dye-sensitized solar cells. *Nano Energy*, 2012, vol. 1, no. 6, pp. 796–804. https://doi.org/10.1016/j.nanoen.2012.08.006
5. Yeung K.L., Yau S.T., Maira A.J., Coronado J.M., Soria J., Yue P.L. The influence of surface properties on the photocatalytic activity of nanostructured TiO_2. *Journal of Catalysis*, 2003, vol. 219, no. 1, pp. 107–116. https://doi.org/10.1016/s0021-9517(03)00187-8
6. Jafari S., Mahyad B., Hashemzadeh H., Janfaza S., Gholikhani T., Tayebi L. Biomedical applications of TiO_2 nanostructures: recent advances. *International Journal of Nanomedicine*, 2020, vol. 15, pp. 3447–3470. https://doi.org/10.2147/ijn.s249441
7. Hoshian S., Jokinen V., Hjort K., Ras R.H.A., Franssila S. Amplified and localized photoswitching of TiO_2 by micro- and nanostructuring. *ACS Applied Materials & Interfaces*, 2015, vol. 7, no. 28, pp. 15593–15599. https://doi.org/10.1021/acsami.5b04309
8. Yavari S.A., van der Stok J., Chai Y.C., Wauthle R., Birgani T.Z., Habibovic P., et al. Bone regeneration performance of surface-treated porous titanium. *Biomaterials*, 2014, vol. 35, no. 24, pp. 6172–6181. https://doi.org/10.1016/j.biomaterials.2014.04.054
9. Yavari S.A., Wauthle R., Böttger A.J., Schrooten J., Weinans H., Zadpoor A.A. Crystal structure and nanotopographical features on the surface of heat-treated and anodized porous titanium biomaterials produced using selective laser melting. *Applied Surface Science*, 2014, vol. 290, pp. 287–294. https://doi.org/10.1016/j.apsusc.2013.11.069
10. Reghunath S., Pinheiro D., Devi K.R.S. A review of hierarchical nanostructures of TiO_2: Advances and applications. *Applied Surface Science Advances*, 2021, vol. 3, pp. 100063. https://doi.org/10.1016/j.apsadv.2021.100063
11. Kaushik V. Negative Index Materials: Metamaterials. *Research Review International Journal of Multidisciplinary*. 2019. V. 4. Issue 4. Available at: http://www.rrjournals.com. (accessed: 25.12.2024)

12. Wang G., Li D., Liao W., Liu T., Li X., An Q., Qu Z. Multifunctional metamaterial with reconfigurable electromagnetic scattering properties for advanced stealth and adaptive applications // Advanced Materials. 2024. V. 36. N 40. P. 2408216. https://doi.org/10.1002/adma.202408216
13. Huang T., Lu J., Zhang X., Xiao R., Yang W., Wu Q. Femtosecond laser fabrication of anatase TiO_2 micro-nanostructures with chemical oxidation and annealing // Scientific Reports. 2017. V. 7. P. 2089. https://doi.org/10.1038/s41598-017-02369-w
14. Варламов П.В., Михайлова Ю.В., Андреева Я.М., Сергеев М.М. Исследование влияния параметров лазерной обработки на спектральные характеристики серебросодержащих пленок диоксида титана // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 5. С. 634–641. https://doi.org/10.17586/2226-1494-2020-20-5-634-641
15. Вейко В.П., Карлагина Ю.Ю., Романов В.В., Яцук Р.М., Егорова Е.Е., Зерницкая Е.А. [и др.]. Лазерная технология структурирования поверхности дентальных титановых имплантатов. Часть 1 // Фотоника. 2020. Т. 14. № 5. С. 462–472. https://doi.org/10.22184/1993-7296.FRos.2020.14.5.462.472
16. Samusev I., Tcibulnikova A., Slezhkin V., Matveeva K., Demin M., Khankaev A., Lyatun I., Bryukhanov V. Transformation of refractive index spectra for titanium rough surfaces // Proceedings of SPIE. 2020. V. 11344. P. 1134422. https://doi.org/10.1117/12.2556727
17. Завестовская И.Н. Лазерное наноструктурирование поверхности материалов // Квантовая электроника. 2010. Т. 40. № 11. С. 942–954.
18. Вейко В.П., Либенсон М.Н. Лазерная обработка. Л.: Лениздат, 1973. 190 с.
19. Ахманов С.А., Емельянов В.И., Воротеев Н.И., Семиногов В.Н. Воздействие мощного лазерного излучения на поверхность полупроводников и металлов: нелинейно-оптические эффекты и нелинейно-оптическая диагностика // Успехи физических наук. 1985. Т. 147. № 12. С. 675–745.
20. Колмычек И.А., Малышева И.В., Новиков В.Б., Майдыковский А.И., Леонтьев А.П., Напольский К.С., Мурзина Т.В. Оптические свойства гиперболических метаматериалов (Миниобзор) // Письма в Журнал экспериментальной и теоретической физики. 2021. Т. 114. № 11-12. С. 727–739. https://doi.org/10.31857/S1234567821230026
21. Purcell E.M. Spontaneous emission probabilities at radio frequencies // Physical Review. 1946. V. 69. P. 681.
22. Климов В.В. Наноплазмоника // Успехи физических наук. 2008. Т. 178. № 8. С. 875–880.
23. Майер С.А. Плазмоника: теория и приложения. М. – Ижевск: НИЦ «Регулярная и хаотическая динамика», 2011. 296 с.
24. Barnes W.L., Dereux A., Ebbesen T.W. Surface plasmon subwavelength optics // Nature. 2003. V. 424. N 6950. P. 824–830. https://doi.org/10.1038/nature01937
25. Kiasat Y., Donato M.G., Hinczewski M., ElKabbash M., Letsou T., Saijaet R., et al. Epsilon-near-zero (ENZ)-based optomechanics // Communications Physics. 2023. V. 6. N 1. P. 69. https://doi.org/10.1038/s42005-023-01186-0
26. Reshef O., De Leon I., Alam M.Z., Boyd R.W. Nonlinear optical effects in epsilon-near-zero media // Nature Reviews Materials. 2019. V. 4. N 8. P. 535–551. https://doi.org/10.1038/s41578-019-0120-5
27. Seetharamdoo D. Resonant negative refractive index metamaterials // Metamaterial. 2012. P. 171–194. https://doi.org/10.5772/35153
28. Padilla W.J., Basov D.N., Smith D.R. Negative refractive index metamaterials // Materials Today. 2006. V. 9. N 7-8. P. 28–35. https://doi.org/10.1016/s1369-7021(06)71573-5
29. Francs G.C.D., Barthes J., Bouhelier A., Weeber J.C., Dereux A., Cuche A., Girard C. Plasmonic Purcell factor and coupling efficiency to surface plasmons. Implications for addressing and controlling optical nanosources // Journal of Optics. 2016. V. 18. N 9. P. 094005. https://doi.org/10.1088/2040-8978/18/9/094005
30. Wang W., Yang X., Gao J. Scaling law of Purcell factor in hyperbolic metamaterial cavities with dipole excitation // Optics Letters. 2019. V. 44. N 3. P. 471–474. https://doi.org/10.1364/ol.44.000471
31. Tan Z.-Q., Lin Q., Du W.-J., Wang L.-L., Liu G.-D. Simultaneously enhance electric and magnetic Purcell factor by strong coupling between toroidal dipole quasi-BIC and electric dipole // Journal of Applied Physics. 2025. V. 137. N 3. P. 033103. https://doi.org/10.1063/5.0251015
32. Давидович М.В. Плазмон-поляритоны Дьяконова вдоль гиперболического метаматериала // Компьютерная оптика. 2021. Т. 45. № 1. С. 48–57. https://doi.org10.18287/2412-6179-CO-673

12. Wang G., Li D., Liao W., Liu T., Li X., An Q., Qu Z. Multifunctional metamaterial with reconfigurable electromagnetic scattering properties for advanced stealth and adaptive applications. *Advanced Materials*, 2024, vol. 36, no. 40, pp. 2408216. https://doi.org/10.1002/adma.202408216
13. Huang T., Lu J., Zhang X., Xiao R., Yang W., Wu Q. Femtosecond laser fabrication of anatase TiO_2 micro-nanostructures with chemical oxidation and annealing. *Scientific Reports*, 2017, vol. 7, pp. 2089. https://doi.org/10.1038/s41598-017-02369-w
14. Varlamov P.V., Mikhailova J.V., Andreeva Ya.M., Sergeev M.M. Effect of laser processing parameters on spectral characteristics of silver-impregnated titanium dioxide thin films. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 5, pp. 634–641. (in Russian). https://doi.org/10.17586/2226-1494-2020-20-5-634-641
15. Veiko V.P., Karlagina YU.YU., Romanov V.V., Yacuk R.M., Egorova E.E., Zernickaya E.A., et al. Laser technology for structuring the surface of dental titanium implants. Part 1. *Photonics Russia*, 2020, vol. 14, no. 5, pp. 462–472. (in Russian). https://doi.org/10.22184/1993-7296.FRos.2020.14.5.462.472
16. Samusev I., Tcibulnikova A., Slezhkin V., Matveeva K., Demin M., Khankaev A., Lyatun I., Bryukhanov V. Transformation of refractive index spectra for titanium rough surfaces. *Proceedings of SPIE*, 2020, vol. 11344, pp. 1134422. https://doi.org/10.1117/12.2556727
17. Zavestovskaya I.N. Laser nanostructuring of materials surfaces. *Quantum Electronics*, 2010, vol. 40, no. 11, pp. 942–954. https://doi.org/10.1070/qe2010v040n11abeh014447
18. Veiko V.P., Libenson M.N. *Laser Processing*. Leningrad, Lenizdat, 1973, 190 p (in Russian)
19. Akhmanov S.A., Emel'Yanov V.I., Koroteev N.I., Seminogov V.N. Interaction of powerful laser radiation with the surfaces of semiconductors and metals: Nonlinear optical effects and nonlinear optical diagnostics. *Soviet Physics Uspekhi*, 1985, vol. 28, no. 12, pp. 1084-1124. https://doi.org/10.1070/PU1985v028n12ABEH003986
20. Kolmychek I.A., Malysheva I.V., Novikov V.B., Maidykovskii A.I., Leontev A.P., Napolskii K.S., Murzina T.V. Optical Properties of Hyperbolic Metamaterials (Brief Review). *JETP Letters*, 2021, vol. 114, no. 11, pp. 653–664. https://doi.org/10.1134/S0021364021230089
21. Purcell E.M. Spontaneous emission probabilities at radio frequencies. *Physical Review*, 1946, vol. 69, pp. 681.
22. Klimov V.V. Nanoplasmonics. *Physics-Uspekhi*, 2008, vol. 51, no. 8, pp. 839–844. https://doi.org/10.1070/PU2008v051n08ABEH006794
23. Maier S.A. *Plasmonics: Fundamentals and Applications*. Springer, 2007. 250 p.
24. Barnes W.L., Dereux A., Ebbesen T.W. Surface plasmon subwavelength optics. *Nature*, 2003, vol. 424, no. 6950, pp. 824–830. https://doi.org/10.1038/nature01937
25. Kiasat Y., Donato M.G., Hinczewski M., ElKabbash M., Letsou T., Saijaet R., et al. Epsilon-near-zero (ENZ)-based optomechanics. *Communications Physics*, 2023, vol. 6, no. 1, pp. 69. https://doi.org/10.1038/s42005-023-01186-0
26. Reshef O., De Leon I., Alam M.Z., Boyd R.W. Nonlinear optical effects in epsilon-near-zero media. *Nature Reviews Materials*, 2019, vol. 4, no. 8, pp. 535–551. https://doi.org/10.1038/s41578-019-0120-5
27. Seetharamdoo D. Resonant negative refractive index metamaterials. *Metamaterial*, 2012, pp. 171–194. https://doi.org/10.5772/35153
28. Padilla W.J., Basov D.N., Smith D.R. Negative refractive index metamaterials. *Materials Today*, 2006, vol. 9, no. 7-8, pp. 28–35. https://doi.org/10.1016/s1369-7021(06)71573-5
29. Francs G.C.D., Barthes J., Bouhelier A., Weeber J.C., Dereux A., Cuche A., Girard C. Plasmonic Purcell factor and coupling efficiency to surface plasmons. Implications for addressing and controlling optical nanosources. *Journal of Optics*, 2016, vol. 18, no. 9, pp. 094005. https://doi.org/10.1088/2040-8978/18/9/094005
30. Wang W., Yang X., Gao J. Scaling law of Purcell factor in hyperbolic metamaterial cavities with dipole excitation. *Optics Letters*, 2019, vol. 44, no. 3, pp. 471–474. https://doi.org/10.1364/ol.44.000471
31. Tan Z.-Q., Lin Q., Du W.-J., Wang L.-L., Liu G.-D. Simultaneously enhance electric and magnetic Purcell factor by strong coupling between toroidal dipole quasi-BIC and electric dipole. *Journal of Applied Physics*, 2025, vol. 137, no. 3, pp. 033103. https://doi.org/10.1063/5.0251015
32. Davidovich M.V. Dyakonov plasmon-polaritones along a hyperbolic metamaterial surface. *Computer Optics*, 2021, vol. 45, no. 1, pp. 48–57. (in Russian). https://doi.org10.18287/2412-6179-CO-673

33. Wu S., Luo X., Long Y., Xu B. Exploring the phase transformation mechanism of titanium dioxide by high temperature in situ method // IOP Conference Series: Materials Science and Engineering. 2019. V. 493. N 1. P. 012010. https://doi.org/10.1088/1757-899x/493/1/012010
34. Mett R.R., Sidabras J.W., Hyde J.S. Radio frequency skin depth concepts in magnetic resonance // Current Topics in Biophysics. 2005. V. 29. N 1-2. P. 83–88.
35. Бродский А.М., Урбах М.И. Оптика шероховатых поверхностей металлов // Журнал экспериментальной и теоретической физики. 1985. Т. 89. № 2. С. 680–695.
36. Ordal M.A., Bell R.J., Alexander R.W., Long L.L., Querry M.R. Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V, and W // Applied Optics. 1985. V. 24. N 24. P. 4493–4499. https://doi.org/10.1364/ao.24.004493
37. Forouhi A.R., Bloomer I. Optical dispersion relations for amorphous semiconductors and amorphous dielectrics // Physical Review B. 1986. V. 34. N 10. P. 7018–7026. https://doi.org/10.1103/physrevb.34.7018
38. Mohamad M., Haq B.U., Ahmed R., Shaari A., Ali N., Hussain R. A density functional study of structural, electronic and optical properties of titanium dioxide: Characterization of rutile, anatase and brookite polymorphs // Materials Science in Semiconductor Processing. 2015. V. 31. P. 405–414. https://doi.org/10.1016/j.mssp.2014.12.027
39. Al-Mudhaffer M.F. Optical properties, interband transition strength and the surface, volume energy loss function of titanium dioxide film // Journal of Basrah Researches (Sciences). 2010. V. 36. N 6. P. 31–38.
40. Li A.Y., Yang Y., Shu X., Wan D., Wei N., Yu X., et al. From titanium sesquioxide to titanium dioxide: oxidation-induced structural, phase, and property evolution // Chemistry of Materials. 2018. V. 30. N 13. P. 4383–4392. https://doi.org/10.1021/acs.chemmater.8b01739
41. Sokolowski-Tinten K., von der Linde D. Generation of dense electron-hole plasmas in silicon // Physical Review B. 2000. V. 61. N 4. P. 2643–2650. https://doi.org/10.1103/physrevb.61.2643
42. Sun J., Yang Y., Khan J.I., Alarousu E., Guo Z., Zhang X., et al. Ultrafast carrier trapping of a metal-doped titanium dioxide semiconductor revealed by femtosecond transient absorption spectroscopy // ACS Applied Materials & Interfaces. 2014. V. 6. N 13. P. 10022–10027. https://doi.org/10.1021/am5026159
43. Highfield J.G., Graetzel M. Discovery of reversible photochromism in titanium dioxide using photoacoustic spectroscopy: implications for the investigation of light-induced charge-separation and surface redox processes in titanium dioxide // The Journal of Physical Chemistry. 1988. V. 92. N 2. P. 332–333. https://doi.org/10.1021/j100313a043
44. Huy A.H., Aradi B., Frauenheim T., Deák P. Calculation of carrier-concentration-dependent effective mass in Nb-doped anatase crystals of TiO_2 // Physical Review B. 2011. V. 83. N 15. P. 155201. https://doi.org/10.1103/physrevb.83.155201

33. Wu S., Luo X., Long Y., Xu B. Exploring the phase transformation mechanism of titanium dioxide by high temperature in situ method. *IOP Conference Series: Materials Science and Engineering*, 2019, vol. 493, no. 1, pp. 012010. https://doi.org/10.1088/1757-899x/493/1/012010
34. Mett R.R., Sidabras J.W., Hyde J.S. Radio frequency skin depth concepts in magnetic resonance. *Current Topics in Biophysics*, 2005, vol. 29, no. 1-2, pp. 83–88.
35. Brodskiiand A.M., Urbakh M.I. Optics of rough surfaces of metals. *Soviet Physics — JETP*, 1985, vol. 62, no. 2, pp. 391–399.
36. Ordal M.A., Bell R.J., Alexander R.W., Long L.L., Querry M.R. Optical properties of fourteen metals in the infrared and far infrared: Al, Co, Cu, Au, Fe, Pb, Mo, Ni, Pd, Pt, Ag, Ti, V, and W. *Applied Optics*, 1985, vol. 24, no. 24, pp. 4493–4499. https://doi.org/10.1364/ao.24.004493
37. Forouhi A.R., Bloomer I. Optical dispersion relations for amorphous semiconductors and amorphous dielectrics. *Physical Review B*, 1986, vol. 34, no. 10, pp. 7018–7026. https://doi.org/10.1103/physrevb.34.7018
38. Mohamad M., Haq B.U., Ahmed R., Shaari A., Ali N., Hussain R. A density functional study of structural, electronic and optical properties of titanium dioxide: Characterization of rutile, anatase and brookite polymorphs. *Materials Science in Semiconductor Processing*, 2015, vol. 31, pp. 405–414. https://doi.org/10.1016/j.mssp.2014.12.027
39. Al-Mudhaffer M.F. Optical properties, interband transition strength and the surface, volume energy loss function of titanium dioxide film. *Journal of Basrah Researches (Sciences)*, 2010, vol. 36, no. 6, pp. 31–38.
40. Li A.Y., Yang Y., Shu X., Wan D., Wei N., Yu X., et al. From titanium sesquioxide to titanium dioxide: oxidation-induced structural, phase, and property evolution. *Chemistry of Materials*, 2018, vol. 30, no. 13, pp. 4383–4392. https://doi.org/10.1021/acs.chemmater.8b01739
41. Sokolowski-Tinten K., von der Linde D. Generation of dense electron-hole plasmas in silicon. *Physical Review B*, 2000, vol. 61, no. 4, pp. 2643–2650. https://doi.org/10.1103/physrevb.61.2643
42. Sun J., Yang Y., Khan J.I., Alarousu E., Guo Z., Zhang X., et al. Ultrafast carrier trapping of a metal-doped titanium dioxide semiconductor revealed by femtosecond transient absorption spectroscopy. *ACS Applied Materials & Interfaces*, 2014, vol. 6, no. 13, pp. 10022–10027. https://doi.org/10.1021/am5026159
43. Highfield J.G., Graetzel M. Discovery of reversible photochromism in titanium dioxide using photoacoustic spectroscopy: implications for the investigation of light-induced charge-separation and surface redox processes in titanium dioxide. *The Journal of Physical Chemistry*, 1988, vol. 92, no. 2, pp. 332–333. https://doi.org/10.1021/j100313a043
44. Huy A.H., Aradi B., Frauenheim T., Deák P. Calculation of carrier-concentration-dependent effective mass in Nb-doped anatase crystals of TiO_2. *Physical Review B*, 2011, vol. 83, no. 15, pp. 155201. https://doi.org/10.1103/physrevb.83.155201

Авторы

Кострина Алёна Андреевна — техник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, sc 59762280900, https://orcid.org/0009-0002-0943-9290, aakostrina@gmail.com

Цибульникова Анна Владимировна — кандидат физико-математических наук, старший научный сотрудник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, sc 57193169565, https://orcid.org/0000-0001-8578-0701, memorgold@mail.ru

Слежкин Василий Анатольевич — кандидат химических наук, доцент, старший научный сотрудник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, sc 6506061436, https://orcid.org/0000-0002-2801-7029, vasiliy.slezhkin@klgtu.ru

Лятун Иван Игоревич — старший научный сотрудник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, sc 56600612300, https://orcid.org/0000-0002-4988-8077, ILyatun@kantiana.ru

Царьков Максим Владимирович — технический руководитель Астрономического сообщества, инженер лаборатории астрофизики, Балтийский федеральный университет имени Иммануила Канта,

Authors

Alena A. Kostrina — Technician, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, sc 59762280900, https://orcid.org/0009-0002-0943-9290, aakostrina@gmail.com

Anna V. Tcibulnikova — PhD (Physics & Mathematics), Senior Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, sc 57193169565, https://orcid.org/0000-0001-8578-0701, memorgold@mail.ru

Vasily A. Slezhkin — PhD (Chemistry), Senior Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, sc 6506061436, https://orcid.org/0000-0002-2801-7029, vasiliy.slezhkin@klgtu.ru

Ivan I. Lyatun — Senior Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, sc 56600612300, https://orcid.org/0000-0002-4988-8077, ILyatun@kantiana.ru

Maksim V. Tsarkov — Technical Director of the Astronomical Society, Engineer of the Astrophysics Laboratory, Immanuel Kant Baltic Federal

Калининград, 236041, Российская Федерация, https://orcid.org/0009-0000-5345-4223, max.offreal@gmail.com
Артамонов Дмитрий Александрович — техник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, https://orcid.org/0000-0003-2141-8900, euroset2016ig98@icloud.com
Курицкий Михаил Сергеевич — лаборант-исследователь, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, https://orcid.org/0000-0002-4125-4678, Mix-Mux@yandex.ru
Самусев Илья Геннадьевич — кандидат физико-математических наук, доцент, директор Научно-образовательного центра «Фундаментальная и прикладная фотоника. Нанофотоника»; Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, sc 12779220200, https://orcid.org/0000-0001-5026-7510, ISamusev@kantiana.ru
Брюханов Валерий Вениаминович — доктор физико-математических наук, профессор, ведущий научный сотрудник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, sc 7003848491, https://orcid.org/0000-0003-4689-7207, Bryukhanov_v.v@mail.ru

University, Kaliningrad, 236041, Russian Federation, https://orcid.org/0009-0000-5345-4223, max.offreal@gmail.com
Dmitry A. Artamonov — Technician, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, https://orcid.org/0000-0003-2141-8900, euroset2016ig98@icloud.com

Michael S. Kuritskij — Laboratory Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, https://orcid.org/0000-0002-4125-4678, Mix-Mux@yandex.ru

Ilia G. Samusev — PhD (Physics & Mathematics), Associate Professor, Head of the Research & Educational Center “Fundamental and Applied Photonics. Nanophotonics”, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, sc 12779220200, https://orcid.org/0000-0001-5026-7510, ISamusev@kantiana.ru

Valery V. Bryukhanov — D.Sc. (Psysics & Mathematics), Professor, Leading Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, sc 7003848491, https://orcid.org/0000-0003-4689-7207, Bryukhanov_v.v@mail.ru

Статья поступила в редакцию 19.08.2025
Одобрена после рецензирования 20.10.2025
Принята к печати 26.11.2025

Received 19.08.2025
Approved after reviewing 20.10.2025
Accepted 26.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2025 Том 25 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2025 Vol. 25 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

НОВЫЕ МАТЕРИАЛЫ И НАНОТЕХНОЛОГИИ

MATERIAL SCIENCE AND NANOTECHNOLOGIES

doi: 10.17586/2226-1494-2025-25-6-1047-1057
УДК 538.958

Фотофизические, оптические и люминесцентные характеристики гетероциклически замещенных кумаринов и их применение в OLED-устройствах

Анна Владимировна Осадченко[1]✉, Сергей Александрович Амброзевич[2], Иван Александрович Захарчук[3], Андрей Александрович Ващенко[4], Даниил Саюзович Дайбаге[5], Леонид Игоревич Кармазин[6], Дмитрий Андреевич Чепцов[7], Валерий Федорович Травень[8], Алина Романовна Слюсаренко[9], Александр Сергеевич Селюков[10]

[1,2,3,4,5,6,10] Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация
[7,8,9] Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация
[10] Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация
[1,10] Российский технологический университет МИРЭА, Москва, 119454, Российская Федерация
[10] Всероссийский институт научной и технической информации РАН, Москва, 125190, Российская Федерация

[1] anna.vl.osadchenko@gmail.com✉, https://orcid.org/0000-0001-9556-4885
[2] s.ambrozevich@mail.ru, https://orcid.org/0000-0002-3906-0735
[3] zakharchukwork@mail.ru, https://orcid.org/0000-0002-1502-6460
[4] vashchenkoaa@lebedev.ru, https://orcid.org/0000-0003-2084-590
[5] daibagya@mail.ru, https://orcid.org/0000-0003-1944-1546
[6] leokarm64@gmail.com, https://orcid.org/0009-0009-3066-3474
[7] dchepcov@yandex.ru, https://orcid.org/0000-0002-9774-7922
[8] valerii.traven@gmail.com, https://orcid.org/0000-0002-2204-7438
[9] alilisha.2006@mail.ru, https://orcid.org/0009-0003-6868-2503
[10] selyukov@lebedev.ru, https://orcid.org/0000-0002-4007-6291

Аннотация
Введение. Развитие органической электроники стимулирует поиск новых материалов. Приоритетной задачей является нахождение соединений, обладающих высокой яркостью, эффективностью и стабильностью свечения. Кумариновые производные рассматриваются как перспективные кандидаты для решения этой задачи. Представлены результаты исследования органических светоизлучающих диодов (Organic Light-Emitting Diode, OLED), в эмиссионном слое которых использован ряд кумариновых красителей с выраженными донорно-акцепторными свойствами. Исследовано влияние структуры синтезированных молекул на фотофизические характеристики, а также на эффективность излучения светодиодов на их основе. **Метод.** Синтезирована серия органических соединений кумаринового ряда: (E)-3-(3-(антрацен-9-ил)акрилоил)кумарин (соединение 1), 4-гидрокси-3-(5-(4-метоксифенил)-1-(п-толил)-4,5-дигидро-1H-пиразол-3-ил)кумарин (соединение 2), 3-(1-ацетил-5-(4-метоксифенил)-4,5-дигидро-1H-пиразол-3-ил)-4-гидроксикумарин (соединение 3), этил 7-(диэтиламино)кумарин-3-карбоксилат (соединение 4), а также известный лазерный краситель Кумарин 6 (3-(бензо[d]тиазол-2-ил)-7-(диэтиламино)кумарин) (соединение Cou), используемый в качестве эталонного соединения. Светодиоды были изготовлены методом термического напыления в вакууме и методом центрифугирования. Исследование спектров флуоресценции и электролюминесценции проводилось при помощи спектрометра Ocean Optics Maya 200 PRO. Для получения кривых затухания люминесценции применен фотоэлектронный умножитель PicoQuant PMA-C 192-N-M. **Основные результаты.** Показано, что спектральные данные (поглощение, фотолюминесценция), а также время-разрешенные измерения (время затухания флуоресценции) указывают на ключевую роль донорно-акцепторных взаимодействий, а также пространственных

© Осадченко А.В., Амброзевич С.А., Захарчук И.А., Ващенко А.А., Дайбаге Д.С., Кармазин Л.И., Чепцов Д.А., Травень В.Ф., Слюсаренко А.Р., Селюков А.С., 2025

эффектов в формировании электронных переходов. Вольт-амперные характеристики подтвердили наличие режимов проводимости, ограниченной пространственным зарядом, и проводимости, ограниченной процессами захвата носителей. Изучение вольт-яркостных характеристик показало, что соединение 2 демонстрирует яркость, сопоставимую с соединением Cou, что делает его наиболее перспективным для дальнейшей оптимизации органических светоизлучающих диодов. Обнаружено, что соединение 4 в составе светодиода обеспечивает белое излучение с координатами цветности, близкими к дневному свету, что делает потенциально возможным его практическое применение в осветительных системах. **Обсуждение.** Полученные данные подтверждают влияние донорно-акцепторных взаимодействий на свойства кумаринов. Степень сопряжения донорных и акцепторных фрагментов напрямую определяют спектральные сдвиги в спектрах поглощения и флуоресценции. Высокая яркость светодиодов на основе соединения 2, сопоставимая с соединением Cou, обусловлена его эффективной донорно-акцепторной системой, оптимизирующей внутримолекулярный перенос заряда и понижающей вероятность безызлучательных переходов. Напротив, ацетильная группа в соединении 3 нарушает конъюгацию, приводя к низкой яркости и малому времени жизни флуоресценции (1,7 нс) из-за безызлучательной релаксации. Способность соединения 4 обеспечивать в диодах белое излучение (цветовая температура около 6410 К, близка к дневному свету) связана с вкладом электронного транспортного слоя в спектр излучения.

Ключевые слова
фотолюминесценция, электролюминесценция, органические светоизлучающие диоды, кумариновые красители, взаимосвязь структура-свойства

Ссылка для цитирования: Осадченко А.В., Амброзевич С.А., Захарчук И.А., Ващенко А.А., Дайбаге Д.С., Кармазин Л.И., Чепцов Д.А., Травень В.Ф., Слюсаренко А.Р., Селюков А.С. Фотофизические, оптические и люминесцентные характеристики гетероциклически замещенных кумаринов и их применение в OLED-устройствах // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1047–1057. doi: 10.17586/2226-1494-2025-25-6-1047-1057

Photophysical, optical and luminescent characteristics of heterocyclic-substituted coumarins and their application in OLED-devices

Anna V. Osadchenko[1]✉, Sergey A. Ambrozevich[2], Ivan A. Zakharchuk[3], Andrey A. Vashchenko[4], Daniil S. Daibagya[5], Leonid I. Karmazin[6], Dmitry A. Cheptsov[7], Valery F. Traven[8], Alina R. Slyusarenko[9], Alexandr S. Selyukov[10]

[1,2,3,4,5,6,10] Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation
[7,8,9] Mendeleev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation
[10] Bauman Moscow State Technical University, Moscow, 105005, Russian Federation
[1,10] Russian Technological University MIREA, Moscow, 119454, Russian Federation
[10] All Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences, Moscow, 125190, Russian Federation

[1] anna.vl.osadchenko@gmail.com✉, https://orcid.org/0000-0001-9556-4885
[2] s.ambrozevich@mail.ru, https://orcid.org/0000-0002-3906-0735
[3] zakharchukwork@mail.ru, https://orcid.org/0000-0002-1502-6460
[4] vashchenkoaa@lebedev.ru, https://orcid.org/0000-0003-2084-590
[5] daibagya@mail.ru, https://orcid.org/0000-0003-1944-1546
[6] leokarm64@gmail.com, https://orcid.org/0009-0009-3066-3474
[7] dchepcov@yandex.ru, https://orcid.org/0000-0002-9774-7922
[8] valerii.traven@gmail.com, https://orcid.org/0000-0002-2204-7438
[9] alilisha.2006@mail.ru, https://orcid.org/0009-0003-6868-2503
[10] selyukov@lebedev.ru, https://orcid.org/0000-0002-4007-6291

Abstract
The development of organic electronics stimulates the search for new materials. The priority task is to find compounds with high brightness, efficiency and stability of luminescence. Coumarin derivatives are considered as promising candidates for solving this problem. This paper presents the results of a study of organic light-emitting diodes in the emission layer of which a number of coumarin dyes with pronounced donor-acceptor properties are used. The aim of the study was to identify the influence of the structure of synthesized molecules on the photophysical characteristics as well as on the emission efficiency of LEDs based on them. A series of organic compounds of the coumarin series has been synthesized: (E)-3-(3-(anthracene-9-yl)acryloyl)coumarin (compound 1), 4-hydroxy-3-(5-(4-methoxyphenyl)-1-(p-tolyl)-4,5-dihydro-1H-pyrazol-3-yl)coumarin (compound 2), 3-(1-acetyl-5-(4-methoxyphenyl)-4,5-dihydro-1H-pyrazol-3-yl)-4-hydroxycoumarin (compound 3), ethyl 7-(diethylamino)Coumarin-3-carboxylate (compound 4) as well as the well-known laser dye Coumarin 6 (3-(benzo[d]thiazole-2-yl)-7-(diethylamino)coumarin) used as a reference compound. The LEDs were produced by vacuum thermal spraying and spin-coating. The fluorescence and electroluminescence spectra were studied using an Ocean Optics Maya 200 PRO spectrometer. A photomultiplier was used to obtain the luminescence decay curves, PicoQuant PMA-C 192-N-M. Spectral data (absorption, photoluminescence) as well as time-resolved measurements (fluorescence attenuation time) indicate the key role of donor-acceptor interactions as well as spatial effects in the formation of electronic transitions. The current-voltage characteristics confirmed the presence of conduction modes limited by spatial charge and conduction limited by carrier capture processes. The study of the voltage-brightness characteristics showed that compound 2 demonstrates brightness comparable to the reference compound Coumarin 6, which makes it the most promising for further optimization of organic light-emitting diodes.

In addition, it is shown that compound 4 in the device provides white emission with chromaticity coordinates close to daylight, which makes it potentially possible for practical use in lighting systems. The data obtained confirm the influence of donor-acceptor interactions on the properties of coumarins. The degree of conjugation of donor and acceptor fragments is directly determined by spectral shifts in the absorption and fluorescence spectra. The high brightness of compound 2-based diodes, comparable to the Cou standard, is due to its efficient donor-acceptor system which optimizes intramolecular charge transfer and increases the likelihood of radiative transitions (long lifetime of the excited state 3.5 ns). On the contrary, the acetyl group in compound 3 disrupts conjugation, leading to low brightness and short fluorescence lifetime (1.7 ns) due to nonradiative relaxation. The ability of compound 4 to provide white radiation in diodes (correlated color temperature 6410 K, close to daylight) is related to the contribution of the electron transport layer to the emission spectrum.

Keywords
photoluminescence, electroluminescence, organic light emitting diodes, coumarin dyes, structure-properties relationship

For citation: Osadchenko A.V., Ambrozevich S.A., Zakharchuk I.A., Vashchenko A.A., Daibagya D.S., Karmazin L.I., Cheptsov D.A., Traven V.F., Slyusarenko A.R., Selyukov A.S. Photophysical, optical and luminescent characteristics of heterocyclic-substituted coumarins and their application in OLED-devices. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1047–1057 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1047-1057

Введение

Современные технологии предъявляют высокие требования к энергоэффективности и экологической безопасности источников освещения и средств отображения информации. В этом контексте органические светоизлучающие диоды (Organic Light-Emitting Diode, OLED) [1–6] представляют собой одно из наиболее перспективных направлений, предлагая существенные преимущества по сравнению с традиционными источниками света, такими как лампы накаливания и люминесцентные лампы. OLED отличаются высокой энергоэффективностью, компактностью, гибкостью и большим углом обзора экранов на их основе, что не только снижает энергопотребление, но и позволяет создавать дисплеи для различных электронных устройств. Кроме того, в отличие от жидкокристаллических экранов, OLED-дисплеи обеспечивают глубокий черный цвет благодаря возможности полного отключения отдельных пикселов, что существенно повышает контрастность изображения. Дополнительными преимуществами являются отсутствие токсичных компонентов, характерных для люминесцентных ламп, а также возможность изготовления гибких и изогнутых панелей, что расширяет как дизайнерские, так и технологические возможности их применения.

Несмотря на достигнутый прогресс, исследования в области OLED продолжаются как с целью создания более эффективных эмиттеров, так и с целью повышения энергоэффективности устройств на их основе, увеличения срока службы таких устройств и улучшения качества цветопередачи. Одним из направлений является применение металлогалогеновых перовскитов (PeLED) в качестве излучающего слоя. В составе светодиодов эти материалы демонстрируют выдающуюся внешнюю квантовую эффективность, достигающую более 20 % для красного [7–9] и около 10 % для синего и зеленого излучений [10–12], а также обеспечивают высокую насыщенность создаваемых цветов за счет узких линий излучения и возможности точной регулировки спектральных характеристик. Помимо этого, значение для внутреннего квантового выхода люминесценции приближается к теоретическому максимуму [7, 9]. Однако широкое применение PeLED ограничено такими факторами, как токсичность, низкая стабильность при воздействии влаги и кислорода, а также сложности в обеспечении длительного срока службы устройств [13–15].

Для преодоления указанных ограничений и повышения эффективности OLED активно исследуются новые классы органических материалов, способных эффективно преобразовывать электрическую энергию в световое излучение. Особое внимание уделяется соединениям кумарина — ряду ароматических соединений, характеризующихся высоким квантовым выходом флуоресценции, значительным стоксовым сдвигом и стабильностью излучения [16–19]. Фотофизические свойства производных кумарина тесно связаны с их молекулярной структурой и степенью сопряжения с донором или акцептором электронов, присоединенным к кумариновому ядру [20]. Гибкость в модификации молекулярной структуры кумаринов позволяет тонко настраивать их оптические и электронные свойства, что делает их перспективными для использования в излучающих слоях OLED. Оптимизация фотофизических характеристик кумаринов также способствует повышению эффективности и долговечности светодиодных устройств.

В настоящей работе представлены исследования OLED, основанных на эмиттерах кумаринового ряда. Проведен анализ влияния молекулярной структуры соединений на фотофизические характеристики люминофоров, а также электролюминесцентные и электрофизические характеристики светодиодов на их основе.

Методика проведения эксперимента

В качестве объектов исследования были рассмотрены следующие материалы: (E)-3-(3-(антрацен-9-ил) акрилоил)кумарин (соединение 1, рис. 1, *a*); 4-гидрокси-3-(5-(4-метоксифенил)-1-(п-толил)-4,5-дигидро-1H-пиразол-3-ил)кумарин (соединение 2, рис. 1, *b*), 3-(1-ацетил-5-(4-метоксифенил)-4,5-дигидро-1H-пиразол-3-ил)-4-гидроксикумарин (соединение 3, рис. 1, *c*), этил 7-(диэтиламино)кумарин-3-карбоксилат (соединение 4, рис. 1, *d*), 3-(бензо[d]тиазол-2-ил)-7-(диэтиламино)кумарин (соединение Cou, рис. 1, *e*) — лазерный краситель Кумарин 6.

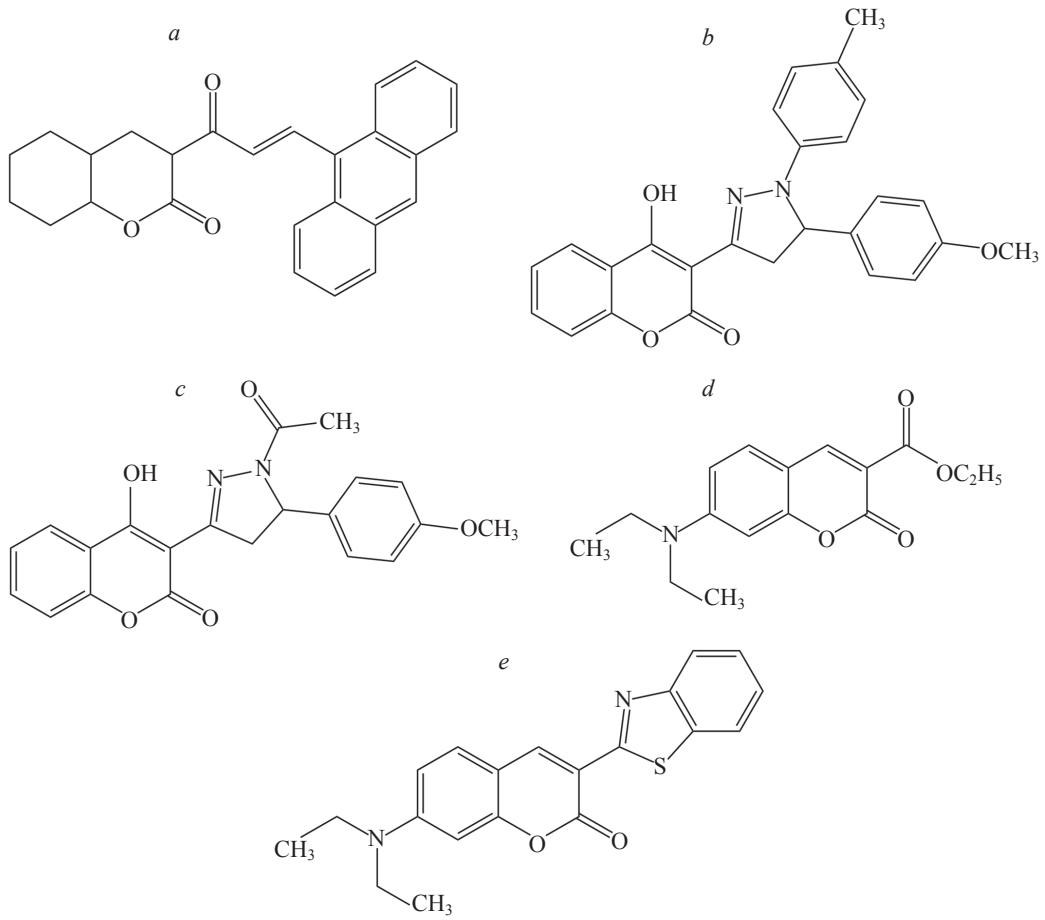

Рис. 1. Структуры исследуемых химических соединений: соединение 1 (*a*); соединение 2 (*b*); соединение 3 (*c*); соединение 4 (*d*); соединение Cou (*e*)

Fig. 1. The structures of the studied chemical compounds: compound 1 (*a*); compound 2 (*b*); compound 3 (*c*); compound 4 (*d*); compound Cou (*e*)

Синтез соединений 1–4 и Cou осуществлялся по методикам, описанным в работах [21–25].

Спектры поглощения были получены с помощью спектрофотометра PerkinElmer LAMBDA 45 с рабочим диапазоном длин волн 190–1100 нм для растворов исследуемых соединений в толуоле с концентрацией порядка 2,5 мг/л. Спектры флуоресценции были получены для растворов соединений в толуоле с концентрацией около 1 мг/л. Измерения спектров фото- и электролюминесценции проводились при помощи спектрометра OceanOptics Maya 2000 Pro, чувствительного в диапазоне 200–1100 нм. Кривые затухания люминесценции были получены при помощи фотоэлектронного умножителя (ФЭУ) PicoQuant PMA-C 192-N-M, чувствительного от 250 до 850 нм, работающего в режиме счета одиночных фотонов. ФЭУ был подключен к плате счета одиночных фотонов PicoQuant TimeHarp 100. В качестве источника возбуждения использовался импульсный лазер PicoQuant LDH-C 400, излучающий на λ = 405 нм с длительностью импульсов 75 пс. При измерении спектров флуоресценции частота повторения импульсов составляла 40 МГц, а средняя мощность излучения — 5 мВт, при исследовании кривых затухания при помощи внешнего генератора задавалась частота повторения импульсов лазера 1 МГц. Измерение характеристик OLED проводилось в перчаточном боксе, заполненном аргоном. Вольт-амперные характеристики регистрировались при помощи двух цифровых мультиметров DT 838. Вольт-яркостные характеристики были получены при помощи яркомера ТКА-ПКМ.

Создание OLED проводилось аналогично методике, описанной в работе [6]. Архитектура полученных светодиодов имела следующий вид: оксид индия-олова (ITO)/поли(3,4-этилендиокситиофен) полистирол сульфонат (PEDOT:PSS)/поли[N,N′-бис(4-бутилфенил)-N,N′-бис(фенил)-бензидин] (poly-TPD)/краситель/2,2′,2″-(1,3,5-бензинитрил)-трис(1-фенил-1-H-бензимидазол) (TPBi)/LiF/Al.

Результаты и их обсуждение

Для соединения 1 (рис. 2) в спектрах поглощения наблюдаются максимумы на λ_{11} = 350,1 нм, λ_{12} = 367,1 нм, λ_{13} =389,6 нм, для соединения 2 — λ_{21} = 440,1 нм и λ_{22} = 460,4 нм, для соединения 3 — λ_{31} = 358,2 нм, λ_{32} = = 373,1 нм и λ_{33} = 393,0 нм, у соединения 4 наблюдается максимум на λ_4 = 412,5 нм, а для соединения Cou — на λ_{cou1} = 439,4 нм и λ_{cou2} = 459,6 нм. Ширина полосы поглощения на полувысоте для соединения 2 составила 63 нм, для соединения 3 — 68 нм, для соединения 4 —

47 нм и для соединения Cou — 63 нм. Для соединения 1 максимумы на λ_{11} и λ_{12} обусловлены π–π*-переходами в сопряженных бензопироновых системах. Однако максимум на λ_{13}, вероятно, возникает за счет вклада массивного антраценового фрагмента или локальных электронных переходов в кумариновом ядре [21]. Для соединения 3 основной вклад в спектр поглощения связан с кумариновым фрагментом (максимумы с λ_{31}, λ_{32}) с характерными π–π*-переходами в сопряженных системах, причем максимум с λ_{33}, вероятно, связан с вкладом ацетильной группы, прикрепленной к пиразольному фрагменту. У спектров поглощения соединений 2, 4 и Cou наблюдается батохромный сдвиг. Причем, у спектров поглощения соединений 4 и Cou отмечены существенные различия, несмотря на наличие у них диэтиламиновой группы. Они обусловлены, в первую очередь, наличием эфирной группы у соединения 4, которая хоть и является акцептором, но гораздо более слабым, чем бензотиазол у соединения Cou [18]. Следовательно, получена большая разница в коэффициентах экстинкции исследуемых соединений. В то же время, у соединений 2 и Cou более схожий вид кривой, что обусловлено превалированием общего хромофорного фрагмента в их структурах. Однако донорно-акцепторные взаимодействия у соединения 2 в этом случае выражены гораздо слабее, по сравнению с соединением Cou, что обусловлено пиразольным фрагментом, обладающим слабыми акцепторными свойствами. Такое различие в электронном взаимодействии напрямую влияет на спектральные характеристики, включая значения молярного коэффициента экстинкции.

Спектры фотолюминесценции исследуемых соединений представлены на рис. 3. Для соединения 1 присутствует пик с максимумом на λ_1 = 570 нм, для соединения 2 — на λ_2 = 544 нм, для соединения 3 — на λ_{31} = 441 нм и λ_{32} = 468 нм, для соединения 4 — на λ_4 = 441 нм, для соединения Cou — на λ_{cou1} = 482 нм и λ_{cou2} = 511 нм. Для соединения 2 стоксов сдвиг был незначительным, около 1 нм, что обусловлено жесткой структурой молекулы, у которой дипольный момент практически не меняется. Для соединения 3 сдвиг составил около 101 нм, что обусловлено наличием ацетильной группы, являющейся слабым акцептором, и пиразольного цикла, являющегося слабым донором, что приводит к реорганизации электронной плотности в возбужденном состоянии [26]. Для соединений 4 и Cou стоксов сдвиг составил — 22 нм и 28 нм, что может быть объяснено взаимодействием растворителя и красителя [27]. Для соединения 1 стоксов сдвиг оказался гораздо большим — 186 нм, что объясняется тем, что электронная плотность преимущественно локализована на антраценовом фрагменте. Наличие двух максимумов в спектре излучения соединений Cou и 3 обусловлено образованием димеров [28]. Было обнаружено, что у соединения 2, по сравнению с соединением 3, несмотря на схожую структуру молекулы, наблюдается значительный батохромный сдвиг (около 100 нм), обусловленный вкладом ацетильной группы, которая в данном случае не усиливает, а скорее частично ограничивает эффективную конъюгацию, повышая энергию перехода [28]. Помимо этого, в спектре флуоресценции для соединения 4, по сравнению с соединением Cou, наблюдается значительный гипсохромный сдвиг (около 40 нм), что обусловлено заменой карбоксильного фрагмента на более сильный акцептор (бензотиазол), что расширяет π-систему и уменьшает энергетический зазор.

Кривые затухания флуоресценции (рис. 4) демонстрируют достаточно простое поведение. Их аппроксимация проводилась моноэкспоненциальной функцией $A \cdot e^{-\frac{t}{\tau}}$. Природа процесса для всех трех соединений

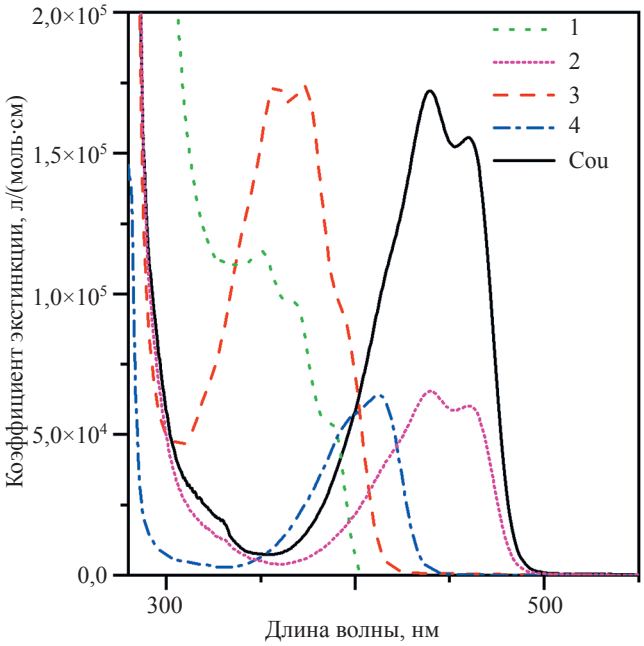

Рис. 2. Спектры поглощения исследуемых соединений 1–4 и Cou

Fig. 2. Absorption spectra of the studied compounds 1–4 and Cou

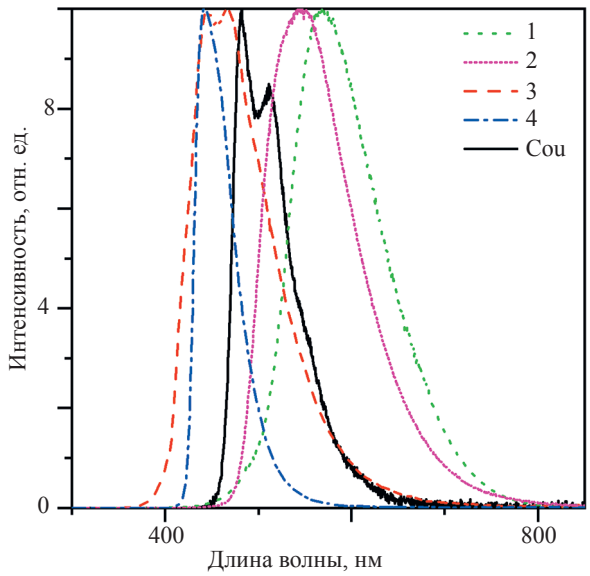

Рис. 3. Спектры флуоресценции исследуемых соединений 1–4 и Cou

Fig. 3. Fluorescence spectra of the studied compounds 1–4 and Cou

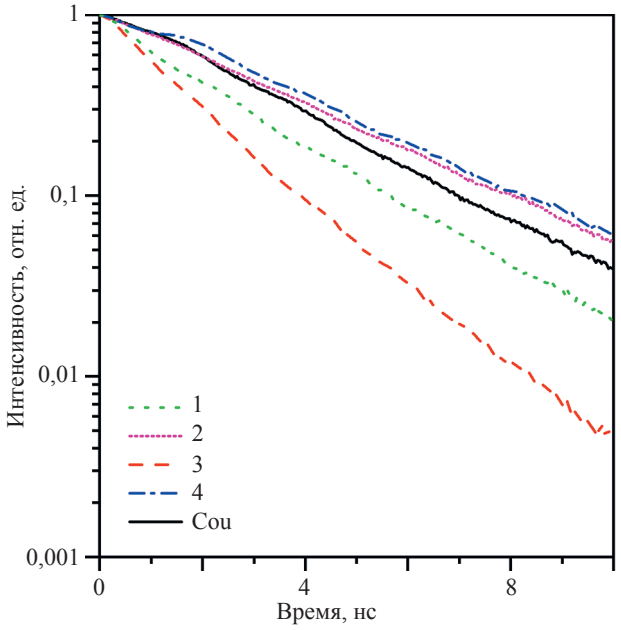

Рис. 4. Кривые затухания люминесценции исследуемых соединений 1–4 и Cou

Fig. 4. Fluorescence spectra of the studied compounds 1–4 and Cou

одна: переход $S_1 \rightarrow S_0$. Были получены характерные времена затухания люминесценции исследуемых соединений кумаринового ряда. Различия во временах затухания флуоресценции для исследуемых соединений 1–4 (τ_1 = 2,5 нс, τ_2 = 3,5 нс, τ_3 = 1,7 нс, τ_4 = 3,8 нс) и Cou (τ_{cou} = 3,1 нс) объясняются балансом между излучательной (k_r) и безызлучательной (k_{nr}) релаксациями возбужденного состояния. Для соединений 1 и Cou наблюдаются характерные времена жизни флуоресценции, типичные для соединений на основе кумарина [29]. Наиболее короткое время жизни флуоресценции было выявлено для соединения 3, что обусловлено влиянием ацетильной группы. Эта группа способствует усилению процессов безызлучательной релаксации за счет делокализации π-электронной системы и повышения конформационной подвижности карбонильного фрагмента [30]. В свою очередь, самое продолжительное время жизни флуоресценции было обнаружено у соединений 4 и 2. Вероятно, это связано с жесткостью их молекулярной структуры, которая ограничивает безызлучательные процессы и способствует преобладанию излучательных переходов.

Были исследованы перспективы создания OLED-устройств на основе изученных органических соединений 1–4 и соединения Cou. Изучены спектры электролюминесценции устройств, изготовленных на основе соединений 1–4 и Cou (рис. 5, *a*), а также получены их вольт-амперные (рис. 5, *b*) и вольт-яркостные характеристики (рис. 5, *c*). В спектре электролюминесценции для соединения 1 наблюдается максимум с λ_1 = 582 нм, для соединения 4 — λ_{41} = 464 нм, λ_{42} = 560 нм, причем, первый из максимумов обусловлен излучением материала красителя, второй — вкладом транспортного слоя TPBi [31]. Для данных красителей наблюдался батохромный сдвиг максимума длины волны электролюминесценции по сравнению с их фотолюминесценцией, который составил 12–20 нм, что может быть обусловлено образованием эксиплексов [32]. Для соединения 2 максимум электролюминесценции расположен на λ_2 = 544 нм, что соответствует спектру фотолюминесценции. Для соединения 3 длина волны максимума электролюминесценции составила λ_3 = 546 нм, по сравнению со спектром фотолюминесценции наблюдается батохромный сдвиг (около 78 нм), который может быть связан с наличием гидроксигруппы [30] (для аминогрупп подобная ситуация также характерна) в составе красителя, при этом при электрическом возбуждении возникает более стабильная таутомерная форма [27], вызывающая сдвиг длины волны максимума излучения. Аналогичная ситуация наблюдается и для соединения Cou: батохромный сдвиг по отношению к фотолюминесценции составил около 100 нм, что обусловлено наличием аминогруппы в составе данного красителя [19].

Вольт-амперные характеристики исследованных OLED-устройств (рис. 5, *b*) описываются зависимостью вида $J \propto V^n$ [33]. Анализ кривых показал, что при напряжениях до 3 В наблюдается режим проводимости, ограниченной пространственным зарядом ($n \approx 2$, Space Charge Limited Conduction, SCLC). Подобное поведение структуры указывает на ситуацию, когда движение носителей заряда в слое ограничивается их накоплением внутри полупроводника. При более высоких напряжениях проводимость становится ограниченной процессами захвата носителей на ловушках ($n > 2$, Trap-Limited Conduction). В этом режиме рост тока с увеличением напряжения происходит быстрее, чем при SCLC, что соответствует большему показателю степени n. Были исследованы вольт-яркостные характеристики (рис. 5, *c*). Оказалось, что у устройств на основе соединений 1, 3 и 4 с увеличением приложенного напряжения яркость возрастает медленнее и достигает меньших значений, чем у соединения Cou. В то же время новое соединение 2 демонстрирует яркость, сопоставимую с соединением Cou, т. е. при прочих равных условиях светоотдача устройства на его основе близка к показателям эталонного материала. Соединение 2, содержащее 4-гидроксихроменовый акцепторный фрагмент, сопряженный с 4-метоксифенилпиразолиновым донорным блоком, формирует эффективную систему внутримолекулярного переноса заряда, вероятно, вследствие чего происходит увеличение вероятности излучательного перехода и улучшение баланса инжекции электронов и дырок, что в совокупности приводит к высокой эффективности электролюминесценции. Напротив, OLED с соединением 3, содержащим ацетильную группу, демонстрирует существенно более низкие показатели яркости из-за нарушения π-сопряжения, что вызывает локализацию π-электронной плотности и увеличение вероятности безызлучательных переходов. В результате, сопряженные системы, такие как соединение Cou и 2, показывают максимальную эффективность электролюминесценции благодаря делокализации возбужденного состояния. С другой стороны, соединение 4 с гибким диэтиламино-фрагментом обладает повышен-

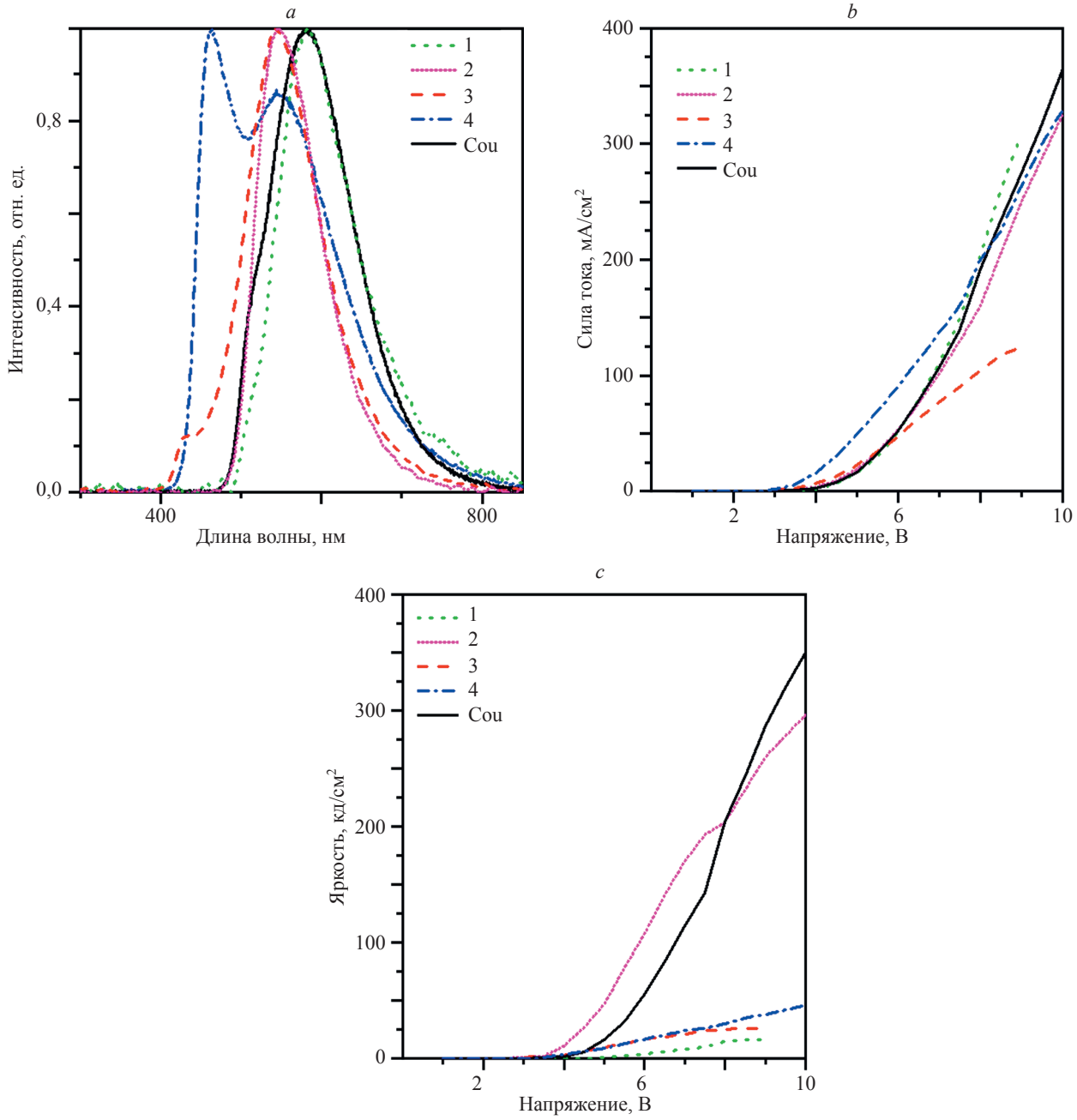

Рис. 5. Оптические и электрофизические характеристики исследуемых соединений: спектры электролюминесценции (*a*); вольт-амперные характеристики (*b*); вольт-яркостные характеристики (*c*)

Fig. 5. Optical and electrophysical characteristics of the studied compounds: electroluminescence spectra (*a*); current-voltage characteristics (*b*); brightness-voltage characteristics (*c*)

ной конформационной подвижностью, что увеличивает вероятность интеркомбинационной и внутренней конверсий, снижая тем самым квантовый выход фотолюминесценции и яркость соответствующего устройства. Стерические факторы также оказывают существенное влияние на оптоэлектронные характеристики. Кроме того, жесткая плоская структура соединения 2 минимизирует колебательные потери энергии, тогда как наличие объемных или гибких заместителей в других соединениях увеличивает конформационную энтропию молекулы. Особенно показателен в этом отношении пример соединения 1, где антраценовый фрагмент, несмотря на свою жесткость, создает дисбаланс в транспорте зарядов, вызывая преобладание дырочной проводимости и, как следствие, сужение рекомбинационной зоны и нелинейную зависимость яркости от напряжения.

Для созданных OLED были получены координаты цветности в рамках модели CIE 1931 [34]. Устройство на основе соединения 1 излучает в оранжево-красной области спектра с высокой чистотой цвета, причем обнаружен небольшой сдвиг в сторону более теплого оттенка, в зависимости от приложенного напряжения. Для электролюминесценции соединения 2 характе-

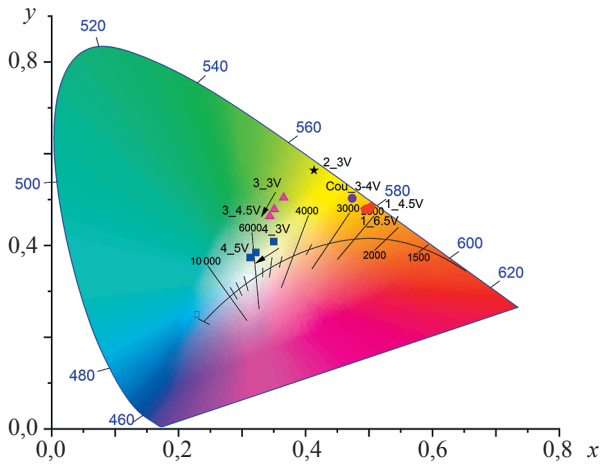

Рис. 6. Хроматическая диаграмма для исследованных соединений в зависимости от приложенных напряжений.

Для соединений: 1_4.5V–1_6.5V напряжения 4,5–6,5 В; 2_3V — 3 В; 3_3V–3_4.5V — 3–4,5 В; 4_3V–4_5V — 3–5 В; Cou_3–4V — 3–4 В

Fig. 6. Chromatic diagram for the studied compounds depending on the applied voltages.

For compounds: 1_4.5V–1_6.5V voltages are 4.5–6.5 V; 2_3V — 3 V; 3_3V–3_4.5V — 3–4.5 V; 4_3V–4_5V — 3–5 V; Cou_3–4V — 3–4 V

рен зелено-желтый оттенок, а излучение устройства на основе соединения 3 смещается к зелено-желтой зоне с меньшей чистотой цвета. Для OLED на основе соединения 4, в результате увеличения приложенного напряжения координаты цветности сдвигаются в область белого света на хроматической диаграмме (рис. 6). Коррелированная цветовая температура для данного устройства составила 6410 К при напряжении 4 В, что указывает на холодный оттенок белого цвета, близкий к естественному дневному свету. Полученные результаты свидетельствуют о том, что соединение 4 является перспективным материалом для создания OLED-устройств с белым излучением. Наконец, OLED на основе соединения Cou демонстрирует теплое, насыщенно-желто-оранжевое излучение.

Заключение

Исследованы фотофизические свойства четырех производных кумарина (соединения 1–4) и проведено их сравнение с соединением Cou, а также созданы многослойные органические светоизлучающие устройства на их основе и изучены их оптические и электрофизические характеристики. Установлено, что донорно-акцепторные взаимодействия и пространственные эффекты в молекулах кумариновых производных определяют их спектральные характеристики, при этом соединение 2 с 4-гидроксихроменовым и 4-метоксифенилпиразолиновым фрагментами продемонстрировало наиболее эффективный внутримолекулярный перенос заряда, что обусловило высокую яркость светодиода на его основе, сопоставимую с соединением Cou. Анализ вольт-амперных характеристик органических светоизлучающих устройств выявил два режима проводимости — ограниченной пространственным зарядом и процессами захвата носителей. Соединение 2 продемонстрировало наилучшие электролюминесцентные характеристики, что делает его перспективным кандидатом для создания высокоэффективных светодиодов. Также заслуживает внимания светодиод на основе соединения 4, показавший белое излучение с координатами цветности, близкими к дневному свету (цветовая температура около 6400 К) за счет синергизма электролюминесценции кумаринового красителя и электронного транспортного слоя, что открывает возможности его применения в осветительных системах.

Литература

1. Song J., Guan Y., Wang C., Li W., Bao X., Niu L. Effect of conductive polymers PEDOT:PSS on exciton recombination and conversion in doped-type BioLEDs // Polymers. 2023. V. 15. N 15. P. 3275. https://doi.org/10.3390/polym15153275
2. Kim J., Jeon M.-G., Yun S., Kirakosyan A., Choi J. Suppressing metal cation diffusion in perovskite light-emitting diodes via blending amino acids with PEDOT:PSS // ACS Photonics. 2025. V. 12. N 2. P. 971–980. https://doi.org/10.1021/acsphotonics.4c02027
3. Liu R., Yu T., Su R., Zhao Y., Zhang D., Zhang S., Su W. Photo-and electroluminescent properties of V-shaped fused-biscoumarins containing tert-butyl group modified imidazole/carbazole groups // Organic Electronics. 2025. V. 139. P. 107208. https://doi.org/10.1016/j.orgel.2025.107208
4. Osadchenko A.V., Ambrozevich S.A., Zakharchuk I.A., Vashchenko A.A., Daibagya D.S., Ryzhov A.V., et al. Organic light-emitting diodes based on Eu(III) complexes involving 1,1,1-trifluoro-4-phenyl-2,4-butanedione with ethanoic and n-butanoic acids // Physics of Wave Phenomena. 2025. V. 33. N 1. P. 72–77. https://doi.org/10.3103/S1541308X24700559
5. Jia Z., Xie X., Guo Z., Kou Z. High-CRI warm white OLEDs based on TADF-doped exciplex co-host structure enabled by efficient reverse intersystem crossing // Organic Electronics. 2025. V. 141. P. 107229. https://doi.org/10.1016/j.orgel.2025.107229

References

1. Song J., Guan Y., Wang C., Li W., Bao X., Niu L. Effect of conductive polymers PEDOT:PSS on exciton recombination and conversion in doped-type BioLEDs. *Polymers*, 2023, vol. 15, no. 15, pp. 3275. https://doi.org/10.3390/polym15153275
2. Kim J., Jeon M.-G., Yun S., Kirakosyan A., Choi J. Suppressing metal cation diffusion in perovskite light-emitting diodes via blending amino acids with PEDOT:PSS. *ACS Photonics*, 2025, vol. 12, no. 2, pp. 971–980. https://doi.org/10.1021/acsphotonics.4c02027
3. Liu R., Yu T., Su R., Zhao Y., Zhang D., Zhang S., Su W. Photo-and electroluminescent properties of V-shaped fused-biscoumarins containing tert-butyl group modified imidazole/carbazole groups. *Organic Electronics*, 2025, vol. 139, pp. 107208. https://doi.org/10.1016/j.orgel.2025.107208
4. Osadchenko A.V., Ambrozevich S.A., Zakharchuk I.A., Vashchenko A.A., Daibagya D.S., Ryzhov A.V., et al. Organic light-emitting diodes based on Eu(III) complexes involving 1,1,1-trifluoro-4-phenyl-2,4-butanedione with ethanoic and n-butanoic acids. *Physics of Wave Phenomena*, 2025, vol. 33, no. 1, pp. 72–77. https://doi.org/10.3103/S1541308X24700559
5. Jia Z., Xie X., Guo Z., Kou Z. High-CRI warm white OLEDs based on TADF-doped exciplex co-host structure enabled by efficient reverse intersystem crossing. *Organic Electronics*, 2025, vol. 141, pp. 107229. https://doi.org/10.1016/j.orgel.2025.107229

6. Осадченко А.В., Амброзевич С.А., Захарчук И.А., Ващенко А.А., Дайбаге Д.С., Рыжов А.В., Певцов Д.Н., Певцов Н.В., Селюков А.С. Электролюминесценция новых координационных соединений ионов европия с β-дикетонами, уксусной и масляной кислотами // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24. № 4. С. 570–576. https://doi.org/10.17586/2226-1494-2024-24-4-570-576
7. Cao Y., Wang N., Tian H., Guo J., Wei Y., Chen H., et al. Perovskite light-emitting diodes based on spontaneously formed submicrometre-scale structures // Nature. 2018. V. 562. N 7726. P. 249–253. https://doi.org/10.1038/s41586-018-0576-2
8. Chiba T., Hayashi Y., Ebe H., Hoshi K., Sato J., Sato S., PuY.-J., Ohisa Kido J. Anion-exchange red perovskite quantum dots with ammonium iodine salts for highly efficient light-emitting devices // Nature Photonics. 2018. V. 12. N 11. P. 681–687. https://doi.org/10.1038/s41566-018-0260-y
9. Lin K., Xing J., Quan L.N., de Arquer F.P.G., Gong X., Lu J., et al. Perovskite light-emitting diodes with external quantum efficiency exceeding 20 per cent // Nature. 2018. V. 562. N 7726. P. 245–248. https://doi.org/10.1038/s41586-018-0575-3
10. Liu Y., Cui J., Du K., Tian H., He Z., Zhou Q., et al. Efficient blue light-emitting diodes based on quantum-confined bromide perovskite nanostructures // Nature Photonics. 2019. V. 13. N 11. P. 760–764. https://doi.org/10.1038/s41566-019-0505-4
11. Li C.H.A., Zhou Z., Vashishtha P., Halpert J.E. The future is blue (LEDs): why chemistry is the key to perovskite displays // Chemistry of Materials. 2019. V. 31. N 16. P. 6003–6032. https://doi.org/10.1021/acs.chemmater.9b01650
12. Wang Q., Wang X., Yang Z., Zhou N., Deng Y., Zhao J., et al. Efficient sky-blue perovskite light-emitting diodes via photoluminescence enhancement // Nature Communications. 2019. V. 10. N 1. P. 5633. https://doi.org/10.1038/s41467-019-13580-w
13. Kang S., Jillella R., Jeong J., Park Y.-I., Pu Y.-J., Park J. Improved electroluminescence performance of perovskite light-emitting diodes by a new hole transporting polymer based on the benzocarbazole moiety // ACS Applied Materials and Interfaces. 2020. V. 12. N 46. P. 51756–51765. https://doi.org/10.1021/acsami.0c16593
14. Zhang X., Guo M., Li J., Dai T., Yang Z., Lou Z., et al. Low-voltage RGB perovskite light-emitting transistors with magnetron sputtered Ta_2O_5 high-k dielectric layer // Organic Electronics. 2025. V. 142. P. 107241. https://doi.org/10.1016/j.orgel.2025.107241
15. Worku M., Ben-Akacha A., Shonde T.B., Liu H., Ma B. The past, present, and future of metal halide perovskite light-emitting diodes // Small Science. 2021. V. 1. N 8. P. 2000072. https://doi.org/10.1002/smsc.202000072
16. Yu T., Zhang P., Zhao Y., Zhang H., Meng J., Fan D., Chen L., Qiu Y. Synthesis, crystal structure and photo- and electro-luminescence of the coumarin derivatives with benzotriazole moiety // Organic Electronics. 2010. V. 11. N 1. P. 41–49. https://doi.org/10.1016/j.orgel.2009.09.023
17. Zhang H., Chai H., Yu T., Zhao Y., Fan D. High-efficiency blue electroluminescence based on coumarin derivative 3-(4-(anthracen-10-yl)phenyl)-benzo[5,6]coumarin // Journal of Fluorescence. 2012. V. 22. N 6. P. 1509–1512. https://doi.org/10.1007/s10895-012-1088-3
18. Shreykar M.R., Sekar N. Stimuli-responsive luminescent coumarin thiazole hybrid dye: synthesis, aggregation induced emission, thermochromism and DFT study // Dyes and Pigments. 2017. V. 142. P. 121–125. https://doi.org/10.1016/j.dyepig.2017.03.028
19. Zhang H., Liu X., Gong Y., Yu T., Zhao Y. Synthesis and characterization of SFX-based coumarin derivatives for OLEDs // Dyes and Pigments. 2021. V. 185. Part A. P. 108969. https://doi.org/10.1016/j.dyepig.2020.108969
20. Zhang H., Luo Q., Mao Y., Zhao Y., Yu T. Synthesis and characterization of coumarin-biphenyl derivatives as organic luminescent materials // Journal of Photochemistry and Photobiology A: Chemistry. 2017. V. 346. P. 10–16. https://doi.org/10.1016/j.jphotochem.2017.05.039
21. Traven V.F., Cheptsov D.A., Svetlova J.I., Ivanov I.V., Cuerva C., Lodeiro C. et al. The role of the intermolecular π···π interactions in the luminescence behavior of novel coumarin-based pyrazoline materials // Dyes and Pigments. 2021. V. 186. P. 108942. https://doi.org/10.1016/j.dyepig.2020.108942
22. Traven V.F., Cheptsov D.A., Bulanova M.V., Solovjova N.P., Chibisova T.A., Dolotov S.M., Ivanov I.V., et al. On the mechanism of photodehydrogenation of aryl(hetaryl)pyrazolines in the presence

6. Osadchenko A.V., Ambrozevich S.A., Zakharchuk I.A., Vashchenko A.A., Daibagya D.S., Ryzhov A.V., Pevtsov D.N., Pevtsov N.V., Selyukov A.S. Electroluminescence of new coordination compounds of europium ions with β-diketones, acetic and butyric acids. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 4, pp. 570–576. (in Russian). https://doi.org/10.17586/2226-1494-2024-24-4-570-576
7. Cao Y., Wang N., Tian H., Guo J., Wei Y., Chen H., et al. Perovskite light-emitting diodes based on spontaneously formed submicrometre-scale structures. *Nature*, 2018, vol. 562, no. 7726, pp. 249–253. https://doi.org/10.1038/s41586-018-0576-2
8. Chiba T., Hayashi Y., Ebe H., Hoshi K., Sato J., Sato S., PuY.-J., Ohisa Kido J. Anion-exchange red perovskite quantum dots with ammonium iodine salts for highly efficient light-emitting devices. *Nature Photonics*, 2018, vol. 12, no. 11, pp. 681–687. https://doi.org/10.1038/s41566-018-0260-y
9. Lin K., Xing J., Quan L.N., de Arquer F.P.G., Gong X., Lu J., et al. Perovskite light-emitting diodes with external quantum efficiency exceeding 20 per cent. *Nature*, 2018, vol. 562, no. 7726, pp. 245–248. https://doi.org/10.1038/s41586-018-0575-3
10. Liu Y., Cui J., Du K., Tian H., He Z., Zhou Q., et al. Efficient blue light-emitting diodes based on quantum-confined bromide perovskite nanostructures. *Nature Photonics*, 2019, vol. 13, no. 11, pp. 760–764. https://doi.org/10.1038/s41566-019-0505-4
11. Li C.H.A., Zhou Z., Vashishtha P., Halpert J.E. The future is blue (LEDs): why chemistry is the key to perovskite displays. *Chemistry of Materials*, 2019, vol. 31, no. 16, pp. 6003–6032. https://doi.org/10.1021/acs.chemmater.9b01650
12. Wang Q., Wang X., Yang Z., Zhou N., Deng Y., Zhao J., et al. Efficient sky-blue perovskite light-emitting diodes via photoluminescence enhancement. *Nature Communications*, 2019, vol. 10, no. 1, pp. 5633. https://doi.org/10.1038/s41467-019-13580-w
13. Kang S., Jillella R., Jeong J., Park Y.-I., Pu Y.-J., Park J. Improved electroluminescence performance of perovskite light-emitting diodes by a new hole transporting polymer based on the benzocarbazole moiety. *ACS Applied Materials and Interfaces*, 2020, vol. 12, no. 46, pp. 51756–51765. https://doi.org/10.1021/acsami.0c16593
14. Zhang X., Guo M., Li J., Dai T., Yang Z., Lou Z., et al. Low-voltage RGB perovskite light-emitting transistors with magnetron sputtered Ta_2O_5 high-k dielectric layer. *Organic Electronics*, 2025, vol. 142, pp. 107241. https://doi.org/10.1016/j.orgel.2025.107241
15. Worku M., Ben-Akacha A., Shonde T.B., Liu H., Ma B. The past, present, and future of metal halide perovskite light-emitting diodes. *Small Science*, 2021, vol. 1, no. 8, pp. 2000072. https://doi.org/10.1002/smsc.202000072
16. Yu T., Zhang P., Zhao Y., Zhang H., Meng J., Fan D., Chen L., Qiu Y. Synthesis, crystal structure and photo- and electro-luminescence of the coumarin derivatives with benzotriazole moiety. *Organic Electronics*, 2010, vol. 11, no. 1, pp. 41–49. https://doi.org/10.1016/j.orgel.2009.09.023
17. Zhang H., Chai H., Yu T., Zhao Y., Fan D. High-efficiency blue electroluminescence based on coumarin derivative 3-(4-(anthracen-10-yl)phenyl)-benzo[5,6]coumarin. *Journal of Fluorescence*, 2012, vol. 22, no. 6, pp. 1509–1512. https://doi.org/10.1007/s10895-012-1088-3
18. Shreykar M.R., Sekar N. Stimuli-responsive luminescent coumarin thiazole hybrid dye: synthesis, aggregation induced emission, thermochromism and DFT study. *Dyes and Pigments*, 2017, vol. 142, pp. 121–125. https://doi.org/10.1016/j.dyepig.2017.03.028
19. Zhang H., Liu X., Gong Y., Yu T., Zhao Y. Synthesis and characterization of SFX-based coumarin derivatives for OLEDs. *Dyes and Pigments*, 2021, vol. 185, part A, pp. 108969. https://doi.org/10.1016/j.dyepig.2020.108969
20. Zhang H., Luo Q., Mao Y., Zhao Y., Yu T. Synthesis and characterization of coumarin-biphenyl derivatives as organic luminescent materials. *Journal of Photochemistry and Photobiology A: Chemistry*, 2017, vol. 346, pp. 10–16. https://doi.org/10.1016/j.jphotochem.2017.05.039
21. Traven V.F., Cheptsov D.A., Svetlova J.I., Ivanov I.V., Cuerva C., Lodeiro C. et al. The role of the intermolecular π···π interactions in the luminescence behavior of novel coumarin-based pyrazoline materials. *Dyes and Pigments*, 2021, vol. 186, pp. 108942. https://doi.org/10.1016/j.dyepig.2020.108942
22. Traven V.F., Cheptsov D.A., Bulanova M.V., Solovjova N.P., Chibisova T.A., Dolotov S.M., Ivanov I.V., et al. On the mechanism of photodehydrogenation of aryl(hetaryl)pyrazolines in the presence

of perchloroalkanes // Photochemistry and Photobiology. 2018. V. 94. N 4. P. 659–666. https://doi.org/10.1111/php.12918
23. Secci D., Carradori S., Bolasco A., Chimenti P., Yáñez M., Ortuso F., Alcaro S., et al. Synthesis and selective human monoamine oxidase inhibition of 3-carbonyl, 3-acyl, and 3-carboxyhydrazido coumarin derivatives // European Journal of Medicinal Chemistry. 2011. V. 46. N 10. P. 4846–4852. https://doi.org/10.1016/j.ejmech.2011.07.017
24. Huang Z.-L., Li N., Sun Y., Wang H., Song H., Xu Z. Synthesis and structure–photophysical property relationships for two coumarinyl-based two-photon induced fluorescent molecules // Journal of Molecular Structure. 2003. V. 657. N 1-3. P. 343–350. https://doi.org/10.1016/S0022-2860(03)00427-7
25. Lee S., Sivakumar K., Shin W-S., Xie F., Wang Q. Synthesis and anti-angiogenesis activity of coumarin derivatives // Bioorganic & Medicinal Chemistry Letters. 2006. V. 16. N 17. P. 4596–4599. https://doi.org/10.1016/j.bmcl.2006.06.007
26. Shreykar M.R., Sekar N. Coumarin-pyrazole hybrid with red shifted ESIPT emission and AIE characteristics — a comprehensive study // Journal of Fluorescence. 2017. V. 27. N 5. P. 1687–1707. https://doi.org/10.1007/s10895-017-2106-2
27. Gawad S.A.A., Sakr M.A.S. Spectroscopic investigation, DFT and TD-DFT calculations of 7-(Diethylamino) Coumarin (C466) // Journal of Molecular Structure. 2022. V. 1248. P. 131413. https://doi.org/10.1016/j.molstruc.2021.131413
28. Yang L., Liu Y., Zhou X., Wu Y., Ma C., Liu W., Zhang C. Asymmetric anthracene-fused BODIPY dye with large Stokes shift: synthesis, photophysical properties and bioimaging // Dyes and Pigments. 2016. V. 126. P. 232–238. https://doi.org/10.1016/j.dyepig.2015.11.028
29. Felorzabihi N., Haley J.C., Bardajee G.R., Winnik M.A. Systematic study of the fluorescence decays of amino-coumarin dyes in polymer matrices // Journal of Polymer Science Part B: Polymer Physics. 2007. V. 45. N 17. P. 2333–2343. https://doi.org/10.1002/polb.21226
30. Altalbawy F.M.A., Abdelkader M.H., Darwish E.S.S., Elnagdi M.H. Synthesis, electronic absorption, fluorescence and live time spectroscopic study of some new 3,7-disubstituted coumarin derivatives as new fluorescent probes // Asian Journal of Chemistry. 2016. V. 28. N 10. P. 2303–2310. https://doi.org/10.14233/ajchem.2016.19975
31. Takizawa S., Montes V.A., Anzenbacher P. Phenylbenzimidazole-based new bipolar host materials for efficient phosphorescent organic light-emitting diodes // Chemistry of Materials. 2009. V. 21. N 12. P. 2452–2458. https://doi.org/10.1021/cm9004954
32. Kotchapadist P., Prachumrak N., Sunonnam T., Namuangruk S., Sudyoadsuk T., Keawin T., et al. Synthesis, characterisation, and electroluminescence properties of N-coumarin derivatives containing peripheral triphenylamine // European Journal of Organic Chemistry. 2015. V. 2015. N 3. P. 496–505. https://doi.org/10.1002/ejoc.201402680
33. Sahoo R.K., Atta S., Singh N.D.P., Jacob C. Influence of functional derivatives of an amino-coumarin/MWCNT composite organic hetero-junction on the photovoltaic characteristics // Materials Science in Semiconductor Processing. 2014. V. 25. P. 279–285. https://doi.org/10.1016/j.mssp.2014.01.001
34. Осадченко А.В., Ващенко А.А., Захарчук И.А., Дайбаге Д.С., Амброзевич С.А., Володин Н.Ю., Чепцов Д.А., Долотов С.М., Травень В.Ф., Авраменко А.И., Семенова С.Л., Селюков А.С. Органические светоизлучающие диоды с новыми красителями на основе кумарина // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22. № 6. С. 1112–1118. https://doi.org/10.17586/2226-1494-2022-22-6-1112-1118

of perchloroalkanes. *Photochemistry and Photobiology*, 2018, vol. 94, no. 4, pp. 659–666. https://doi.org/10.1111/php.12918
23. Secci D., Carradori S., Bolasco A., Chimenti P., Yáñez M., Ortuso F., Alcaro S., et al. Synthesis and selective human monoamine oxidase inhibition of 3-carbonyl, 3-acyl, and 3-carboxyhydrazido coumarin derivatives. *European Journal of Medicinal Chemistry*, 2011, vol. 46, no. 10, pp. 4846–4852. https://doi.org/10.1016/j.ejmech.2011.07.017
24. Huang Z.-L., Li N., Sun Y., Wang H., Song H., Xu Z. Synthesis and structure–photophysical property relationships for two coumarinyl-based two-photon induced fluorescent molecules. *Journal of Molecular Structure*, 2003, vol. 657, no. 1-3, pp. 343–350. https://doi.org/10.1016/S0022-2860(03)00427-7
25. Lee S., Sivakumar K., Shin W-S., Xie F., Wang Q. Synthesis and anti-angiogenesis activity of coumarin derivatives. *Bioorganic & Medicinal Chemistry Letters*, 2006, vol. 16, no. 17, pp. 4596–4599. https://doi.org/10.1016/j.bmcl.2006.06.007
26. Shreykar M.R., Sekar N. Coumarin-pyrazole hybrid with red shifted ESIPT emission and AIE characteristics — a comprehensive study. *Journal of Fluorescence*, 2017, vol. 27, no. 5, pp. 1687–1707. https://doi.org/10.1007/s10895-017-2106-2
27. Gawad S.A.A., Sakr M.A.S. Spectroscopic investigation, DFT and TD-DFT calculations of 7-(Diethylamino) Coumarin (C466). *Journal of Molecular Structure*, 2022, vol. 1248, pp. 131413. https://doi.org/10.1016/j.molstruc.2021.131413
28. Yang L., Liu Y., Zhou X., Wu Y., Ma C., Liu W., Zhang C. Asymmetric anthracene-fused BODIPY dye with large Stokes shift: synthesis, photophysical properties and bioimaging. *Dyes and Pigments*, 2016, vol. 126, pp. 232–238. https://doi.org/10.1016/j.dyepig.2015.11.028
29. Felorzabihi N., Haley J.C., Bardajee G.R., Winnik M.A. Systematic study of the fluorescence decays of amino-coumarin dyes in polymer matrices. *Journal of Polymer Science Part B: Polymer Physics*, 2007, vol. 45, no. 17, pp. 2333–2343. https://doi.org/10.1002/polb.21226
30. Altalbawy F.M.A., Abdelkader M.H., Darwish E.S.S., Elnagdi M.H. Synthesis, electronic absorption, fluorescence and live time spectroscopic study of some new 3,7-disubstituted coumarin derivatives as new fluorescent probes. *Asian Journal of Chemistry*, 2016, vol. 28, no. 10, pp. 2303–2310. https://doi.org/10.14233/ajchem.2016.19975
31. Takizawa S., Montes V.A., Anzenbacher P. Phenylbenzimidazole-based new bipolar host materials for efficient phosphorescent organic light-emitting diodes. *Chemistry of Materials*, 2009, vol. 21, no. 12, pp. 2452–2458. https://doi.org/10.1021/cm9004954
32. Kotchapadist P., Prachumrak N., Sunonnam T., Namuangruk S., Sudyoadsuk T., Keawin T., et al. Synthesis, characterisation, and electroluminescence properties of N-coumarin derivatives containing peripheral triphenylamine. *European Journal of Organic Chemistry*, 2015, vol. 2015, no. 3, pp. 496–505. https://doi.org/10.1002/ejoc.201402680
33. Sahoo R.K., Atta S., Singh N.D.P., Jacob C. Influence of functional derivatives of an amino-coumarin/MWCNT composite organic hetero-junction on the photovoltaic characteristics. *Materials Science in Semiconductor Processing*, 2014, vol. 25, pp. 279–285. https://doi.org/10.1016/j.mssp.2014.01.001
34. Osadchenko A.V., Vashchenko A.A., Zakharchuk I.A., Daibagya D.S., Ambrozevich S.A., Volodin N.Yu., Cheptsov D.A., Dolotov S.M., Traven V.F., Avramenko A.I., Semenova S.L., Selyukov A.S. Organic light-emitting diodes with new dyes based on coumarin. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1112–1118. (in Russian). https://doi.org/10.17586/2226-1494-2022-22-6-1112-1118

Авторы

Осадченко Анна Владимировна — младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация; ассистент, Российский технологический университет МИРЭА, Москва, 119454, Российская Федерация, sc 57439684100, https://orcid.org/0000-0001-9556-4885, anna.vl.osadchenko@gmail.com

Амброзевич Сергей Александрович — кандидат физико-математических наук, старший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, sc 12789274600, https://orcid.org/0000-0002-3906-0735, s.ambrozevich@mail.ru

Authors

Anna V. Osadchenko — Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation; Assistant, Russian Technological University MIREA, Moscow, 119454, Russian Federation, sc 57439684100, https://orcid.org/0000-0001-9556-4885, anna.vl.osadchenko@gmail.com

Sergey A. Ambrozevich — PhD (Physics & Mathematics), Senior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, sc 12789274600, https://orcid.org/0000-0002-3906-0735, s.ambrozevich@mail.ru

Захарчук Иван Александрович — младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, sc 57672815700, https://orcid.org/0000-0002-1502-6460, zakharchukia@yandex.ru

Ващенко Андрей Александрович — кандидат физико-математических наук, старший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, sc 35800121600, https://orcid.org/0000-0003-2084-5900, vashchenkoaa@lebedev.ru

Дайбаге Даниил Саюзович — младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, sc 57673090900, https://orcid.org/0000-0003-1944-1546, daibagya@mail.ru

Кармазин Леонид Игоревич — младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, https://orcid.org/0009-0009-3066-3474, leokarm64@gmail.com

Чепцов Дмитрий Андреевич — кандидат химических наук, доцент, Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация, sc 56711985600, https://orcid.org/0000-0002-9774-7922, dchepcov@yandex.ru

Травень Валерий Федорович — доктор химических наук, профессор, заведующий кафедрой, Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация, sc 57208522812, https://orcid.org/0000-0002-2204-7438, valerii.traven@gmail.com

Слюсаренко Алина Романовна — студент, Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация, https://orcid.org/0009-0003-6868-2503, alilisha.2006@mail.ru

Селюков Александр Сергеевич — кандидат физико-математических наук, младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация; доцент, Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация; ассистент, Российский технологический университет МИРЭА, Москва, 119454, Российская Федерация; специалист II категории, Всероссийский институт научной и технической информации РАН, 125190, Москва, Российская Федерация, sc 55787344500, https://orcid.org/0000-0002-4007-6291, selyukov@lebedev.ru

Ivan A. Zakharchuk — Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, sc 57672815700, https://orcid.org/0000-0002-1502-6460, zakharchukia@yandex.ru

Andrey A. Vashchenko — PhD (Physics & Mathematics), Senior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, sc 35800121600, https://orcid.org/0000-0003-2084-5900, vashchenkoaa@lebedev.ru

Daniil S. Daibagya — Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, sc 57673090900, https://orcid.org/0000-0003-1944-1546, daibagya@mail.ru

Leonid I. Karmazin — Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, https://orcid.org/0009-0009-3066-3474, leokarm64@gmail.com

Dmitry A. Cheptsov — PhD (Chemistry), Associate Professor, Mendeleev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation, sc 56711985600, https://orcid.org/0000-0002-9774-7922, dchepcov@yandex.ru

Valery F. Traven — D.Sc. (Chemistry), Professor, Head of Department, Mendeleev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation, sc 57208522812, https://orcid.org/0000-0002-2204-7438, valerii.traven@gmail.com

Alina R. Slyusarenko — Student, Mendeleev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation, https://orcid.org/0009-0003-6868-2503, alilisha.2006@mail.ru

Alexandr S. Selyukov — PhD (Physics & Mathematics), Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation; Associate Professor, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation; Assistant, Russian Technological University MIREA, Moscow, 119454, Russian Federation; 2nd Category Specialist, All Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences, Moscow, 125190, Russian Federation, sc 55787344500, https://orcid.org/0000-0002-4007-6291, selyukov@lebedev.ru

Статья поступила в редакцию 31.07.2025
Одобрена после рецензирования 28.09.2025
Принята к печати 25.11.2025

Received 31.07.2025
Approved after reviewing 28.09.2025
Accepted 25.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1058-1066
УДК 621.315.592

Влияние скорости эпитаксиального роста сильно напряженных квантовых ям InGaAs на фотолюминесценцию гетероструктур на подложках GaAs

Денис Сергеевич Папылев[1]✉, Андрей Владимирович Бабичев[2], Андрей Геннадьевич Гладышев[3], Леонид Яковлевич Карачинский[4], Владислав Васильевич Андрюшкин[5], Иннокентий Игоревич Новиков[6], Арсений Олегович Ефимов[7], Антон Юрьевич Егоров[8], Алексей Михайлович Надточий[9]

[1,2,4,5,6,7] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
[3,5,8] ООО «Коннектор Оптикс», Санкт-Петербург, 194292, Российская Федерация
[9] Национальный исследовательский университет «Высшая школа экономики» (Санкт-Петербургский филиал), Санкт-Петербург, 190008, Российская Федерация

[1] dspapylev@itmo.ru✉, https://orcid.org/0009-0001-3683-5558
[2] a.babichev@itmo.ru, https://orcid.org/0000-0002-3463-4744
[3] andrey.gladyshev@connector-optics.com, https://orcid.org/0000-0002-9448-2471
[4] lkarachinsky@itmo.ru, https://orcid.org/0000-0002-5634-8183
[5] vvandriushkin@itmo.ru, https://orcid.org/0000-0002-7471-8627
[6] innokenty.novikov@itmo.ru, https://orcid.org/0000-0003-1983-0242
[7] aoefimov@itmo.ru, https://orcid.org/0009-0009-2212-0437
[8] anton.egorov@connector-optics.com, https://orcid.org/0000-0002-0789-4241
[9] al.nadtochy@mail.ioffe.ru, https://orcid.org/0000-0003-0982-907X

Аннотация
Введение. Эпитаксия сильно напряженных квантовых ям InGaAs с мольной долей индия более 35 % является технологически сложной задачей. Эффективность люминесценции таких квантовых ям в значительной степени зависит от их структурного качества. Изготовление гетероструктур с упруго-напряженными эпитаксиальными слоями требует оптимизации параметров эпитаксиального роста, одним из которых является скорость роста эпитаксиального слоя. **Метод.** Методом молекулярно-пучковой эпитаксии на подложках GaAs изготовлены гетероструктуры с квантовыми ямами $In_xGa_{1-x}As$ ($0{,}37 \le x < 0{,}41$) с различной скоростью осаждения слоя от 0,24 Å/с до 3,3 Å/с. Методом рентгеновской дифрактометрии определены фактические толщина и состав квантовой ямы, а также исследовано структурное качество гетероструктур. Проведены измерения спектров фотолюминесценции изготовленных гетероструктур при температурах 20 К и 300 К с различной мощностью оптической накачки. На основании зависимости интенсивности фотолюминесценции от мощности накачки произведены расчеты рекомбинационных токов и оценено время безызлучательной рекомбинации в квантовых ямах исследуемых гетероструктур. **Основные результаты.** Содержание InAs в квантовых ямах изготовленных гетероструктур определено в диапазоне 37,0–40,6 %. Исходя из анализа кривых качания рентгеновской дифракции определено ухудшение структурного качества гетероструктур при низкой (0,24 Å/с) скорости роста InGaAs. Измерения методом спектроскопии фотолюминесценции показали значительно более высокую интенсивность фотолюминесценции квантовых ям с умеренной скоростью роста InGaAs (0,9–2,5 Å/с) по сравнению с остальными образцами. Рассчитанные значения времени безызлучательной рекомбинации в квантовых ямах, выращенных при умеренных скоростях роста, составили порядка 10^{-6} с при температуре 20 К и 10^{-9} с при 300 К. В случае увеличения и уменьшения скорости роста время безызлучательной рекомбинации сокращается. **Обсуждение.** Полученные результаты показали достижение наилучшего структурного качества сильно напряженных слоев InGaAs при поддержании скорости роста InGaAs в диапазоне 0,9–2,5 Å/с. Представленные результаты могут быть применены при оптимизации параметров эпитаксиального роста

© Папылев Д.С., Бабичев А.В., Гладышев А.Г., Карачинский Л.Я., Андрюшкин В.В., Новиков И.И., Ефимов А.О., Егоров А.Ю., Надточий А.М., 2025

сильно напряженных квантовых ям InGaAs при изготовлении монолитного вертикально-излучающего лазера на подложке GaAs спектрального диапазона 1200–1300 нм.

Ключевые слова
напряженные квантовые ямы, фотолюминесценция, структурные дефекты, время безызлучательной рекомбинации, молекулярно-пучковая эпитаксия, скорость роста, InGaAs

Благодарности
Работа авторов из Университета ИТМО по анализу данных рентгеновской дифракции и фотолюминесценции поддержана Министерством науки и высшего образования Российской Федерации (государственное задание № FSER-2025-0025).
А.М. Надточий благодарит Программу фундаментальных исследований НИУ ВШЭ за поддержку работы.

Ссылка для цитирования: Папылев Д.С., Бабичев А.В., Гладышев А.Г., Карачинский Л.Я., Андрюшкин В.В., Новиков И.И., Ефимов А.О., Егоров А.Ю., Надточий А.М. Влияние скорости эпитаксиального роста сильно напряженных квантовых ям InGaAs на фотолюминесценцию гетероструктур на подложках GaAs // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1058–1066. doi: 10.17586/2226-1494-2025-25-6-1058-1066

Influence of the growth rate of a highly strained InGaAs quantum well on the photoluminescence of a heterostructures grown on GaAs substrate

Denis S. Papylev[1]✉, Andrey V. Babichev[2], Andrey G. Gladyshev[3], Leonid Ya. Karachinsky[4], Vladislav V. Andryushkin[5], Innokenty I. Novikov[6], Arseny O. Efimov[7], Anton Yu. Egorov[8], Alexey M. Nadtochiy[9]

[1,2,4,5,6,7] ITMO University, Saint Petersburg, 197101, Russian Federation
[3,5,8] Connector Optics LLC, Saint Petersburg, 194292, Russian Federation
[9] National Research University Higher School of Economics, Saint Petersburg, 190008, Russian Federation

[1] dspapylev@itmo.ru✉, https://orcid.org/0009-0001-3683-5558
[2] a.babichev@itmo.ru, https://orcid.org/0000-0002-3463-4744
[3] andrey.gladyshev@connector-optics.com, https://orcid.org/0000-0002-9448-2471
[4] lkarachinsky@itmo.ru, https://orcid.org/0000-0002-5634-8183
[5] vvandriushkin@itmo.ru, https://orcid.org/0000-0002-7471-8627
[6] innokenty.novikov@itmo.ru, https://orcid.org/0000-0003-1983-0242
[7] aoefimov@itmo.ru, https://orcid.org/0009-0009-2212-0437
[8] anton.egorov@connector-optics.com, https://orcid.org/0000-0002-0789-4241
[9] al.nadtochy@mail.ioffe.ru, https://orcid.org/0000-0003-0982-907X

Abstract
Epitaxy of highly strained InGaAs quantum wells with a mole fraction of indium exceeding 35 % is a technologically challenging task. The structural quality of these elastically strained epitaxial layers greatly affects the photoluminescence efficiency of quantum wells. Therefore, in order to achieve high structural quality, optimization of the epitaxial growth parameters is required, one of which is the growth rate of the epitaxial layer. Heterostructures containing the $In_xGa_{1-x}As$ ($0.37 \le x < 0.41$) quantum well were produced on GaAs substrates by molecular beam epitaxy with different growth rates of InGaAs ranging from 0.24 to 3.3 Å/s. The actual thickness and composition of the quantum well were determined by X-ray diffractometry, as well as the structural quality of the heterostructures was investigated. The photoluminescence spectra of the manufactured heterostructures were measured at temperatures of 20 K and 300 K at different optical pumping powers. Based on the dependence of photoluminescence intensity on pumping power, recombination currents were calculated and the time of non-radiative recombination in the studied structures was estimated. The InAs content in the quantum wells of the manufactured heterostructures ranged from 37.0 % to 40.6 %. Based on analysis of X-ray rocking curves, deterioration of structural quality at low deposition rate of 0.24 Å/s was observed. The photoluminescence spectroscopy measurements showed a significantly higher photoluminescence intensity of quantum wells at moderate growth rates of InGaAs (0.9–2.5 Å/s) compared to other samples. The calculated values for the non-radiative recombination lifetime of quantum wells produced at these moderate growth rates were in the order of 10^{-6} s at 20 K and 10^{-9} s at 300 K. At higher or lower growth rates, the values of the non-radiative recombination lifetime decreased. The results obtained demonstrate the achievement of the best structural quality for highly strained InGaAs layers produced at 0.9–2.5 Å/s growth rate. These results can be used to optimize the parameters of epitaxial growth processes for highly strained quantum wells based on InGaAs, for fabricating monolithic vertical-cavity surface-emitting lasers, based on GaAs substrate, operating in the 1200–1300 nm spectral range.

Keywords
strained quantum wells, photoluminescence, structural defects, non-recombination lifetime, molecular-beam epitaxy, growth rate, InGaAs

Acknowledgements
The work of the authors from ITMO University on the X-ray diffraction and photoluminescence data analysis was supported by the Ministry of Science and Higher Education of the Russian Federation (State Assignment No. FSER-2025-0025).
Nadtochiy A. M. thanks the HSE Fundamental Research Program for supporting the work.

For citation: Papylev D.S., Babichev A.V., Gladyshev A.G., Karachinsky L.Ya., Andryushkin V.V., Novikov I.I., Efimov A.O., Egorov A.Yu., Nadtochiy A.M. Influence of the growth rate of a highly strained InGaAs quantum well on the photoluminescence of a heterostructures grown on GaAs substrate. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1058–1066 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1058-1066

Введение

При формировании активной области полупроводниковых лазеров на основе решеточно-согласованных с подложкой материалов спектральный диапазон излучения ограничен шириной запрещенной зоны таких материалов. Для расширения спектрального диапазона в качестве активной области могут использоваться упруго-напряженные квантовые ямы (КЯ) InGaAs на подложке арсенида галлия (GaAs). Так, при использовании сильно напряженных КЯ InGaAs с мольной долей индия (In), достигающей 40 %, может быть достигнуто излучение с длиной волны свыше 1200 нм [1, 2]. Применение такого подхода перспективно для изготовления на подложке GaAs монолитных вертикально-излучающих лазеров спектрального диапазона 1200–1300 нм [3–5], востребованных в современной сенсорике и телекоммуникациях.

Основным препятствием для изготовления гетероструктур вертикально-излучающих лазеров спектрального диапазона 1200–1300 нм на основе КЯ InGaAs является сложность эпитаксии сильно напряженных слоев высокого структурного качества. Характерное при псевдоморфном росте состава $In_{0,4}Ga_{0,6}As$ на подложке GaAs рассогласование постоянных решеток подложки a_{sub} и эпитаксиального слоя a_{film}, равное $(a_{film} - a_{sub})/a_{sub} \approx 2{,}9$ %, приводит к возникновению значительных упругих напряжений сжатия в латеральной плоскости [6].

В зависимости от ростовых параметров релаксация напряжений может происходить через переход от планарного к островковому режиму роста [7], а также через образование дислокаций несоответствия [8]. В результате образование структурных дефектов приводит к формированию центров безызлучательной рекомбинации, способствующих снижению эффективности люминесценции в КЯ [9]. Для предотвращения релаксации напряжений в сжатом слое КЯ InGaAs могут использоваться барьеры арсенид-фосфида галлия (GaAsP) [10], позволяющие компенсировать напряжения [11]. Однако резкость гетерограниц КЯ при таком подходе может ухудшаться за счет формирования промежуточного слоя арсенид-фосфида галлия-индия (InGaAsP), который образуется в результате диффузии атомов In в барьер GaAsP [12], что может привести к уширению спектра люминесценции.

В связи со сложностями эпитаксии сильно напряженных слоев InGaAs возникает необходимость оптимизации параметров эпитаксиального роста, влияющих на структурное качество псевдоморфных КЯ: температуры подложки [13], соотношения потоков материалов V/III групп и скорости роста эпитаксиального слоя [14]. Кроме этого, для достижения высокого структурного качества слоев используется прерывание роста с изменением температуры подложки перед эпитаксией InGaAs ввиду подверженности атомов In к переиспарению при оптимальной для роста GaAs температуре [15].

В настоящей работе методами рентгеновской дифрактометрии и спектроскопии люминесценции исследуется влияние скорости роста методом молекулярно-пучковой эпитаксии сильно напряженных слоев InGaAs с мольной долей In более 35 % на структурное качество КЯ. Основываясь на зависимости интенсивности фотолюминесценции (ФЛ) от мощности оптической накачки приведена расчетная оценка времени безызлучательной рекомбинации для КЯ, сформированных при различных скоростях эпитаксиального роста слоев.

Материалы и методы

Для проведения исследования методом молекулярно-пучковой эпитаксии на подложках GaAs (001) с буферным слоем GaAs 300 нм были выращены пять гетероструктур с КЯ. Гетероструктуры содержали одиночные КЯ $In_{0,4}Ga_{0,6}As$ толщиной 7 нм, которые расположены между обкладочными слоями GaAs толщиной 100 нм, прилегающими к КЯ, и слоями $Al_{0,4}Ga_{0,6}As$ толщиной 50 нм. Состав слоев тройных растворов определялся по предварительно откалиброванным линейным зависимостям скорости роста бинарных составов от температуры эффузионных источников материалов. Образцы различались скоростью роста слоя InGaAs. Для образцов 1, 2 и 3 скорость роста составляла 0,24 Å/с, 0,9 Å/с и 2,5 Å/с соответственно, а для образцов 4 и 5 — 3,3 Å/с. Было выявлено, что содержание In в образце 4 значительно отличается от остальных, поэтому был выращен образец 5 с уточненными температурами эффузионных источников и такой же скоростью роста слоя. Эпитаксия слоев GaAs и арсенида алюминия-галлия (AlGaAs) производилась при температуре подложки 853 К, при эпитаксии слоя InGaAs и прикрывающего его слоя GaAs толщиной 2 нм температура подложки была понижена до 673 К. Соотношение потоков материалов III и V групп поддерживалось равным 1/10.

Фактический состав и толщина КЯ определялись методом рентгеновской дифрактометрии на дифрактометре высокого разрешения PANAlytical X'Pert PRO MRD с гибридным четырех-кристальным щелевым Ge (400) монохроматором при $\omega/2\Theta$ симметричном сканировании отражения (004). Измеренные кривые качания моделировались с использованием программного обеспечения X'Pert Epitaxy. Образцы исследовались методом спектроскопии ФЛ при температурах 20 К и 300 К. Оптическая накачка образцов осуществлялась излучением Nd:YAG лазера 532 нм, сфокусированным в пятно размером 65 мкм. Максимальная мощность оптической накачки составляла 150 мВт и варьировалась посредством использования нейтральных светофильтров. Излучение ФЛ анализировалось с применением монохроматора МДР-23 и детектировалось одноканаль-

ным охлаждаемым Ge-фотодиодом в режиме синхронного детектирования.

Результаты и обсуждение

Измеренные рентгенодифракционные кривые качания (рис. 1) содержат пик состава InGaAs, расположенный левее высокоинтенсивного узкого пика подложки GaAs. Отметим, что на кривых качания присутствуют высокочастотные толщинные осцилляции, образованные интерференцией излучения, отраженного от интерфейсов верхних слоев GaAs/AlGaAs. Такой же эффект создает низкочастотные толщинные осцилляции от тонкого слоя InGaAs, формирующие пик между пиками подложки и состава InGaAs.

Для образцов 2–5, экспериментальные кривые качания хорошо описываются моделируемыми кривыми. В характерной для образца 1 кривой качания слабо выражен пик состава InGaAs, а также отсутствуют толщинные осцилляции как от КЯ, так и от слоев GaAs/AlGaAs, что указывает на ухудшение структурного качества гетероинтерфейсов в структуре, являющегося причиной диффузного рассеяния рентгеновского излучения [16]. Ухудшение структурного качества слоев образца 1 при низкой скорости роста может быть связано с переходом к островковому режиму роста, обусловленному увеличением длины поверхностной диффузии атомов при повышении времени экспозиции эпитаксиальной поверхности [7].

Результаты определения толщины и состава КЯ методом рентгеновской дифрактометрии представлены в таблице. Для образца 1 достоверное определение толщины КЯ не представляется возможным. В результате моделирования кривых качания определено, что оценочная толщина КЯ в гетероструктурах достигает 7,2–7,9 нм. Содержание In при этом составляет от 37,0 % для образца 2 и до 40,6 % для образца 4.

Представленные на рис. 2 результаты измерений спектров ФЛ образца 3 при различных мощностях оптической накачки демонстрируют типовые для всех исследуемых образцов спектры ФЛ, за исключением отличительных особенностей, описанных далее. Аналогично спектрам образца 3 при температуре 20 К образцы 1, 2, 4 и 5 демонстрируют одиночный пик ФЛ, не изменяющий своего положения в диапазоне низких мощностей накачки. Для образцов 2 и 3 при повышении мощности накачки до 4,5 мВт (рис. 2, *a*) происходит коротковолновой сдвиг пика ФЛ, указывающий на насыщение основного и возбуждение более высокоэнергетических излучательных переходов. Для образца 1 коротковолновый сдвиг наблюдается при мощности накачки 150 мВт. Для образцов 4 и 5 при используемых мощностях накачки сдвига длины волны ФЛ не наблюдается. Длина волны пика ФЛ при малых мощностях накачки для образцов 1, 2, 3 и 5 при температуре 20 К определена в диапазоне длин волн 1105–1110 нм. Для образца 4 длина волны пика ФЛ составила 1141 нм, что связано с повышенным содержанием In в КЯ. При температуре 300 К измеренные спектры образцов 2, 4 и 5 аналогично образцу 3 (рис. 2, *b*) также демонстрируют одиночный, но значительно уширенный пик излучения КЯ. При этом не наблюдается коротковолнового сдвига

Таблица. Скорости роста и измеренные параметры КЯ

Table. Growth rate and measured parameters of quantum wells

Номер образца	Скорость роста, Å/с	Содержание InAs, %	Толщина КЯ, нм
1	0,24	38,6	—
2	0,90	37,0	7,9
3	2,50	38,2	7,6
4	3,30	40,6	7,8
5	3,30	38,5	7,2

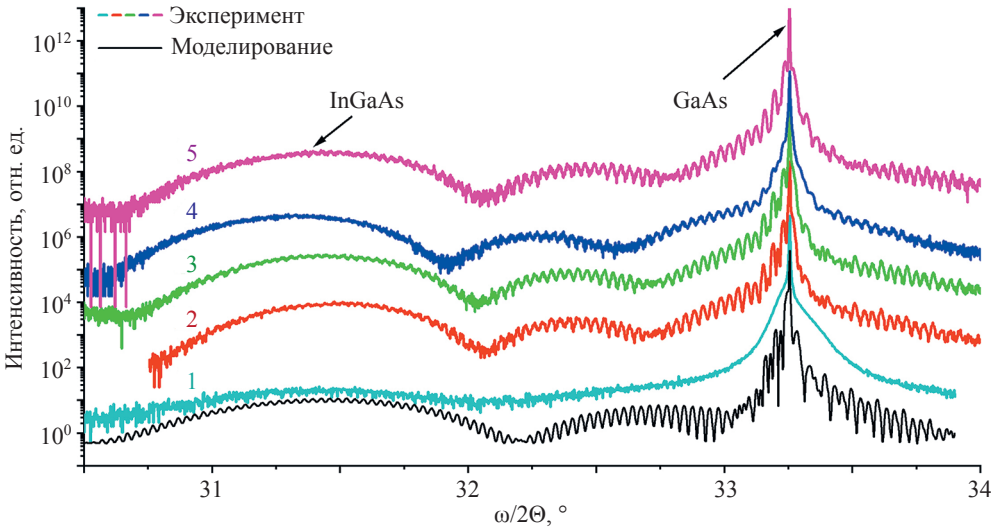

Рис. 1. Экспериментально измеренные и смоделированная для образца 1 кривые качания рентгеновской дифракции. Кривые качания сдвинуты по вертикальной оси для наглядности

Fig. 1. The X-ray diffraction rocking curves measured and modeled for sample 1. The swing curves are shifted along the vertical axis for clarity

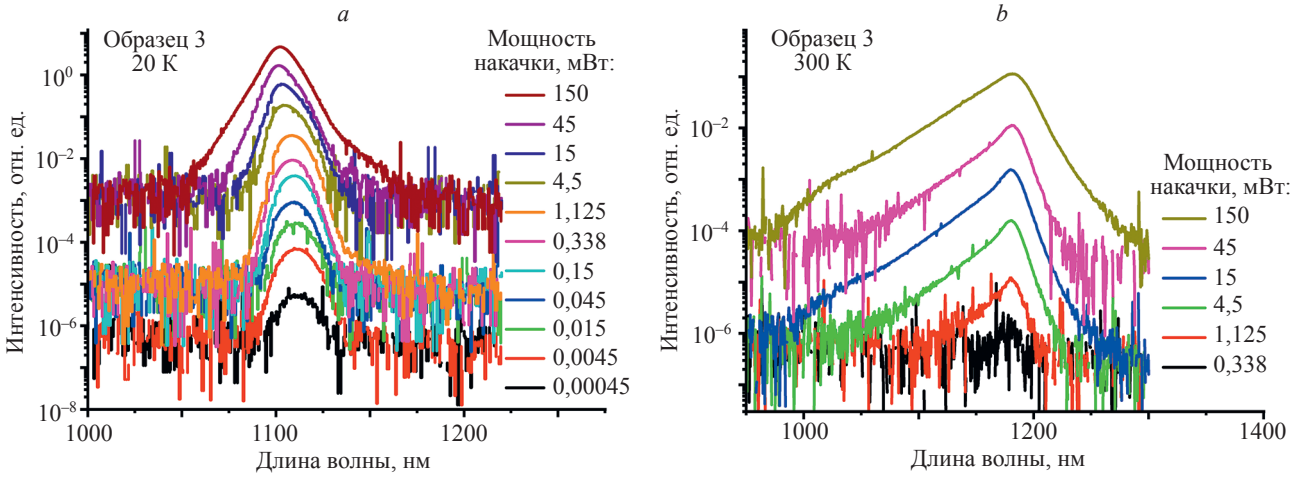

Рис. 2. Спектры фотолюминесценции квантовых ям при температурах 20 К (*a*) и 300 К (*b*) при различной мощности накачки

Fig. 2. Photoluminescence spectra measured at 20 K (*a*) and 300 K (*b*) at different pumping power

пика излучения во всем диапазоне мощности накачки. Однако в спектре ФЛ от образца 1 не было обнаружено различимого пика от КЯ.

Зависимости интегральной интенсивности ФЛ от мощности оптической накачки представлены на рис. 3. При одинаковых мощностях накачки образцы 2 и 3 демонстрируют значительно более высокую интегральную интенсивность ФЛ по сравнению с образцами 1, 4 и 5. Наблюдаемые на зависимостях загибы при увеличении мощности накачки у образцов 1, 2 и 3 коррелируют с ранее отмеченными на спектрах более высокоэнергетическими переходами. Наблюдаемый при высокой мощности накачки резкий загиб зависимости у образца 1, а также полная деградация пика ФЛ от КЯ при комнатной температуре может указывать на разные по своей природе структурные дефекты, являющиеся причиной безызлучательной рекомбинации носителей заряда при чрезмерно низких и высоких скоростях роста напряженных слоев.

Расчет времени безызлучательной рекомбинации

Для оценки времени безызлучательной рекомбинации использовалась модель расчета рекомбинационных токов в гетероструктуре в зависимости от мощности накачки. По причине отсутствия легирования гетероструктур основными механизмами рекомбинации в структурах являются рекомбинация Шокли–Рида–Холла и излучательная рекомбинация, влияние Оже-рекомбинации не должно быть существенным при относительно низких концентрациях носителей заряда. В используемой модели принято условие электронейтральности в системе КЯ и барьерах GaAs:

$$p_{QW} - n_{QW} + p_B - n_B = 0,$$

где p_{QW}, n_{QW} — концентрации в КЯ дырок и электронов соответственно; p_B, n_B — соответствующие концентрации дырок и электронов в барьерах.

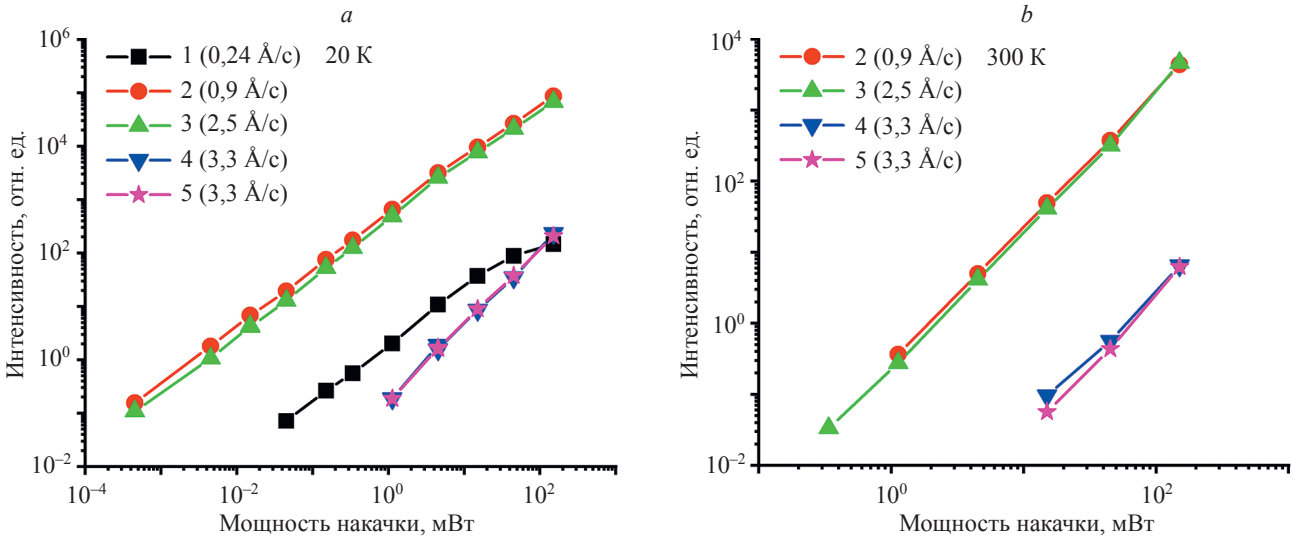

Рис. 3. Зависимости интегральной интенсивности фотолюминесценции от мощности накачки при температурах 20 К (*a*) и 300 К (*b*)

Fig. 3. Dependences of the integral photoluminescence intensity on the pumping power at temperatures 20 K (*a*) and 300 K (*b*)

Полный рекомбинационный ток J_{exc}, соответствующий мощности оптической накачки, определяется суммой безызлучательных J_{Bnr}, J_{QWnr} и излучательных J_{Brad}, J_{QWrad} рекомбинационных токов в барьерах и КЯ соответственно:

$$J_{exc} = J_{Brad} + J_{QWrad} + J_{Bnr} + J_{QWnr};$$
$$J_{Brad} = qBn_Bp_B;$$
$$J_{QWrad} = qBn_{QW}p_{QW};$$
$$J_{Bnr} = qdn_B/\tau_{Bnr},$$
$$J_{QWnr} = qn_{QW}/\tau_{QWnr},$$

где q — элементарный заряд; B — коэффициент излучательной рекомбинации, определенный в работе [17]; τ_{Bnr}, τ_{QWnr} — время безызлучательной рекомбинации в барьерах и КЯ соответственно. В КЯ рекомбинационные токи определяются двумерной концентрацией носителей заряда p_{QW} и n_{QW}, в случае барьера значения тока, полученные с использованием объемных концентраций носителей заряда p_B и n_B домножаются на толщину барьерных слоев d.

Концентрации электронов (дырок) в КЯ и в барьерах вычисляются следующим образом:

$$n(p)_{QW} = \frac{m_{QWn(p)}k_BT}{\pi\hbar^2}\ln\left(\exp\left(\frac{\mu_{n(p)}}{k_BT}\right)+1\right);$$
$$n(p)_B = 2\left(\frac{m_{Bn(p)}k_BT}{2\pi\hbar^2}\right)^{2/3}\exp\left(\frac{\mu_{n(p)}-\Delta E_{Bn(p)}}{k_BT}\right);$$

где m — эффективная масса носителей заряда; k_B — постоянная Больцмана; $\hbar$ — приведенная постоянная Планка; μ — энергия квази-уровня Ферми в зоне проводимости и валентной зоне, отчитываемые от дна (потолка) КЯ; ΔE_B — высота барьера. Расчеты производились при постоянном значении времени безызлучательной рекомбинации в барьере GaAs, равной 1 нс при температурах 20 К и 300 К [18]. На рис. 4 представлены результаты расчетов, демонстрирующие расчетные зависимости роста относительной интенсивности ФЛ с ростом мощности накачки при различном времени безызлучательной рекомбинации.

По результатам расчетов зависимости интенсивности ФЛ от мощности накачки для исследуемых в работе образцов 1–5 определено τ_{QWnr}, при котором достигается наилучшее совпадение расчетных и экспериментальных зависимостей. Полученные результаты оценки времени безызлучательной рекомбинации представлены на рис. 5 в виде зависимости от используемой при изготовлении гетероструктуры скорости роста КЯ. Исходя из полученных результатов, наибольшее влияние скорости роста на время безызлучательной рекомбинации наблюдается при температуре 20 К. При умеренных скоростях роста (0,9–2,5 Å/с) время безызлучательной рекомбинации составляет порядка 10^{-6} с, при минимальной скорости роста (0,24 Å/с) – $3\cdot10^{-7}$ с. Увеличение скорости роста до 3,3 Å/с приводит к более значительному снижению времени безызлучательной рекомбинации до $(4–5)\cdot10^{-9}$ с. При комнатной температуре тепловое уширение значительно влияет на процессы рекомбинации — наблюдается меньшее изменение времени безызлучательной рекомбинации при различных скоростях роста. Для образцов 2–5 при увеличении скорости роста КЯ наблюдается снижение времени безызлучательной рекомбинации с $3\cdot10^{-9}$ с до $8\cdot10^{-10}$ с.

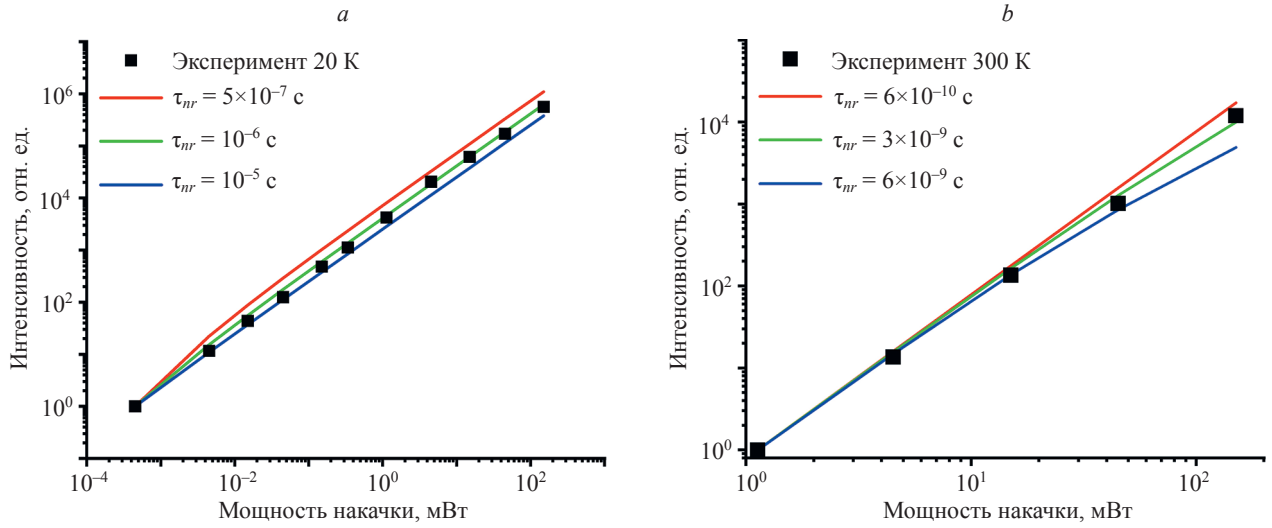

Рис. 4. Расчетные зависимости интенсивности фотолюминесценции от мощности накачки при различном времени безызлучательной рекомбинации при температурах 20 К (*a*) и 300 К (*b*)

Fig. 4. Calculated dependences of the photolumenescence intensity on the pumping power at different non-radiative lifetime values at temperatures 20 K (*a*) and 300 K (*b*)

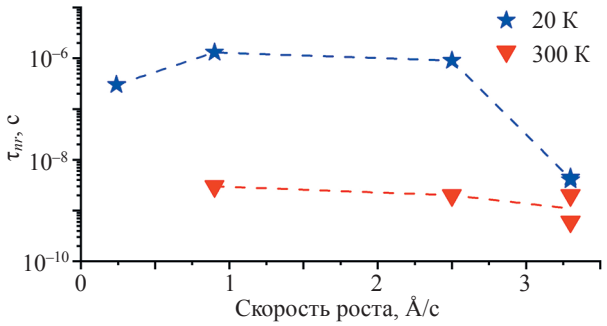

Рис. 5. Расчетные значения времени безызлучательной рекомбинации в зависимости от скорости роста InGaAs в образцах 1–5 при 20 К и образцах 2–5 при 300 К

Fig. 5. Calculated values of non-radiative recombination lifetimes vs. the InGaAs growth rate in samples 1–5 at 20 K and samples 2–5 at 300 K

Заключение

В результате определено, что наиболее подходящая скорость роста методом молекулярно-пучковой эпитаксии сильно напряженных слоев $In_{0,4}Ga_{0,6}As$ на подложке GaAs находится в диапазоне 0,9–2,5 Å/c. Как увеличение, так и уменьшение скорости роста относительно этого диапазона приводит к значительному уменьшению времени безызлучательной рекомбинации носителей заряда, в результате чего интенсивность фотолюминесценции квантовых ям снижается. Более подробное изучение влияния скорости роста слоя квантовых ям возможно при проведении дополнительных измерений и расчетов зависимостей интенсивности фотолюминесценции от мощности накачки, а также при исследовании образцов гетероструктур с квантовыми ямами, выращенных при промежуточных скоростях роста диапазона 0,9–2,5 Å/c. Полученные результаты позволяют изготавливать сильно напряженные квантовые ямы InGaAs, являющие перспективными для производства монолитных вертикально-излучающих лазеров спектрального диапазона 1200–1300 нм.

Литература

1. Sundgren P., Berggren J., Goldman P., Hammar M. Highly strained InGaAs⁄GaAs multiple quantum-wells for laser applications in the 1200–1300 nm wavelength regime // Applied Physics Letters. 2005. V. 87. N 7. P. 071104. https://doi.org/10.1063/1.2010615
2. Sundgren P., von Würtemberg R.M., Berggren J., Hammar M., Ghisoni M., Oscarsson V., Ödling E., Malmquist J. High-performance 1.3 μm InGaAs vertical cavity surface emitting lasers // Electronics Letters. 2003. V. 39. N 15. P. 1128–1129. https://doi.org/10.1049/el:20030733
3. Kondo T., Arai M., Onomura A., Miyamoto T., Koyama F. 1.23 μm long wavelength highly strained GaInAs/GaAs quantum well laser // Proc. of the 15th Annual Meeting of the IEEE Lasers and Electro-Optics Society. 2002. P. 618–619. https://doi.org/10.1109/LEOS.2002.1159459
4. Westbergh P., Söderberg E., Gustavsson J.S., Modh P., Larsson A., Zhang Z., Berggren J., Hammar M. Single mode 1.3 um InGaAs VCSELs for access network applications // Proceedings of SPIE. 2008. V. 6997. P. 69970Y. https://doi.org/10.1117/12.781294
5. Takeda K., Miyamoto T., Kondo T., Uchiyama Y., Kitabayashi N., Uchida T., Matsutani A., Koyama F. Wavelength extension effect on lasing characteristics of highly-strained GaInAs/GaAs vertical-cavity surface-emitting lasers with cavity detuning // Japanese Journal of Applied Physics. 2006. V. 45. N 8S. P. 6691. https://doi.org/10.1143/JJAP.45.6691
6. Deki R., Sasaki T., Takahasi M. Strain relaxation and compositional separation during growth of InGaAs/GaAs (001) // Journal of Crystal Growth. 2017. V. 468. P. 241–244. https://doi.org/10.1016/j.jcrysgro.2017.01.028
7. Grandjean N., Massies J. Epitaxial growth of highly strained $In_xGa_{1-x}As$ on GaAs (001): the role of surface diffusion length // Journal of Crystal Growth. 1993. V. 134. N 1-2. P. 51–62. https://doi.org/10.1016/0022-0248(93)90008-k
8. Chang K.H., Bhattacharya P.K., Gibala R. Characteristics of dislocations at strained heteroepitaxial InGaAs/GaAs interfaces // Journal of Applied Physics. 1989. V. 66. N 7. P. 2993–2998. https://doi.org/10.1063/1.344183
9. Yang V.K., Ting S.M., Groenert M.E., Bulsara M.T., Currie M.T., Leitz C.W., Fitzgerald E.A. Comparison of luminescent efficiency of InGaAs quantum well structures grown on Si, GaAs, Ge, and SiGe virtual substrate // Journal of Applied Physics. 2003. V. 93. N 9. P. 5095–5102. https://doi.org/10.1063/1.1563031
10. Sun W., Kim H., Mawst L.J., Tansu N. Interplay of GaAsP barrier and strain compensation in InGaAs quantum well at near-critical thickness // Journal of Crystal Growth. 2020. V. 531. P. 125381. https://doi.org/10.1016/j.jcrysgro.2019.125381

References

1. Sundgren P., Berggren J., Goldman P., Hammar M. Highly strained InGaAs⁄GaAs multiple quantum-wells for laser applications in the 1200–1300 nm wavelength regime. *Applied Physics Letters*, 2005, vol. 87, no. 7, pp. 071104. https://doi.org/10.1063/1.2010615
2. Sundgren P., von Würtemberg R.M., Berggren J., Hammar M., Ghisoni M., Oscarsson V., Ödling E., Malmquist J. High-performance 1.3 μm InGaAs vertical cavity surface emitting lasers. *Electronics Letters*, 2003, vol. 39, no. 15, pp. 1128–1129. https://doi.org/10.1049/el:20030733
3. Kondo T., Arai M., Onomura A., Miyamoto T., Koyama F. 1.23 μm long wavelength highly strained GaInAs/GaAs quantum well laser. *Proc. of the 15th Annual Meeting of the IEEE Lasers and Electro-Optics Society*, 2002, pp. 618–619. https://doi.org/10.1109/LEOS.2002.1159459
4. Westbergh P., Söderberg E., Gustavsson J.S., Modh P., Larsson A., Zhang Z., Berggren J., Hammar M. Single mode 1.3 um InGaAs VCSELs for access network applications. *Proceedings of SPIE*, 2008, vol. 6997, pp. 69970Y. https://doi.org/10.1117/12.781294
5. Takeda K., Miyamoto T., Kondo T., Uchiyama Y., Kitabayashi N., Uchida T., Matsutani A., Koyama F. Wavelength extension effect on lasing characteristics of highly-strained GaInAs/GaAs vertical-cavity surface-emitting lasers with cavity detuning. *Japanese Journal of Applied Physics*, 2006, vol. 45, no. 8S, pp. 6691. https://doi.org/10.1143/JJAP.45.6691
6. Deki R., Sasaki T., Takahasi M. Strain relaxation and compositional separation during growth of InGaAs/GaAs (001). *Journal of Crystal Growth*, 2017, vol. 468, pp. 241–244. https://doi.org/10.1016/j.jcrysgro.2017.01.028
7. Grandjean N., Massies J. Epitaxial growth of highly strained $In_xGa_{1-x}As$ on GaAs (001): the role of surface diffusion length. *Journal of Crystal Growth*, 1993, vol. 134, no. 1-2, pp. 51–62. https://doi.org/10.1016/0022-0248(93)90008-k
8. Chang K.H., Bhattacharya P.K., Gibala R. Characteristics of dislocations at strained heteroepitaxial InGaAs/GaAs interfaces. *Journal of Applied Physics*, 1989, vol. 66, no. 7, pp. 2993–2998. https://doi.org/10.1063/1.344183
9. Yang V.K., Ting S.M., Groenert M.E., Bulsara M.T., Currie M.T., Leitz C.W., Fitzgerald E.A. Comparison of luminescent efficiency of InGaAs quantum well structures grown on Si, GaAs, Ge, and SiGe virtual substrate. *Journal of Applied Physics*, 2003, vol. 93, no. 9, pp. 5095–5102. https://doi.org/10.1063/1.1563031
10. Sun W., Kim H., Mawst L.J., Tansu N. Interplay of GaAsP barrier and strain compensation in InGaAs quantum well at near-critical thickness. *Journal of Crystal Growth*, 2020, vol. 531, pp. 125381. https://doi.org/10.1016/j.jcrysgro.2019.125381

11. Sweeney S.J., Eales T.D., Adams A.R. The impact of strained layers on current and emerging semiconductor laser systems // Journal of Applied Physics. 2019. V. 125. N 8. P. 082538. https://doi.org/10.1063/1.5063710
12. Dong H., Sun J., Ma S., Liang J., Jia Z., Liu X., Xu B. Interfacial relaxation analysis of InGaAs/GaAsP strain-compensated multiple quantum wells and its optical property // Superlattices and Microstructures. 2018. V. 114. P. 331–339. https://doi.org/10.1016/j.spmi.2017.12.049
13. Jasik A., Wnuk A., Wójcik-Jedlińska A., Jakieła R., Muszalski J., Strupiński W., Bugajski M. The influence of the growth temperature and interruption time on the crystal quality of InGaAs/GaAs QW structures grown by MBE and MOCVD methods // Journal of Crystal Growth. 2008. V. 310. N 11. P. 2785–2792. https://doi.org/10.1016/j.jcrysgro.2008.02.018
14. Jasik A., Wnuk A., Gaca J., Wójcik M., Wójcik-Jedlińska A., Muszalski J., Strupiński W. The influence of the growth rate and V/III ratio on the crystal quality of InGaAs/GaAs QW structures grown by MBE and MOCVD methods // Journal of Crystal Growth. 2009. V. 311. N 19. P. 4423–4432. https://doi.org/10.1016/j.jcrysgro.2009.07.032
15. Wang Q., Wang H., Zhang B., Wang X., Liu W., Wang J., Wang J., Fan J., Zou Y., Ma X. Integrated fabrication of a high strain InGaAs/GaAs quantum well structure under variable temperature and improvement of properties using MOCVD technology // Optical Materials Express. 2021. V. 11. N 8. P. 2378–2388. https://doi.org/10.1364/OME.431015
16. Chuev M.A., Lomov A.A., Imamov R.M. Simultaneous analysis of double-crystal X-ray rocking curves from a set of crystallographic planes // Crystallography Reports. 2006. V. 51. N 2. P. 178–191. https://doi.org/10.1134/S1063774506020027
17. Asryan L.V. Spontaneous radiative recombination and nonradiative Auger recombination in quantum-confined heterostructures // Quantum Electronics. 2005. V. 35. N 12. P. 1117–1120. https://doi.org/10.1070/QE2005v035n12ABEH013093
18. Gurioli M., Vinattieri A., Colocci M., Deparis C., Massies J., Neu G., Bosacchi A., Franchi S. Temperature dependence of the radiative and nonradiative recombination time in GaAs/$Al_xGa_{1-x}As$ quantum-well structures // Physical Review B. 1991. V. 44. N 7. P. 3115–3124. https://doi.org/10.1103/PhysRevB.44.3115

11. Sweeney S.J., Eales T.D., Adams A.R. The impact of strained layers on current and emerging semiconductor laser systems. *Journal of Applied Physics*, 2019, vol. 125, no. 8, pp. 082538. https://doi.org/10.1063/1.5063710
12. Dong H., Sun J., Ma S., Liang J., Jia Z., Liu X., Xu B. Interfacial relaxation analysis of InGaAs/GaAsP strain-compensated multiple quantum wells and its optical property. *Superlattices and Microstructures*, 2018, vol. 114, pp. 331–339. https://doi.org/10.1016/j.spmi.2017.12.049
13. Jasik A., Wnuk A., Wójcik-Jedlińska A., Jakieła R., Muszalski J., Strupiński W., Bugajski M. The influence of the growth temperature and interruption time on the crystal quality of InGaAs/GaAs QW structures grown by MBE and MOCVD methods. *Journal of Crystal Growth*, 2008, vol. 310, no. 11, pp. 2785–2792. https://doi.org/10.1016/j.jcrysgro.2008.02.018
14. Jasik A., Wnuk A., Gaca J., Wójcik M., Wójcik-Jedlińska A., Muszalski J., Strupiński W. The influence of the growth rate and V/III ratio on the crystal quality of InGaAs/GaAs QW structures grown by MBE and MOCVD methods. *Journal of Crystal Growth*, 2009, vol. 311, no. 19, pp. 4423–4432. https://doi.org/10.1016/j.jcrysgro.2009.07.032
15. Wang Q., Wang H., Zhang B., Wang X., Liu W., Wang J., Wang J., Fan J., Zou Y., Ma X. Integrated fabrication of a high strain InGaAs/GaAs quantum well structure under variable temperature and improvement of properties using MOCVD technology. *Optical Materials Express*, 2021, vol. 11, no. 8, pp. 2378–2388. https://doi.org/10.1364/OME.431015
16. Chuev M.A., Lomov A.A., Imamov R.M. Simultaneous analysis of double-crystal X-ray rocking curves from a set of crystallographic planes. *Crystallography Reports*, 2006, vol. 51, no. 2, pp. 178–191. https://doi.org/10.1134/S1063774506020027
17. Asryan L.V. Spontaneous radiative recombination and nonradiative Auger recombination in quantum-confined heterostructures. *Quantum Electronics*, 2005, vol. 35, no. 12, pp. 1117–1120. https://doi.org/10.1070/QE2005v035n12ABEH013093
18. Gurioli M., Vinattieri A., Colocci M., Deparis C., Massies J., Neu G., Bosacchi A., Franchi S. Temperature dependence of the radiative and nonradiative recombination time in GaAs/$Al_xGa_{1-x}As$ quantum-well structures. *Physical Review B*, 1991, vol. 44, no. 7, pp. 3115–3124. https://doi.org/10.1103/PhysRevB.44.3115

Авторы

Папылев Денис Сергеевич — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 59241769900, https://orcid.org/0009-0001-3683-5558, dspapylev@itmo.ru

Бабичев Андрей Владимирович — кандидат физико-математических наук, старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 56819964000, https://orcid.org/0000-0002-3463-4744, a.babichev@itmo.ru

Гладышев Андрей Геннадьевич — кандидат физико-математических наук, главный специалист по эпитаксии, ООО «Коннектор Оптикс», 194292, Санкт-Петербург, Российская Федерация, sc 6701588481, https://orcid.org/0000-0002-9448-2471, andrey.gladyshev@connector-optics.com

Карачинский Леонид Яковлевич — доктор технических наук, ведущий научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 6603558355, https://orcid.org/0000-0002-5634-8183, lkarachinsky@itmo.ru

Андрюшкин Владислав Васильевич — кандидат технических наук, ведущий специалист по эпитаксии, ООО «Коннектор Оптикс», 194292, Санкт-Петербург, Российская Федерация; старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57209329441, https://orcid.org/0000-0002-7471-8627, vvandriushkin@itmo.ru

Новиков Иннокентий Игоревич — кандидат физико-математических наук, старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 7202658480, https://orcid.org/0000-0003-1983-0242, innokenty.novikov@itmo.ru

Ефимов Арсений Олегович — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0009-2212-0437, aoefimov@itmo.ru

Егоров Антон Юрьевич — доктор физико-математических наук, технический директор, ООО «Коннектор Оптикс», 194292, Санкт-

Authors

Denis S. Papylev — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 59241769900, https://orcid.org/0009-0001-3683-5558, dspapylev@itmo.ru

Andrey V. Babichev — PhD (Physics & Mathematics), Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 56819964000, https://orcid.org/0000-0002-3463-4744, a.babichev@itmo.ru

Andrey G. Gladyshev — PhD (Physics & Mathematics), Chief Epitaxy Specialist, Connector Optics LLC, Saint Petersburg, 194292, Russian Federation, sc 6701588481, https://orcid.org/0000-0002-9448-2471, andrey.gladyshev@connector-optics.com

Leonid Ya. Karachinsky — D.Sc., Leading Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 6603558355, ttps://orcid.org/0000-0002-5634-8183, lkarachinsky@itmo.ru

Vladislav V. Andryushkin — PhD, Leading Specialist in Epitaxy, Connector Optics LLC, Saint Petersburg, 194292, Russian Federation; Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57209329441, https://orcid.org/0000-0002-7471-8627, vvandriushkin@itmo.ru

Innokenty I. Novikov — PhD (Physics & Mathematics), Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 7202658480, https://orcid.org/0000-0003-1983-0242, innokenty.novikov@itmo.ru

Arseny O. Efimov — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0009-2212-0437, aoefimov@itmo.ru

Anton Yu. Egorov — D.Sc. (Physics & Mathematics), Chief Technology Officer, Connector Optics LLC, Saint Petersburg, 194292, Russian

Петербург, Российская Федерация, SC 7202362169, https://orcid.org/0000-0002-0789-4241, anton.egorov@connector-optics.com
Надточий Алексей Михайлович — кандидат физико-математических наук, ведущий научный сотрудник, Национальный исследовательский университет «Высшая школа экономики», Санкт-Петербург, 190008, Российская Федерация, SC 26642604500, https://orcid.org/0000-0003-0982-907X, al.nadtochy@mail.ioffe.ru

Federation, SC 7202362169, https://orcid.org/0000-0002-0789-4241, anton.egorov@connector-optics.com
Alexey M. Nadtochiy — PhD (Physics & Mathematics), Leading Researcher, National Research University Higher School of Economics, Saint Petesburg, 190008, Russian Federation, SC 26642604500, https://orcid.org/0000-0003-0982-907X, al.nadtochy@mail.ioffe.ru

Статья поступила в редакцию 30.09.2025
Одобрена после рецензирования 10.10.2025
Принята к печати 21.11.2025

Received 30.09.2025
Approved after reviewing 10.10.2025
Accepted 21.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1067-1079
УДК 534-16

Оптимизации конструкции резонансно-частотного МЭМС датчика давления на основе численного моделирования

Сергей Евгеньевич Вторушин[1]✉, Алёна Алексеевна Таловская[2], Евгений Сергеевич Барбин[3], Иван Владимирович Кулинич[4], Мария Сергеевна Вайсбеккер[5]

[1,2,3,4,5] Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация
[3] Институт оптики атмосферы им. В.Е. Зуева Сибирского отделения Российской академии наук, Томск, 634055, Российская Федерация

[1] sergei.e.vtorushin@tusur.ru✉, https://orcid.org/0009-0008-9407-9552
[2] alena.a.talovskaia@tusur.ru, https://orcid.org/0009-0001-6796-1135
[3] bar_es@mail.ru, https://orcid.org/0000-0001-5904-0216
[4] kulinich@tusur.ru, https://orcid.org/0009-0001-0768-3495
[5] mariya_vaisbekker@mail.ru, https://orcid.org/0009-0002-1078-4649

Аннотация
Введение. Кремниевые микроэлектромеханические датчики давления резонансно-частотного типа выделяются высокой линейностью и стабильностью выходных характеристик, что делает их особенно перспективными для прецизионных измерений. Представлено исследование влияния геометрии и напряженно-деформированного состояния мембран на чувствительность резонансно-частотных датчиков давления, а также разработка рекомендаций по оптимальному размещению резонаторов и выбору технологического маршрута формирования мембран. **Метод.** С использованием трехмерных моделей мембран различных геометрических форм выполнено численное моделирование их напряженно-деформированного состояния под действием статического давления с помощью метода конечных элементов. Метод позволяет выявить зоны локализации деформаций, наиболее подходящие для размещения резонаторов. Для изготовления тестовых образцов кремниевых мембран применено жидкостное травление с предварительным утонением пластин и последующей финишной механической обработкой. **Основные результаты.** Показано, что максимальная чувствительность достигается при позиционировании резонатора в зонах пиковых растягивающих и сжимающих напряжений. Проведен анализ зависимости формы мембраны на распределение напряжений и отклик резонатора, что позволило выявить оптимальные места расположения резонатора с точки зрения технологических допусков и чувствительности. Выполнено сравнение методов подготовки мембраны: химическое и механическое утонение с последующей полировкой. На основании проведенных измерений параметров шероховатости для мембран, изготовленных разными методами, описана более оптимальная технология их подготовки. **Обсуждение.** Полученные результаты позволяют оптимизировать геометрию и технологический процесс изготовления резонансно-частотного датчика давления, что способствует повышению чувствительности, расширению технологических допусков, снижению затрат на производство и повышению надежности устройств в промышленной эксплуатации.

Ключевые слова
мембрана, резонатор, механические напряжения, чувствительность, травление, шероховатость

Благодарности
Результаты исследований получены при поддержке Министерства науки и высшего образования Российской Федерации (тема № FEWM-2024-0008) (численное моделирование, расчет допусков).
Работа выполнена в рамках программы развития Томского государственного университета систем управления и радиоэлектроники на 2025–2036 годы программы стратегического академического лидерства «Приоритет 2030» (механическое утонение мембраны).

© Вторушин С.Е., Таловская А.А., Барбин Е.С., Кулинич И.В., Вайсбеккер М.С., 2025

Ссылка для цитирования: Вторушин С.Е., Таловская А.А., Барбин Е.С., Кулинич И.В., Вайсбеккер М.С. Оптимизации конструкции резонансно-частотного МЭМС датчика давления на основе численного моделирования // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1067–1079. doi: 10.17586/2226-1494-2025-25-6-1067-1079

Optimization of the resonant frequency MEMS pressure sensor based on numerical simulation

Sergey E. Vtorushin[1]✉, Alena A. Talovskaia[2], Evgenii S. Barbin[3], Ivan V. Kulinich[4], Mariya S. Vaisbekker[5]

[1,2,3,4,5] Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation
[3] V.E. Zuev Institute of Atmospheric Optics, RAS, Siberian Branch, Tomsk, 634055, Russian Federation

[1] sergei.e.vtorushin@tusur.ru✉, https://orcid.org/0009-0008-9407-9552
[2] alena.a.talovskaia@tusur.ru, https://orcid.org/0009-0001-6796-1135
[3] bar_es@mail.ru, https://orcid.org/0000-0001-5904-0216
[4] kulinich@tusur.ru, https://orcid.org/0009-0001-0768-3495
[5] mariya_vaisbekker@mail.ru, https://orcid.org/0009-0002-1078-4649

Abstract
Silicon microelectromechanical pressure sensors of the resonant-frequency type are distinguished by high linearity and stability of their output characteristics, making them particularly promising for precision measurements. This paper presents a study of the influence of membrane geometry and stress-strain state on the sensitivity of resonant-frequency pressure sensors. Recommendations for optimal resonator placement and the selection of a process route for membrane formation are also developed. Using three-dimensional models of membranes of various geometric shapes, numerical simulation of their stress-strain state under static pressure was performed using the finite element method. This method allowed us to identify the zones of localized deformation most suitable for resonator placement. Wet etching with preliminary wafer thinning and subsequent finishing machining was used to fabricate test samples of silicon membranes. It is shown that maximum sensitivity is achieved by positioning the resonator in zones of peak tensile and compressive stresses. An analysis of the membrane shape relationship to stress distribution and resonator response was conducted, enabling the identification of optimal resonator locations in terms of manufacturing tolerances and sensitivity. Membrane preparation methods were compared: chemical and mechanical thinning followed by polishing. Based on roughness measurements for membranes manufactured using different methods, the optimal preparation technology was described. The obtained results enable optimization of the geometry and manufacturing process of the resonant-frequency pressure sensor, which contributes to increased sensitivity, wider manufacturing tolerances, reduced production costs, and improved reliability in industrial operation.

Keywords
membrane, resonator, mechanical stress, sensitivity, etching, roughness

Acknowledgements
The results were obtained within the state assignment of the Ministry of Science and Higher Education of the Russian Federation (theme no. FEWM-2024-0008) (numerical modeling, tolerance calculation).
This project was executed within the scope of the TUSUR Development Program for 2025-2036, as part of the Strategic Academic Leadership Program, “Priority 2030” (mechanical thinning of the membrane).

For citation: Vtorushin S.E., Talovskaia A.A., Barbin E.S., Kulinich I.V., Vaisbekker M.S. Optimization of the resonant frequency MEMS pressure sensor based on numerical simulation. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1067–1079 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1067-1079

Введение

Микроэлектромеханические (МЭМС) датчики давления находят широкое применение в промышленности, энергетике и нефтегазовом секторе. Особое место среди них занимают резонансно-частотные датчики давления. В отличие от емкостных и тензорезистивных, резонансно-частотные датчики давления обладают высокой стабильностью, практически полным отсутствием гистерезиса, но при этом имеют более низкую чувствительность [1–5]. Актуальной задачей является повышение чувствительности резонансно-частотного датчика давления путем оптимизации геометрии их чувствительного элемента, совершенствованием технологических процессов, проектирования схемотехнических и алгоритмических методов обработки первичной информации и других. Особое внимание уделяется вопросам технологических погрешностей изготовления, так как эти ошибки могут существенно снижать чувствительность датчика [6].

На сегодняшний день известно достаточное количество научных публикаций, связанных с исследованием нагруженных мембран МЭМС датчиков давления и степенью их влияния на чувствительность первичных преобразователей тензорезистивного, емкостного, частотно-резонансного и других типов.

В работах [5, 7–10] представлены результаты по оптимизации геометрии мембран и первичных преобразователей. Представлено решение задачи поиска более оптимальной формы мембраны на основании анализа их напряженно-деформированного состояния, а также поиска наиболее эффективного расположения преобразователей на мембранах. В [5, 7–10] выполнен анализ мембран квадратной и круглой форм, а также

мембраны с модификацией их планарной поверхности путем введения концентраторов напряжений.

В [4, 11–13] приведено подробное описание напряженно-деформированного состояния мембраны МЭМС датчика давления и обобщенные математические модели нагруженных статическим давлением мембран круглой и квадратной форм. Рассмотрены модели, описывающие поведение тензорезистивных преобразователей в зависимости от их расположения на мембране. Полученные математические модели позволяют аналитически достаточно быстро построить картину распределения прогибов и напряжений по поверхности мембран с учетом их геометрических особенностей.

Большинство работ по исследованию геометрии мембран МЭМС датчиков давления, как правило, ограничены двумя-тремя формами и направлены на описание математических моделей нагруженного состояния мембран. В настоящей работе исследован расширенный набор вариантов геометрии мембран с целью определения наиболее подходящей формы для обеспечения не только высокой чувствительности, но и возможности расширения технологических допусков позиционирования резонаторов на мембране. В работе поставлена следующая задача: методами численного моделирования определить степень влияния механического напряжения мембраны на чувствительность резонансно-частотного преобразователя в зависимости от геометрии мембраны и места его расположения. Метод конечных элементов, который относится к численным методам, в значительной степени сокращает время расчета, обеспечивая при этом высокую точность.

Актуальность темы данной работы подтверждается объемом мировых научных публикаций в области разработки МЭМС датчиков давления, где задачи повышения чувствительности и снижение технологических погрешностей остаются ключевыми задачами при проектировании МЭМС. Новизна настоящей работы основана на анализе работ по схожей тематике, в которых недостаточно уделено внимание оценке возможных погрешностей расположения первичного преобразователя и как эта погрешность влияет на точность чувствительного элемента датчика. При формировании резонаторов рассовмещение топологий может достигать десятков микрометров в зависимости от используемого оборудования. Это приводит к ошибкам позиционирования первичного преобразователя на мембране и снижению его чувствительности.

Проведенные исследования позволят определить наиболее эффективное расположение резонатора на мембране на основе анализа градиента напряжений в мембранах. Наиболее эффективным расположением будет считаться такое, при котором сохраняется высокая чувствительность и при этом будут расширены допуски технологических процессов совмещения топологий при формировании резонатора.

Описание модели и метода исследования

В качестве предмета исследования выбрана трехмерная модель кристалла квадратной формы с размерами сторон 3 мм и толщиной 500 мкм. В кристалле сформирована мембрана определенной геометрии площадью 1,96 мм^2 и толщиной 110 мкм. На плоской стороне мембраны расположен резонатор, представляющий из себя балку длиной 100 мкм с площадью сечения 0,5 мкм^2. Материал модели (мембрана и резонатор) — монокристаллический кремний с ориентацией пластины в направлении (100).

В качестве форм для мембраны выбрано 8 различных форм, три из которых наиболее часто встречающихся в МЭМС датчиках давления [4, 5]: квадратная, круглая, круглая с жестким центром. Также для анализа выбраны 5 дополнительных форм: треугольная, прямоугольная, ромбовидная, эллиптическая, гексагональная. Выбор таких форм обусловлен системным подходом, позволяющим охватить особенности ключевых форм мембран в МЭМС датчиках давления. Предложенные формы мембран позволят оценить характер градиента напряжений в зависимости от количества сторон, соотношения размеров сторон, их ориентации относительно резонатора, плавности контура и других особенностей формы мембраны.

Исследование модели проводилось методом конечных элементов в программе ANSYS. Для определения чувствительности использовался связанный анализ, включающий в себя определение отклика мембраны на статическую нагрузку с дальнейшим расчетом собственных частот и форм колебаний резонатора при напруженной и ненагруженной мембране. В качестве граничных условий использовалось ограничение трех степеней свободы для основания кристалла. В качестве статической нагрузки на мембрану прикладывалось давление 100 кПа. Такое давление обеспечивает деформацию мембраны в пределах 1 мкм (менее 1 % от ее толщины), а напряжения, возникающие в мембране и резонаторе, не превышают предела прочности кремния с учетом коэффициента запаса по прочности $SF = 2$. При таких условиях имеет место линейный характер и обратимость деформаций, что позволяет руководствоваться линейной теорией упругости. На рис. 1 приведен пример конечно-элементной модели кристалла с резонатором на квадратной мембране.

Принцип работы такого чувствительного элемента основан на частотно-резонансном методе детектирования. При воздействии давления мембрана изгибается, из-за чего резонатор испытывает продольные деформации, и его собственная частота изменяется. При этом смещение собственной частоты резонатора будет пропорционально приложенному давлению к мембране.

Исследование напряженного состояния нагруженных мембран

В процессе исследования образцов мембран различных форм резонатор располагался на их поверхностях сначала в центре, затем перемещался на шаг 2 мкм в направлении периферии мембраны. В каждой точке расположения резонатора рассчитывалась его чувствительность, а также строился график напряжений на мембране вдоль этого направления.

Для резонатора чувствительность определяется по сдвигу собственной частоты, возникающему при появ-

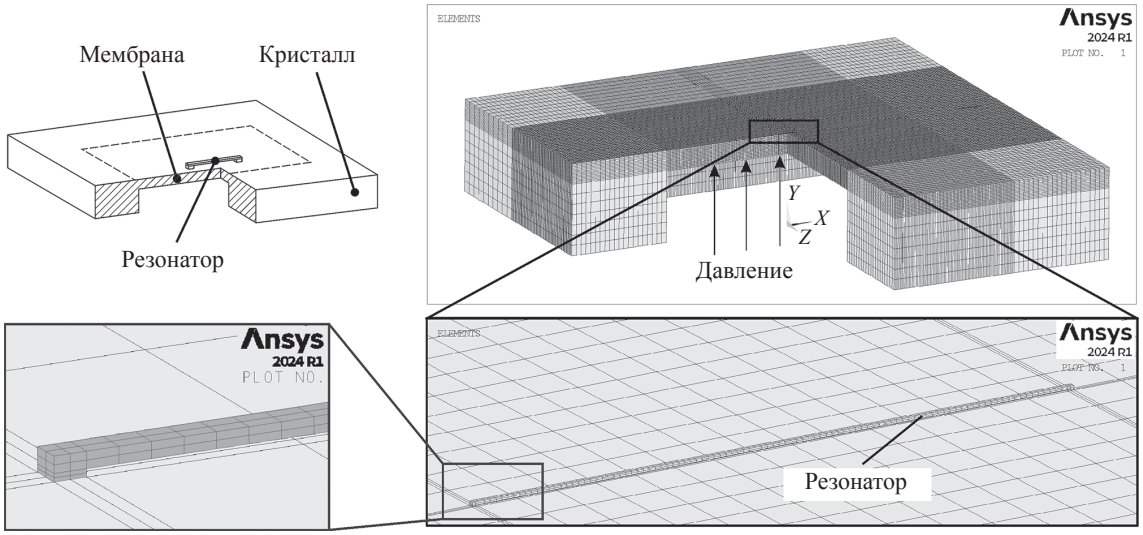

Рис. 1. Конечно-элементная модель мембраны квадратной формы

Fig. 1. Finite-element model of the square membrane

лении в них механических напряжений. Эти напряжения образуются в резонаторах в результате деформации мембраны под действием внешнего измеряемого давления. Смещение собственной частоты основного тона резонатора имеет вид

$$f = f_0 \sqrt{1 \pm \frac{\sigma}{\sigma_e}},$$

где f_0 — собственная частота резонатора при отсутствии напряжений; σ — напряжения в резонаторе; σ_e — критические напряжения, приводящие резонатор к потере устойчивости [14].

Согласно данному выражению, добиться максимального смещения собственной частоты можно, обеспечив накопление максимальных напряжений в резонаторе. Так как деформации от мембраны передаются на резонатор, то, чем больше деформации мембраны в зоне его расположения, тем больше деформации в самом резонаторе. Именно поэтому важно иметь представление о картине распределения напряжений по мембране. При этом расположение резонатора в зонах пиковых значений механических напряжений может негативно сказываться на чувствительности резонатора из-за технологических погрешностей совмещения топологий при литографии. Картина градиента напряжений на мембране может содержать несколько локальных максимумов, которые могут существенно различаться по характеру их изменения по поверхности мембраны. Для некоторых экстремумов характерен резкий перепад напряжений на малом расстоянии, в то время как для других — плавный спад, создающий обширную область с устойчиво высоким уровнем напряжений.

На основании теории пластин и оболочек, распределение напряжений на мембране такое, что максимальные растягивающие напряжения локализуются в центре мембраны (рис. 2, *а*: точка с координатой 0), а максимальные сжимающие напряжения — по периферии мембраны (рис. 2, *а*: локальные экстремумы в окрестностях координаты, равной 1), что подтверждается результатами конечного-элементного анализа.

На рис. 2 для каждой исследуемой мембраны приведены диаграммы распределения напряжений, а также графики нормальных напряжений вдоль ключевых направлений, представляющих интерес для исследования. В качестве ключевых направлений выбраны радиальные направления от центра к периферии мембран. Радиальное размещение пары резонаторов вдоль этих направлений позволяет реализовать дифференциальное детектирование полезного сигнала с датчика.

Для квадратной мембраны (рис. 2, *а*) можно выделить два ключевых направления: перпендикулярное одной из сторон мембраны (*X*-направление) и диагональное (*XZ*-направление). На графике для квадратной формы мембраны (рис. 2, *а*) отчетливо видна разница распределения напряжений в зависимости от направления. График распределения напряжений представляет собой нормированную кривую по значениям напряжений и координатам. Началу координат соответствует геометрический центр мембраны, где для симметричных конструкций локализуются максимумы растягивающих напряжений (положительные значения). Нормированная координата, равная единице, соответствует периферии мембраны — область жесткого закрепления. В *X*-направлении, вблизи области закрепления, концентрируются сжимающие напряжения, характеризующиеся отрицательными значениями. В диагональном направлении (*XZ*-направление), где нормированной координате, равной 1, соответствует угол мембраны, область сжимающих напряжений имеет более плавный градиент, а сами напряжения меньше по модулю, чем в *X*-направлении. При этом максимум сжатия смещен от угла к центру.

Похожая тенденция наблюдается и у мембран других геометрических форм. Для многоугольных мембран (рис. 2, *а–е*) центральная часть характеризуется растягивающими напряжениями, а крутизна графика в этих зонах зависит от выбранного направления. В областях

сжимающих напряжений, направления, перпендикулярные сторонам, демонстрируют более резкий спад напряжений, в то время как диагональные направления имеют меньший градиент, а сами сжимающие напряжения ниже по величине.

Такая картина распределения напряжений позволяет прогнозировать оптимальные зоны размещения резонатора. Для обеспечения высокой чувствительности резонатор следует размещать в радиальных направлениях, перпендикулярных сторонам мембраны,

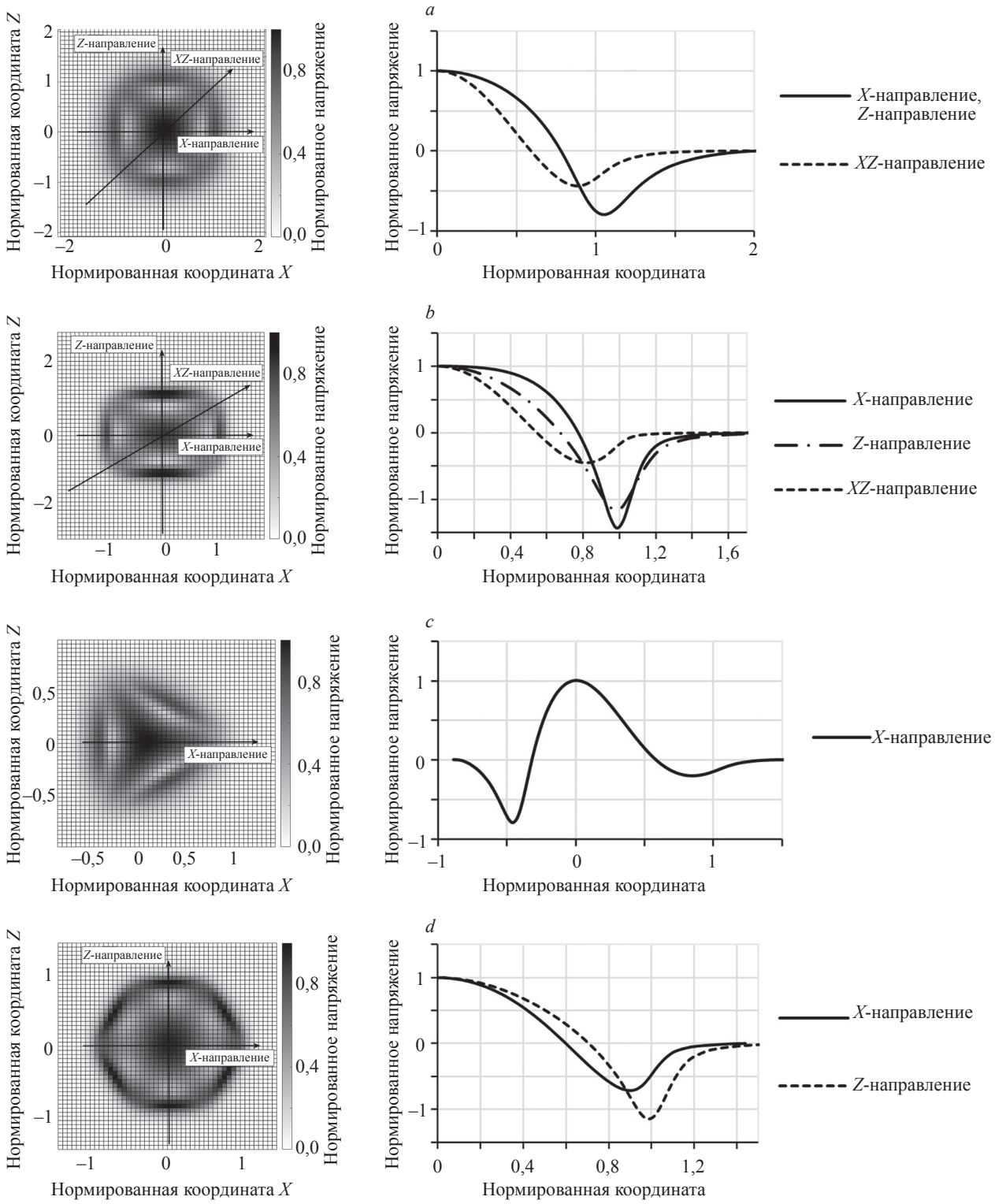

Рис. 2. Диаграммы и графики напряженного состояния мембран различных форм: квадратная (*a*); прямоугольная (*b*); треугольная (*c*); гексагональная (*d*); ромбовидная (*e*); эллиптическая (*f*); круглая (*g*); круглая с жестким центром (*h*)

Fig. 2. Pictures and profiles of the stress state of various shapes membranes: square (*a*); rectangular (*b*); triangular (*c*); hexagonal (*d*); rhomb-shaped (*e*); elliptical (*f*); round (*g*); round with rigid center (*h*)

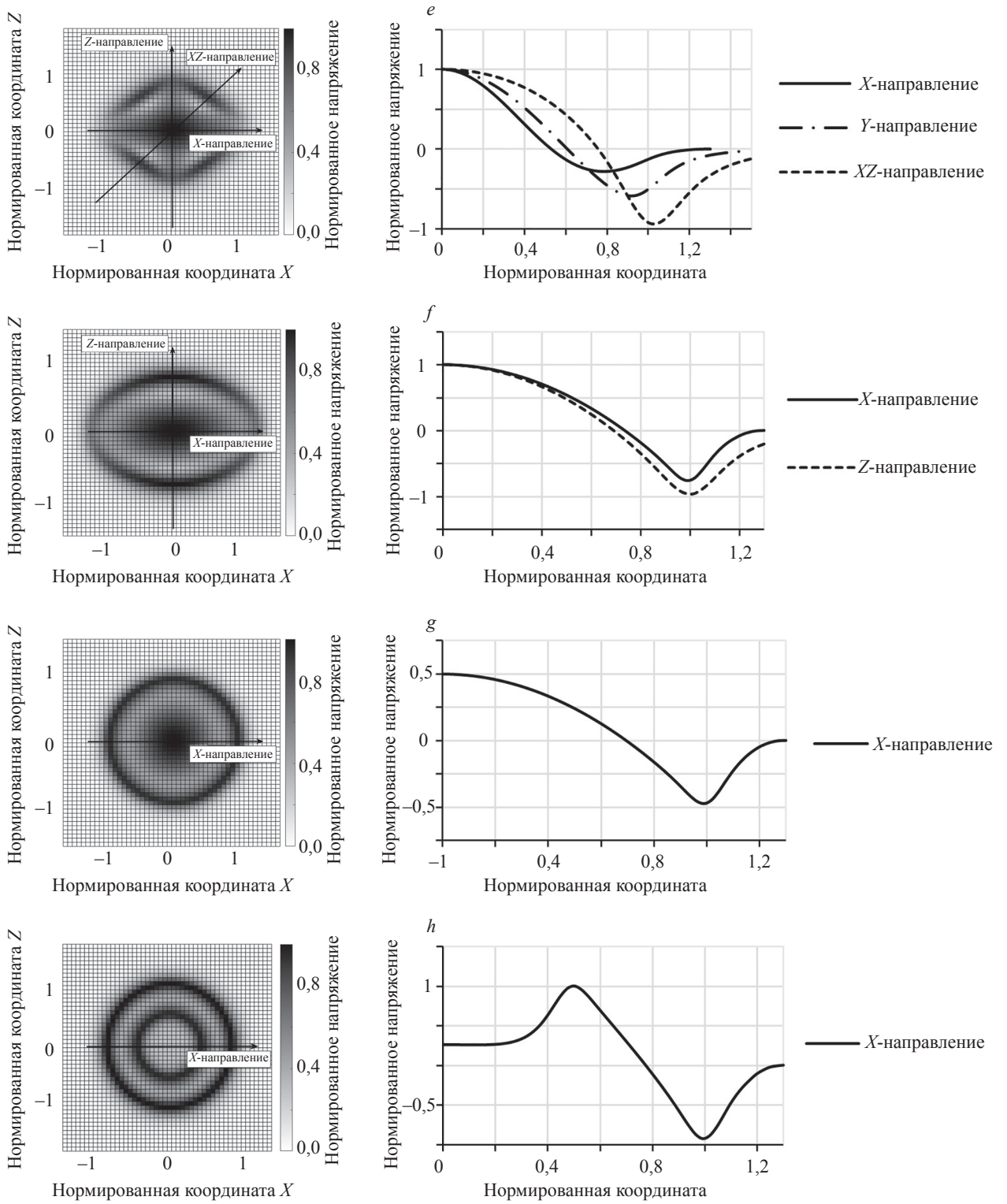

Рис. 2. Продолжение

Fig. 2. Continuation

где напряжения по модулю максимальны. Однако при размещении резонатора на периферии возникают повышенные требования к точности совмещения топологий во время технологического процесса изготовления. Ошибка позиционирования может привести к смещению рисунка топологии резонатора от области максимальных напряжений и, как следствие, к снижению чувствительности. При размещении резонатора в диагональных направлениях требования к допускам могут быть менее жесткими, поскольку градиент напряжений ниже, и даже при небольшой величине рассовмещения топологий чувствительность остается высокой.

Оценку чувствительности резонатора в зависимости от его расположения вдоль ключевых направле-

ний можно провести путем расчета его собственной частоты в нагруженном и ненагруженном состояниях. Величина чувствительности при этом может быть рассчитана следующим образом:

$$S = \frac{f - f_0}{P},$$

где f_0 — собственная частота ненагруженного резонатора; f — собственная частота резонатора при нагрузке мембраны статическим давлением; P — приложенное статическое давление к мембране.

На рис. 3 представлены кривые чувствительности резонатора в зависимости от его положения на поверхности мембраны.

Графики чувствительности показывают корреляцию величин чувствительности с напряжениями на мембранах. Максимумы напряжений соответствуют максимумам чувствительности, что согласуется с предположениями о размещении резонатора в местах концентрации напряжений на мембране.

На основании полученных данных можно оценить технологические допуски на позиционирование резонатора в зависимости от его расположения на мембране и формы мембраны. Такую оценку можно сделать на основании критерия максимально допустимого спада чувствительности. Критической величиной спада чувствительности можно считать величину температурного дрейфа датчика, который, как правило, составляет 1–10 % от чувствительности датчика к давлению [15–18]. Критический спад чувствительности, используемый для расчета допусков, целесообразно устанавливать на уровне фундаментальных погрешностей датчика, таких как температурный дрейф. Это позволяет гарантировать, что производственный разброс не станет доминирующим источником погрешности. Исходя из этого, можно определить границы технологического допуска на позиционирование резонатора, в пределах которого спад чувствительности по давлению не будет превышать 10 %. В соответствии с рис. 3 рассчитаны величины допустимых погрешностей для некоторых вариантов мембран при двух вариантах критической величины спада чувствительности: 1 % и 10 % (табл. 1).

На основании полученных результатов (табл. 1) и графиков чувствительности (рис. 3) можно сделать вывод, что допуски на совмещение топологий будут тем шире, чем плавнее градиент механических напряжений в окрестностях планируемого расположе-

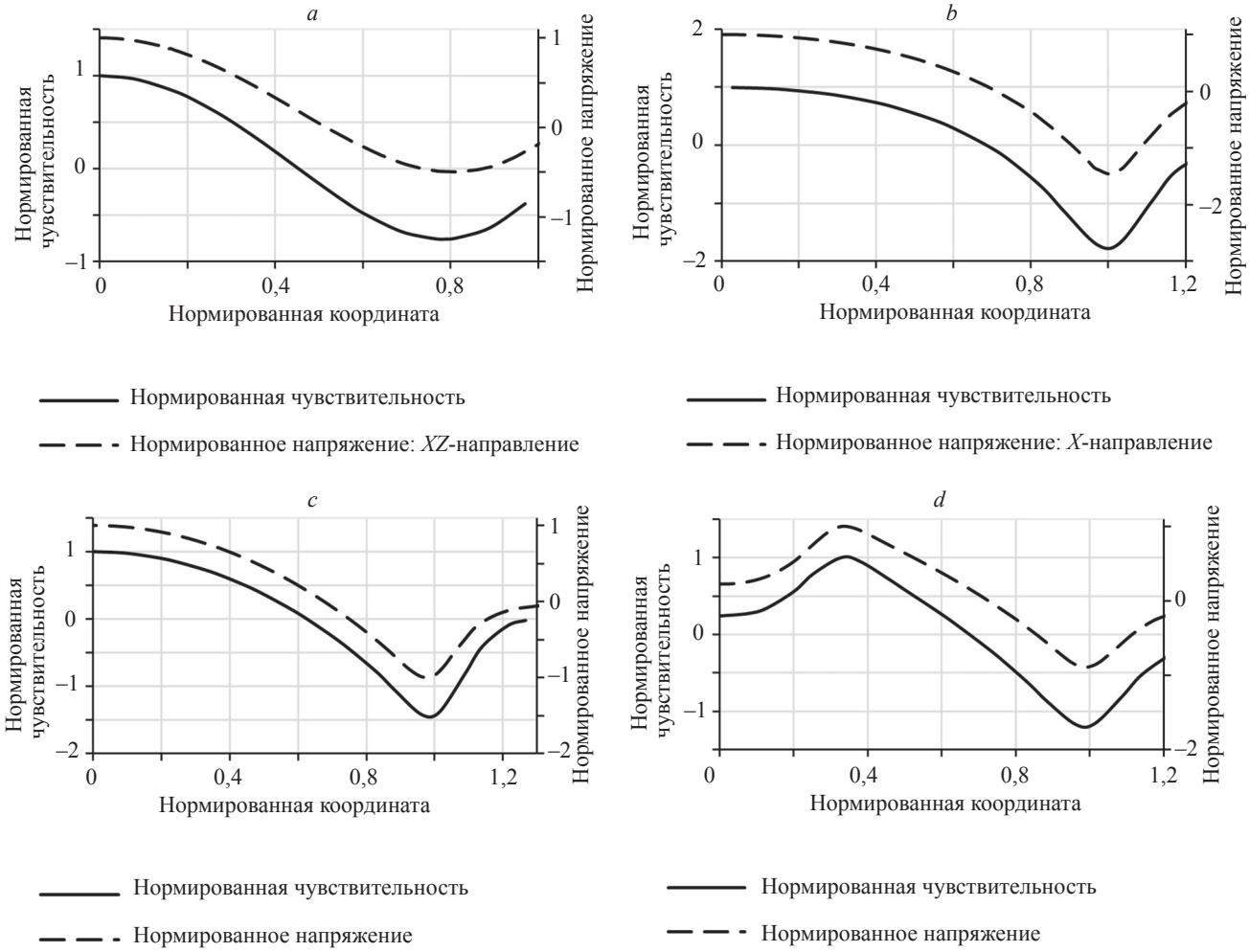

Рис. 3. Графики чувствительности для мембран: квадратной формы (*XZ*-направление) (*a*); квадратной формы (*X*-направление) (*b*); круглой (*c*); круглой с жестким центром (*d*)

Fig. 3. Sensitivity diagrams for membranes: square *XZ*-direction (*a*), square *X*-direction (*b*), round (*c*), round with a rigid center (*d*)

Таблица 1. Сравнительная таблица характеристик чувствительного элемента
Table 1. Comparison table of the sensitive element characteristics

Параметр	Форма мембраны							
	Квадратная, 1,4 × 1,4 мм (*X*-направление)		Квадратная, 1,4 × 1,4 мм (*XZ*-направление)		Круглая, R = 0,7898 мм		Круглая с жестким центром, R_1 = 0,5 мм, R = 0,7898 мм	
Критическая величина спада чувствительности, %	1	10	1	10	1	10	1	10
Технологический допуск (центральное расположение), мкм	±40,3	±177,5	±29,9	±127,4	±39,1	±155,9	±14,6	±45,0
Технологический допуск (краевое расположение), мкм	±11,9	±37,8	±26,5	±84,8	±13,0	±42,2	±13,8	±45,0
Абсолютная чувствительность (центральное расположение), Гц/кПа	144,5		144,3		159,4		169,3	
Абсолютная чувствительность (краевое расположение), Гц/кПа	255,2		109,2		226,9		198,2	

Примечание. R — внешний радиус круглой мембраны и мембраны с жестким центром, R_1 — радиус жесткого центра.

ния резонатора. Например, для квадратной мембраны (1,4 × 1,4 мм) при критическом спаде чувствительности в 1 %, допустимое смещение резонатора в центральной зоне при его ориентации вдоль *X*-направления составляет ±40,3 мкм, в то время как при ориентации в *XZ*-направлении — ±29,9 мкм. В случае размещения резонатора на периферии мембраны допуски шире при ориентации резонатора в *XZ*-направлении: ±26,5 мкм по сравнению с ±11,9 мкм в *X*-направлении. В то же время краевое расположение резонатора характеризуется заметно более высоким уровнем чувствительности 255,2 Гц/Па при ориентации вдоль *X*-направления в сравнении с 109,2 Гц/Па при ориентации вдоль *XZ*-направления.

Круглая мембрана, как показали расчеты, является наиболее технологичной. Ее центральная зона характеризуется высокими допусками (±155,9 мкм), что снимает проблему ориентации резонатора по причине симметрии напряжений вдоль любого радиального направления. Однако при периферийном размещении допуски резко снижаются (±42,2 мкм), становясь сравнимыми с наихудшими случаями для других форм мембран.

Результаты изготовления тестовых образцов

Для формирования мембраны резонансно-частотного датчика давления была выбрана квадратная форма с диагональным расположением резонаторов. Технологически такая конструкция мембраны может быть изготовлена целиком из кремния методами жидкостного или, например, плазменного травления. Резонаторы изготавливаются также из кремния или другого материала методами электронно-лучевого испарения. Основной технологической операцией при производстве такой мембраны является создание упругой области на кристалле кремния. В отличие от плазмохимического травления, жидкостные травители обладают большей доступностью, обеспечивают селективность при анизотропном травлении кремния ориентации (100), а также являются более экономически эффективными при обработке на значительную глубину.

В качестве маски для жидкостного травления применен оксид кремния толщиной 0,5 мкм, сформированный в атмосфере сухого кислорода [19]. В качестве травителя использован тетраметиламмония гидроксид (ТМАГ) [20], который позволяет травить пластину кремния (100) со скоростью до 0,5 мкм/мин при температуре 70 °C. При этом время травления пластины толщиной 450 мкм на глубину 250 мкм составляет 15 ч, что снижает эффективность данного метода. Для оценки возможности ускорения процесса травления было опробовано два подхода:

— повышение температуры травления до 90 °C — увеличивает скорость до 2 мкм/мин, однако снижает селективность, требуя утолщения маски;
— использование травителя на основе гидроксида калия (KOH) — повышает скорость до 8 мкм/мин, сохраняя селективность процесса [21].

Помимо увеличения скорости травления, сокращение времени технологической обработки кремния без ухудшения характеристик устройства может быть достигнуто за счет предварительного утонения кремниевой пластины.

Для утонения кремния использовался ультразвуковой метод (УЗМ) в 32 % растворе KOH, что позволило равномерно уменьшить толщину кремниевой подложки без образования крупных дефектов [22, 23]. Время травления при этом составило приблизительно 1,5 ч. Добавление в раствор изопропилового спирта (ИПС) в соотношении 20 % KOH:80 % (H_2O:ИПС = 4:1) привело к увеличению продолжительности травления и визуальному повышению шероховатости поверхности (рис. 4).

Оценка качества подготовленной поверхности проводилась с использованием сканирующего электронного микроскопа Raith 150. Наилучших результатов с точки зрения однородности и шероховатости полученной поверхности удалось добиться при утонении кремниевой пластины в 32 % растворе KOH с использованием ультразвуковой ванны.

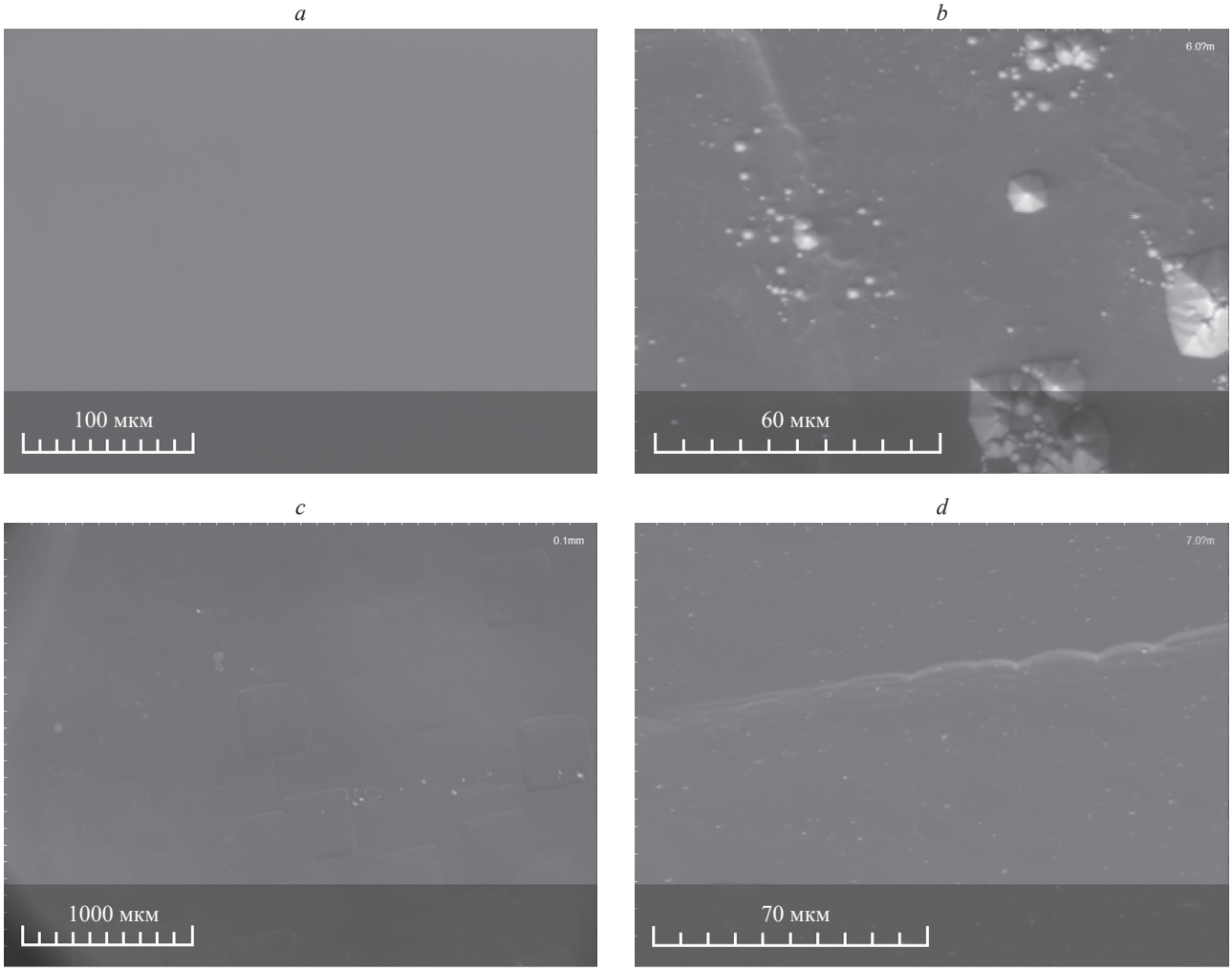

Рис. 4. Фотографии поверхности кремниевой пластины сканирующим электронным микроскопом: до утонения (*a*), утонение с использованием 32 % KOH (*b*), утонение с использованием 32 % KOH + УЗМ (*c*), утонение с использованием 32 % KOH + УЗМ+ ИПС (*d*)

Fig. 4. Scanning electron microscope photograph of the surface of a silicon wafer: before thinning (*a*), thinning with 32 % KOH solution (*b*), thinning with 32 % KOH solution + USB (*c*), thinning with 32 % KOH solution + USM + IPA (*d*)

В качестве альтернативных методов утонения применялись механическая обработка на шлифовальном станке и механическое утонение с последующей полировкой. В результате были получены четыре образца мембран на пластинах, обработанных различными способами утонения. Для химического утонения использовались две пластины, одна из которых имела предварительно полированную поверхность (рис. 5).

Пластины после обработки имеют различную шероховатость (*Ra*) — среднее отклонение от идеальной поверхности, величина которой может существенно влиять на качество формообразования резонаторов на поверхности мембраны. Показатели шероховатости мембран, приведенные в табл. 2, получены с использованием профилометра Profilm 3D.

На основании полученных результатов (табл. 2) можно сделать вывод, что оптимальной технологией

Таблица 2. Параметры шероховатости изготовленных мембран

Table 2. Roughness parameters of manufactured membranes

Метод утонения мембраны	Шероховатость *Ra*, нм	
	вокруг мембраны	внутри мембраны
Химическое утонение (не полированная пластина)	710	530
Химическое утонение (полированная пластина)	190	210
Механическое утонение	320	580
Механическое утонение с последующей полировкой	2	10

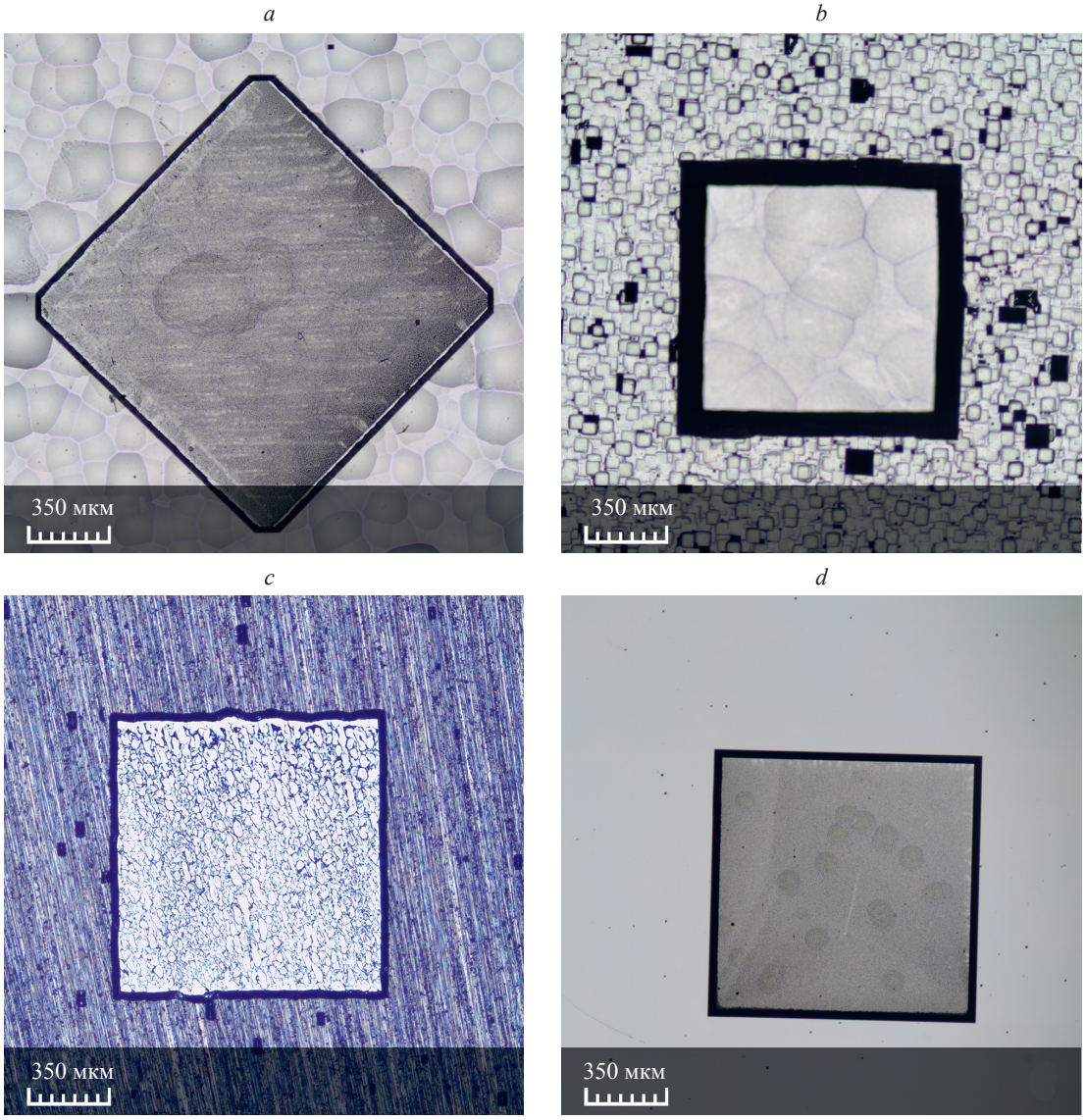

Рис. 5. Образцы мембран резонансно-частотного датчика давления при разных методах утонения: химическое утонение (не полированная пластина) (*a*); химическое утонение (полированная пластина) (*b*); механическое утонение (*c*); механическое утонение с последующей полировкой (*d*)

Fig. 5. Samples of membranes of a resonant frequency pressure sensor with different thinning methods: chemical thinning (non-polished plate) (*a*); chemical thinning (polished plate) (*b*); mechanical thinning (*c*); mechanical thinning followed by polishing (*d*)

формирования мембраны является предварительное механическое утонение пластины с последующей полировкой поверхности, а затем химическое травление с использованием ТМАГ через маску из термически окисленного кремния.

Обсуждение результатов

Исследование напряженно-деформированного состояния мембран показало, что зоны с плавным градиентом напряжений демонстрируют существенно более высокие технологические допуски на совмещение топологий при формировании резонатора, но при этом имеют более низкий уровень чувствительности. Количественно это подтверждается данными из табл. 1. Для квадратной формы мембраны при расположении резонатора в центре его целесообразно ориентировать в *X*-направлении, при размещении на периферии мембраны — в диагональном *XZ*-направлении. В случае, если требуются широкие допуски, диагональное направление может быть более выгодным решением, хотя при этом абсолютная чувствительность ниже.

Круглая мембрана и мембрана с жестким центром на периферии имеют высокую крутизну кривой чувствительности, а следовательно, более узкие допуски, что накладывает более жесткие требования на позиционирование резонатора, однако позволяют добиться наивысшей чувствительности в случае расположения резонатора на периферии мембраны.

Анализ графиков напряжений пяти дополнительных форм (треугольная, прямоугольная, гексагональная, ромбовидная, эллиптическая) позволил выявить общую закономерность: наиболее плавные градиенты напряжений, а следовательно, и наиболее широкие

технологические допуски, наблюдаются в зонах резкого изменения геометрии мембраны, а именно — в углах многогранных структур. В таких областях формируются обширные зоны с относительно высоким и плавно меняющимся напряжением, в отличие от остроконечных экстремумов вдоль прямых сторон. Для данных форм мембран рекомендации по расположению резонаторов будут аналогичны трем предыдущим формам (квадратная, круглая, круглая с жестким центром).

Для реализации опытного образца была выбрана квадратная мембрана на основании результатов в табл. 1. В ходе отработки технологии формирования мембраны было установлено, что комбинированный метод, включающий механическое утонение с последующей полировкой и финишным химическим травлением ТМАГ, позволяет получить поверхность с наименьшей шероховатостью. Этот параметр является критическим, так как непосредственно влияет на качество формирования резонатора и его характеристики.

Заключение

Проведенное исследование демонстрирует решающую роль пространственного градиента механических напряжений, при проектировании микроэлектромеханических датчиков давления. Установлено, что размещение резонатора в зонах с плавным изменением напряжений (например, в окрестностях углов многогранных мембран) позволяет существенно расширить технологические допуски на совмещение топологий при литографии, снизив влияние производственного разброса на чувствительность устройства.

С практической точки зрения, для реализации данного подхода был обоснован выбор мембраны квадратной формы для технологической реализации. Выбранный метод подготовки поверхности мембраны позволит изготавливать образцы чувствительного элемента резонансно-частотного датчика давления с мембранами минимальной шероховатости, что обеспечит больший процент годных образцов после формирования резонаторов на поверхности мембраны.

Результаты исследования с точки зрения практической значимости представляют ценность для инженеров проектировщиков микроэлектромеханических датчиков давления, так как позволяют обеспечить обоснованный подход к проектированию.

Дальнейшие исследования планируется проводить по отработке формообразования резонаторов малой толщины методами плазменного напыления.

Литература

1. Jha C.M., Bahl G., Melamud R., Chandorkar S.A., Hopcroft M.A., Kim B., Agarwal M., Salvia J., Mehta H., Kenny T.W. High resolution microresonator-based digital temperature sensor // Applied Physics Letters. 2007. V. 91. N 7. P. 074101 https://doi.org/10.1063/1.2768629
2. Кудрявцева Д.А. Использование кремниевого резонатора в резонансных преобразователях давления // Труды международного симпозиума “Надежность и качество.” 2015. Т. 2. С. 118–121.
3. Lu Y., Zhang S., Yan P., Li Y., Yu J., Chen D., Wang J., Xie B., Chen J. Resonant pressure micro sensors based on dual double ended tuning fork resonators // Micromachines. 2019. V. 10. N 9. P. 560. https://doi.org/10.3390/mi10090560
4. Андреев К.А., Тиняков Ю.Н., Шахнов В.А.. Математические модели гибридных чувствительных элементов датчиков давления // Датчики и системы. 2013. № 9 (172). С. 2–9.
5. Тиняков Ю.Н., Николаева А.С. О расчете мембран датчиков давления // Вестник МГТУ им. Н.Э. Баумана. Серия Приборостроение. 2015. № 6 (105). С. 135–142. https://doi.org/10.18698/0236-3933-2015-6-135-142
6. Clark S.K., Wise K.D. Pressure sensitivity in anisotropically etched thin-diaphragm pressure sensors // Proc. of the IEEE Transactions on Electron Devices. 1979. V. 26. N 12. P. 1887–1896. https://doi.org/10.1109/T-ED.1979.19792
7. Yu Z., Zhao Y., Li L., Tian B., Li C. Geometry optimization for micropressure sensor considering dynamic interference // Review of Scientific Instruments. 2014. V. 85. N 9. P. 095002. https://doi.org/10.1063/1.4895999
8. Гулиева Д.А., Цыпин Б.В., Кучумов Е.В. Повышение чувствительности струнного первичного преобразователя путем изменения габаритно-массовых характеристик чувствительного элемента // Известия высших учебных заведений. Поволжский регион. Технические науки. 2020. N 3 (55). С. 88–97. https://doi.org/10.21685/2072-3059-2020-3-9
9. Munas F.R., Amarasinghe Y.W.R., Kumarage P., Dao D.V., Dau V.T. Design and simulation of MEMS based piezoresitive pressure sensor for microfluidic applications // Proc. of the Moratuwa Engineering Research Conference (MERCon). 2018. P. 215–220. https://doi.org/10.1109/MERCon.2018.8421908
10. Тун П.В., Симонов Б.М., Тимошенков С.П. Исследование возможностей повышения чувствительности МЭМС-датчика давления емкостного типа с мембранами различных геометрических

References

1. Jha C.M., Bahl G., Melamud R., Chandorkar S.A., Hopcroft M.A., Kim B., Agarwal M., Salvia J., Mehta H., Kenny T.W. High resolution microresonator-based digital temperature sensor. *Applied Physics Letters*, 2007, vol. 91, no. 7, pp. 074101 https://doi.org/10.1063/1.2768629
2. Kudriavtceva D.A. Use of silicon resonator in resonant pressure transducers. *Transactions of the International Symposium on Reliability and Quality*, 2015, vol. 2, pp. 118–121. (in Russian)
3. Lu Y., Zhang S., Yan P., Li Y., Yu J., Chen D., Wang J., Xie B., Chen J. Resonant pressure micro sensors based on dual double ended tuning fork resonators. *Micromachines*, 2019, vol. 10, no. 9, pp. 560. https://doi.org/10.3390/mi10090560
4. Andreev K.A., Tinyakov Yu.N., Shakhnov V.A. Mathematical models hybrid sensing elements transducers pressure. *Sensors & Systems*, 2013, no. 9 (172), pp. 2–9. (in Russian)
5. Tinyakov Yu.N., Nikolaeva A.S. Computation of pressure sensor membrane. Herald of the Bauman Moscow State Technical University. *Series Instrument Engineering*, 2015, no. 6 (105). pp. 135–142. (in Russian). https://doi.org/10.18698/0236-3933-2015-6-135-142
6. Clark S.K., Wise K.D. Pressure sensitivity in anisotropically etched thin-diaphragm pressure sensors. *Proc. of the IEEE Transactions on Electron Devices*, 1979, vol. 26, no. 12, pp. 1887–1896. https://doi.org/10.1109/T-ED.1979.19792
7. Yu Z., Zhao Y., Li L., Tian B., Li C. Geometry optimization for micropressure sensor considering dynamic interference. *Review of Scientific Instruments*, 2014, vol. 85, no. 9, pp. 095002. https://doi.org/10.1063/1.4895999
8. Gulieva D.A., Tsypin B.V., Kuchumov E.V. Increasing of the string primary converter sensitivity by changing the overall and mass characteristics of the sensitive element. *University Proceedings. Volga Region. Technical Sciences*, 2020, no. 3 (55), pp. 88–97. (in Russian). https://doi.org/10.21685/2072-3059-2020-3-9
9. Munas F.R., Amarasinghe Y.W.R., Kumarage P., Dao D.V., Dau V.T. Design and simulation of MEMS based piezoresitive pressure sensor for microfluidic applications. *Proc. of the Moratuwa Engineering Research Conference (MERCon)*, 2018, pp. 215–220. https://doi.org/10.1109/MERCon.2018.8421908
10. Tun P.V., Simonov B.M., Timoshenkov S.P. Investigation of the possibilities of increasing the sensitivity of a capacitive-type MEMS

форм // Известия высших учебных заведений. Электроника. 2023. Т. 28. № 2. С. 222–231. https://doi.org/10.24151/1561-5405-2023-28-2-222-231
11. Волков В.С., Французов М.В., Рыблова Е.А. Аналитическое и численное моделирование чувствительных элементов полупроводниковых датчиков давления // Измерение. Мониторинг. Управление. Контроль. 2016. № 2 (16). С. 110–117.
12. Цибизов П.Н. Чувствительные элементы для микроэлектронных датчиков давления информационно-измерительных систем: диссертация на соискание ученой степени кандидата технических наук. Пенза, 2007. 173 с.
13. Цыпин Б.В., Арискина Е.В., Щипанов В.Д., Ярославцева Д.А., Волков В.С., Баринов И.Н. Моделирование характеристик чувствительных элементов микромеханических датчиков давлений для эксплуатации в особо жестких условиях // Измерение. Мониторинг. Управление. Контроль. 2013. № 2 (4). С. 30–36.
14. Li Y., Lu Y., Xie B., Chen J., Wang J., Chen D. A micromachined resonant differential pressure sensor // IEEE Transactions on Electron Devices. 2020. V. 67. N 2. P. 640–645. https://doi.org/10.1109/TED.2019.2957880
15. Lu Y., Yan P., Xiang C., Chen D., Wang J., Xie B., Chen J. A resonant pressure microsensor with the measurement range of 1 MPa based on sensitivities balanced dual resonators // Sensors. 2019. V. 19. N 10. P. 2272. https://doi.org/10.3390/s19102272
16. Xiang C., Lu Y., Cheng C., Wang J., Chen D., Chen J. A resonant pressure microsensor with a wide pressure measurement range // Micromachines. 2021. V. 12. N. 4. P. 382. https://doi.org/10.3390/mi12040382
17. Yan P., Lu Y., Xiang C., Wang J., Chen D., Chen J. A temperature-insensitive resonant pressure micro sensor based on silicon-on-glass vacuum packaging // Sensors. 2019. V. 19. N 18. P. 3866. https://doi.org/10.3390/s19183866
18. Harada K., Ikeda K., Kuwayama H., Murayama H. Various applications of resonant pressure sensor chip based on 3-D micromachining // Sensors and Actuators A: Physical. 1999. V. 73. N 3. P. 261–266. https://doi.org/10.1016/S0924-4247(98)00245-3
19. Данилина Т.И., Смирнова К.И., Илюшин В.А., Величко А.А. Процессы Микро- и Нанотехнологии: учебное пособие. Томск: ТУСУР, 2004. 260 с.
20. Sundaram K.B., Vijayakumar A., Subramanian G. Smooth etching of silicon using TMAH and isopropyl alcohol for MEMS applications // Microelectronic Engineering. 2005. V. 77. N 3-4. P. 230–241. https://doi.org/10.1016/j.mee.2004.11.004
21. Mitsumori K. Silicon etching // Scientific Wet Process Technology for Innovative LSI/FPD Manufacturing. 2018. P. 252–263. https://doi.org/10.1201/9781315221076
22. Pal P., Swarnalatha V., Rao A.V.N., Pandey A.K., Tanaka H., Sato K. High speed silicon wet anisotropic etching for applications in bulk micromachining: a review // Micro and Nano Systems Letters. 2021. V. 9. N 1. P. 4. https://doi.org/10.1186/s40486-021-00129-0
23. Yu X., Ye Y., Zhu P., Wu L., Shen R., Zhu C. Wet anisotropic etching characteristics of Si{111} in KOH-based solution // ACS Omega. 2025. V. 10. N 3. P. 2940–2948. https://doi.org/10.1021/acsomega.4c09272

pressure sensor with membranes of various geometric shapes. Proceedings of Universities. *Electronics*, 2023, vol. 28, no. 2, pp. 222–231. (in Russian). https://doi.org/10.24151/1561-5405-2023-28-2-222-231
11. Volkov V.S., Frantsuzov M.V., Ryblova E.A. Analytical and numerical simulation of semiconductor piezoresistive pressure sensing elements. Measuring. *Monitoring. Management. Control*, 2016, no. 2 (16). pp. 110–117. (in Russian)
12. Tcibizov P.N. *Sensitive Elements for Microelectronic Pressure Sensors in Information-Measuring Systems*. Penza, PNZGU, 2007, 173 p. (in Russian)
13. Tcypin B.V., Ariskina E.V., Shchipanov V.D., Iaroslavtceva D.A., Volkov V.S., Barinov I.N. Simulation of sensing element characteristic for micromechanical pressure sensor for use for heavy duty. *Measuring. Monitoring. Management. Control*, 2013, no. 2 (4), pp. 30–36. (in Russian)
14. Li Y., Lu Y., Xie B., Chen J., Wang J., Chen D. A micromachined resonant differential pressure sensor. *IEEE Transactions on Electron Devices*, 2020, vol. 67, no. 2, pp. 640–645. https://doi.org/10.1109/TED.2019.2957880
15. Lu Y., Yan P., Xiang C., Chen D., Wang J., Xie B., Chen J. A resonant pressure microsensor with the measurement range of 1 MPa based on sensitivities balanced dual resonators. *Sensors*, 2019, vol. 19, no. 10, pp. 2272. https://doi.org/10.3390/s19102272
16. Xiang C., Lu Y., Cheng C., Wang J., Chen D., Chen J. A resonant pressure microsensor with a wide pressure measurement range. *Micromachines*, 2021, vol. 12, no. 4, pp. 382. https://doi.org/10.3390/mi12040382
17. Yan P., Lu Y., Xiang C., Wang J., Chen D., Chen J. A temperature-insensitive resonant pressure micro sensor based on silicon-on-glass vacuum packaging. *Sensors*, 2019, vol. 19, no. 18, pp. 3866. https://doi.org/10.3390/s19183866
18. Harada K., Ikeda K., Kuwayama H., Murayama H. Various applications of resonant pressure sensor chip based on 3-D micromachining. *Sensors and Actuators A: Physical*, 1999, vol. 73, no. 3, pp. 261–266. https://doi.org/10.1016/S0924-4247(98)00245-3
19. Danilina T.I., Smirnova K.I., Iliushin V.A., Velichko A.A. *Study Guide on Micro- and Nanotechnology Processes*. Tomsk, TUSUR, 2004, 260 p. (in Russian)
20. Sundaram K.B., Vijayakumar A., Subramanian G. Smooth etching of silicon using TMAH and isopropyl alcohol for MEMS applications. *Microelectronic Engineering*, 2005, vol. 77, no. 3-4, pp. 230–241. https://doi.org/10.1016/j.mee.2004.11.004
21. Mitsumori K. Silicon etching. *Scientific Wet Process Technology for Innovative LSI/FPD Manufacturing*, 2018, pp. 252–263. https://doi.org/10.1201/9781315221076
22. Pal P., Swarnalatha V., Rao A.V.N., Pandey A.K., Tanaka H., Sato K. High speed silicon wet anisotropic etching for applications in bulk micromachining: a review. *Micro and Nano Systems Letters*, 2021, vol. 9, no. 1, pp. 4. https://doi.org/10.1186/s40486-021-00129-0
23. Yu X., Ye Y., Zhu P., Wu L., Shen R., Zhu C. Wet anisotropic etching characteristics of Si{111} in KOH-based solution. *ACS Omega*, 2025, vol. 10, no. 3, pp. 2940–2948. https://doi.org/10.1021/acsomega.4c09272

Авторы

Вторушин Сергей Евгеньевич — научный сотрудник, Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация, sc 56669651800, https://orcid.org/0009-0008-9407-9552, sergei.e.vtorushin@tusur.ru

Таловская Алёна Алексеевна — младший научный сотрудник, Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация, https://orcid.org/0009-0001-6796-1135, alena.a.talovskaia@tusur.ru

Барбин Евгений Сергеевич — кандидат технических наук, заведующий лабораторией, Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация; научный сотрудник, Институт оптики атмосферы им. В.Е. Зуева Сибирского отделения Российской академии наук, Томск, 634055, Российская Федерация, sc 56039318500, https://orcid.org/0000-0001-5904-0216, bar_es@mail.ru

Authors

Sergey E. Vtorushin — Scientific Researcher, Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation, sc 56669651800, https://orcid.org/0009-0008-9407-9552, sergei.e.vtorushin@tusur.ru

Alena A. Talovskaia — Junior Researcher, Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation, https://orcid.org/0009-0001-6796-1135, alena.a.talovskaia@tusur.ru

Evgenii S. Barbin — PhD, Head of Laboratory, Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation; Scientific Researcher, V.E. Zuev Institute of Atmospheric Optics, sc 56039318500, https://orcid.org/0000-0001-5904-0216, bar_es@mail.ru

Кулинич Иван Владимирович — кандидат технических наук, доцент, старший научный сотрудник, Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация, sc 57205437118, https://orcid.org/0009-0001-0768-3495, kulinich@tusur.ru

Вайсбеккер Мария Сергеевна — доцент, научный сотрудник, Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация, https://orcid.org/0009-0002-1078-4649, mariya_vaisbekker@mail.ru

Ivan V. Kulinich — PhD, Associate Professor, Senior Researcher, Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation, sc 57205437118, https://orcid.org/0009-0001-0768-3495, kulinich@tusur.ru

Mariya S. Vaisbekker — Associate Professor, Scientific Researcher, Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation, https://orcid.org/0009-0002-1078-4649, mariya_vaisbekker@mail.ru

Статья поступила в редакцию 03.07.2025
Одобрена после рецензирования 08.10.2025
Принята к печати 23.11.2025

Received 03.07.2025
Approved after reviewing 08.10.2025
Accepted 23.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2025 Том 25 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2025 Vol. 25 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

АВТОМАТИЧЕСКОЕ УПРАВЛЕНИЕ И РОБОТОТЕХНИКА

AUTOMATIC CONTROL AND ROBOTICS

doi: 10.17586/2226-1494-2025-25-6-1080-1088
УДК 618.51

Адаптивная компенсация неизвестных возмущений в многоканальных линейных системах с различной величиной задержки в каналах управления

Конг Винь Ты[1]✉, Наталия Александровна Дударенко[2]

[1,2] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] congvinhvkdn@gmail.com✉, https://orcid.org/0009-0005-5106-8560
[2] dudarenko@itmo.ru, https://orcid.org/0000-0002-3553-0584

Аннотация
Введение. Рассмотрена задача компенсации неизвестных внешних возмущений для класса линейных стационарных многоканальных систем с различной величиной запаздывания в каналах управления. Предполагается, что внешние возмущения представляют собой гармонические сигналы с неизвестными частотами, фазами, амплитудами и смещениями, которые одновременно воздействуют как на вход, так и на выход системы. Для решения поставленной задачи используется метод прямой компенсации возмущений, основанный на принципе внутренней модели, в сочетании с классическим методом развязки по состоянию Фальба–Воловича, что позволяет повысить скорость сходимости выходных сигналов даже при небольшом коэффициенте адаптации. **Метод.** С целью устранения перекрестных взаимодействий между управляющими контурами к системе применяется метод развязки каналов на основе обратной связи по состоянию, разработанный на базе классического подхода Фальба–Воловича. Затем строится наблюдатель для оценки вектора состояния модели внешнего возмущения и на основе полученных оценок синтезируется адаптивный закон управления с расширением памяти регрессора для компенсации внешних возмущений на основе принципа внутренней модели. Стабилизация системы осуществляется одновременно с развязкой каналов управления, что позволяет перейти к задаче компенсации внешних неизвестных возмущений, минуя фазу проектирования стабилизирующего компонента управляющего сигнала. Предлагаемый метод не предполагает ограничений на наблюдаемость и устойчивость объекта управления. **Основные результаты.** Для компенсации неизвестных внешних возмущений для класса линейных стационарных многоканальных систем с различными запаздываниями в каналах управления предложен адаптивный алгоритм с расширением памяти регрессора, сочетающийся с методом развязки каналов управления по состоянию Фальба–Воловича. Показана эффективность предлагаемого подхода на примере численного моделирования в программном пакете MATLAB/Simulink. Полученные графики переходных процессов выявили ограниченность всех сигналов замкнутой системы и асимптотическую устойчивость выходных переменных при наличии различных запаздываний в каналах управления многоканальной системы, функционирующей в условиях неизвестных внешних возмущений. **Обсуждение.** Предлагаемый подход позволяет получить улучшенную скорость сходимости процессов и может быть применен в технических системах и комплексах, математическое описание которых задается в виде линейной многоканальной системы с различной величиной запаздывания в каналах управления.

Ключевые слова
адаптивное управление, гармонические возмущения, компенсация возмущений, развязка каналов, запаздывания в каналах управления, метод Фальба–Воловича, многоканальная система

Ссылка для цитирования: Ты К.В., Дударенко Н.А. Адаптивная компенсация неизвестных возмущений в многоканальных линейных системах с различной величиной задержки в каналах управления // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1080–1088 doi: 10.17586/2226-1494-2025-25-6-1080-1088

© Ты К.В., Дударенко Н.А., 2025

Adaptive compensation of unknown disturbances in MIMO linear systems with distinct input delays

Cong Vinh Tu[1]✉, Natalia A. Dudarenko[2]

[1,2] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] congvinhvkdn@gmail.com✉, https://orcid.org/0009-0005-5106-8560
[2] dudarenko@itmo.ru, https://orcid.org/0000-0002-3553-0584

Abstract
The paper considers the problem of compensation for unknown external disturbances for a class of linear stationary multidimensional systems with distinct input delays. It is assumed that external disturbances are harmonic signals with unknown frequencies, phases, amplitudes, and biases that simultaneously affect both the input and output of the system. To solve the problem, the direct disturbance compensation method based on the internal model principle is used in combination with the classical Falb-Wolovich linear state feedback decoupling method which allows increasing the convergence rate of output signals with a small adaptation parameter. In order to eliminate cross-interactions between control loops, the channel decoupling method based on Falb-Wolovich linear state feedback decoupling approach is applied to the system. Then, an observer is constructed to estimate the state vector of the external disturbance model and, based on the estimations, an adaptive control law with the memory regressor extension is designed to compensate for external disturbances based on the internal model principle. The system is stabilized simultaneously with the decoupling of control channels, which allows one to proceed to the problem of compensating for external unknown disturbances, bypassing the design phase of the stabilizing component of the control signal. There are no restrictions on the observability and stability of the control plant. An adaptive algorithm with the memory regressor extension combined with the Falb-Volovich linear state feedback decoupling method is proposed to compensate for unknown external disturbances for a class of linear stationary multidimensional systems with distinct control delays. The efficiency of the proposed approach is illustrated by an example of numerical simulation in the MATLAB/Simulink environment. The resulting transient response plots demonstrate that the proposed algorithm ensures the boundedness of all closed-loop signals and the asymptotic stability of the output variables in the presence of distinct input delays under external harmonic disturbances. The proposed approach allows obtaining an improved rate of convergence of processes and can be applied in engineering systems and complexes the mathematical description of which is given in the form of linear multidimensional systems with distinct input delays.

Keywords
adaptive control, harmonic disturbance, disturbances compensation, decoupling, control delays, Falb-Wolovich method, MIMO system

For citation: Tu C.V., Dudarenko N.A. Adaptive compensation of unknown disturbances in MIMO linear systems with distinct input delays. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1080–1088 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1080-1088

Введение

Системы управления промышленными процессами, как правило, являются многоканальными системами управления. Наличие внутренних взаимосвязей между каналами управления создает значительные трудности при проектировании таких систем. Распространенной практической проблемой является также наличие временных запаздываний различной величины в каналах управления многоканальной системы, что может быть вызвано, например, задержками при отработке команды исполнительных механизмов или задержками при передаче данных с датчиков. Эти взаимодействия и задержки создают серьезные сложности при разработке систем управления. Для решения проблемы часто применяют методы развязки каналов управления [1], позволяющие свести исходную модель системы к модельному представлению с сепаратными каналами управления. При эффективной развязке процессов можно напрямую применять стратегии управления, разработанные для одноканальных систем с одним входом и одним выходом.

Методы развязки [1] обычно классифицируются по двум категориям: статические и динамические. Методы статической развязки просты при проектировании, но эффективны только в установившемся режиме. Напротив, методы динамической развязки, такие как идеальная, упрощенная и обратная развязки, способны скомпенсировать взаимодействия как в установившемся, так и в переходном режиме работы системы. Однако они имеют ряд ограничений, более сложны при реализации и чувствительны к погрешностям модели. Для преодоления этих ограничений были предложены подходы к развязке на основе обратной связи по состоянию [1–8]. В частности, динамический метод обратной связи по состоянию, основанный на подходе Фальба–Воловича [4], который является простым в реализации и позволяет обеспечить эффективную работу системы даже при наличии неопределенностей модели и внешних возмущений [9]. Данный подход применен в настоящей работе для решения задачи компенсации неизвестных внешних возмущений в многоканальных системах.

Один из распространенных подходов к компенсации внешних возмущений основывается на использовании метода внутренней модели [10–12], при котором внешние воздействия моделируются как выходные сигналы автономного линейного генератора. В работах [13–15] задача решается в два этапа. На первом этапе формируются оценки параметров внешних возмущений, а на втором — решается задача синтеза закона управления на основе полученных оценок, что является существенным достоинством данного подхода и позволяет приме-

нять различные алгоритмы управления и компенсации возмущений [16, 17]. Тем не менее, существенным ограничением остается требование к соблюдению условий устойчивого возбуждения регрессора [18].

В работах [19–22] рассмотрен подход к прямой компенсации возмущений в линейных системах с запаздыванием на входе. Этот метод для ослабления условия о постоянном возбуждении регрессора базируется на построении наблюдателя специальной формы, оценивающего неизвестные возмущения на основе состояния системы или ее выходного сигнала. На основе этих оценок формируется закон прямого управления, обеспечивающий желаемое поведение замкнутой системы. Расширение данного подхода на случай неустойчивых многоканальных систем с различной величиной запаздывания в каналах управления представлено в работе [23].

В настоящей работе предлагается новый метод прямой компенсации неизвестных внешних возмущений для неустойчивых многоканальных систем с различными запаздываниями в каналах управления. Сначала, с целью устранения перекрестных взаимодействий между управляющими контурами, к системе применяется метод развязки каналов на основе обратной связи по состоянию, разработанный на базе классического подхода Фальба–Воловича. Затем строится наблюдатель для оценки вектора состояния модели внешнего возмущения и на основе полученных оценок синтезируется закон управления для компенсации внешних возмущений. Стабилизация системы осуществляется одновременно с развязкой каналов, что позволяет сразу перейти к задаче компенсации возмущений, минуя фазу проектирования стабилизирующего компонента управляющего сигнала [22, 23]. Предлагаемый метод не предполагает ограничений на наблюдаемость и устойчивость объекта управления.

Постановка задачи

Рассмотрим следующую линейную систему:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}_\tau + \mathbf{D}\mathbf{v}(t), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{R}\mathbf{v}(t), \end{cases} \tag{1}$$

где $\mathbf{x}(t) \in \mathbb{R}^n$, $\mathbf{y}(t) \in \mathbb{R}^m$, $\mathbf{v}(t) \in \mathbb{R}^k$ и $\mathbf{u}_\tau \in \mathbb{R}^m$ — вектор измеряемых переменных состояния, вектор измеряемых переменных выхода, вектор неизмеряемых ограниченных внешних возмущений и вектор управляющих (входных) сигналов, характеризующихся величиной задержки соответственно. Вектор управляющих сигналов задается в виде $\mathbf{u}_\tau = [u_1(t-\tau_1)u_2(t-\tau_2)\ \ldots\ u_m(t-\tau_m)]^T$, где $\tau_i \in \mathbb{R}^+$, $u_i \equiv 0$ при $t < \tau_i$, $i = 1, 2, \ldots, m$. Матрицы $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{m\times n}$, $\mathbf{D} \in \mathbb{R}^{n\times k}$ считаются известными. Матрица $\mathbf{A}$ может быть неустойчивой, матрица $\mathbf{R} \in \mathbb{R}^{m\times k}$ может быть неизвестной, а n, m, k обозначают размерности соответствующих векторов, при этом $m \le n$. Без ограничения общности будем считать, что $\tau_1 \le \tau_2 \le \ldots \le \tau_m$.

Рассматриваемый объект управления и вектор возмущающих воздействия рассматриваются с учетом следующих допущений.

Допущение 1. Пара матриц ($\mathbf{A}$, $\mathbf{B}$) управляема.

Допущение 2. Вектор внешнего возмущения $\mathbf{v}(t)$ имеет следующий вид:

$$\mathbf{v}(t) = [\mathrm{v}_1(t), \mathrm{v}_2(t), \ldots, \mathrm{v}_k(t)]^T,$$

$$\mathrm{v}_\alpha(t) = \sum_{l=0}^{r} R_l \sin(\omega_l t + \varphi_l) + R_{0l},$$

где $\mathrm{v}_\alpha(t)$ при $\alpha = 1, \ldots, k$ представляет собой соответствующие компоненты векторного внешнего возмущения. R_l, ω_l, φ_l и R_{0l} являются неизвестными амплитудами, частотами, фазами и смещениями внешнего возмущения соответственно.

Примечание. Знание точного количества гармонических составляющих внешнего возмущения не требуется. В предлагаемом подходе используется только информация о верхней границе количества гармоник r, которая априори известна. Алгоритм остается эффективным и обеспечивает достижение цели управления в случае, если фактическое количество гармоник меньше или равно значению r [22]. В текущей работе используется каноническая форма модельного описания возмущения [16], что позволяет решать задачу компенсации внешних воздействий, в том числе и для случая гармонических составляющих одинаковой частоты.

Задача управления заключается в разработке закона управления с обратной связью по состоянию, обеспечивающего ограниченность всех сигналов в замкнутой системе и достижение цели управления

$$\lim_{t\to\infty} \|\mathbf{y}(t)\| = 0\,, \tag{2}$$

где $\|*\|$ — евклидова норма вектора «*».

Для реализации предлагаемого метода на первом этапе проектирования закона управления осуществляется развязка каналов системы на основе линейной обратной связи по состоянию в соответствии с методом Фальба–Воловича. Далее конструируется наблюдатель для оценки вектора состояния внешнего возмущения. Полученные оценки используются для формирования закона управления, обеспечивающего компенсацию внешних возмущений в многомерной системе с различной величиной запаздывания в каналах управления.

Синтез развязки каналов многомерной системы методом Фальба–Воловича

Рассмотрим метод линейной обратной связи по состоянию, основанный на классическом подходе Фальба–Воловича, обеспечивающий одновременно развязку передаточных функций каналов линейных многоканальных систем, заданных уравнением вида (1), и стабилизацию процессов с желаемыми показателями качества.

Предположим, что $\mathbf{G}(s)$ является матрицей передаточных функций исходной системы размером $m \times m$, а $\mathbf{W}(s)$ представляет собой желаемую матрицу передаточных функций для развязанной системы. Передаточные матрицы $\mathbf{G}(s)$ и $\mathbf{W}(s)$ зададим в следующей форме

$$\mathbf{G}(s) = \begin{bmatrix} g_{11}(s) & g_{12}(s) & \cdots & g_{1m}(s) \\ g_{21}(s) & g_{22}(s) & \cdots & g_{2m}(s) \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1}(s) & g_{m2}(s) & \cdots & g_{mm}(s) \end{bmatrix}, \tag{3}$$

$$\mathbf{W}(s) = \begin{bmatrix} W_{11}(s) & 0 & \cdots & 0 \\ 0 & W_{22}(s) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W_{mm}(s) \end{bmatrix}. \quad (4)$$

Рассмотрим представление системы (1) в пространстве состояний без учета внешних возмущений в следующем виде:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}_d(t), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t). \end{cases} \quad (5)$$

Для получения желаемой матрицы передаточных функций $\mathbf{W}(s)$ развязанной системы применим следующий закон управления, первоначально предложенный в работе [4]:

$$\mathbf{u}_d(t) = -\mathbf{K}\mathbf{x}(t) + \mathbf{F}\mathbf{u}_\tau, \quad (6)$$

где $\mathbf{K}$ и $\mathbf{F}$ — вещественные постоянные матрицы соответствующих размерностей.

Представим матрицу $\mathbf{C}$ в виде:

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}^1 \\ \mathbf{c}^2 \\ \vdots \\ \mathbf{c}^m \end{bmatrix},$$

где $\mathbf{c}^i$ — i-ая строка матрицы $\mathbf{C}$.

Введем в рассмотрение целые числа σ_i при $i = 1, \ldots, m$, такие что

$$\sigma_i = \begin{cases} \min(h\colon \mathbf{c}^i\mathbf{A}^{h-1}\mathbf{B} \neq \mathbf{0}^T,\ h = 1, 2, \ldots, n-1), \\ n-1, \text{ если } \mathbf{c}^i\mathbf{A}^{h-1}\mathbf{B} = \mathbf{0}^T,\ h = 1, 2, \ldots, n, \end{cases} \quad (7)$$

где $\mathbf{0}^T \in \mathbb{R}^m$ — строковый вектор, все элементы которого равны нулю, а также матрицы $\mathbf{B}^*$ и $\mathbf{C}^*$ вида

$$\mathbf{B}^* = \begin{bmatrix} \mathbf{c}^1\mathbf{A}^{\sigma_1-1}\mathbf{B} \\ \mathbf{c}^2\mathbf{A}^{\sigma_2-1}\mathbf{B} \\ \vdots \\ \mathbf{c}^m\mathbf{A}^{\sigma_m-1}\mathbf{B} \end{bmatrix}, \ \mathbf{C}^* = \begin{bmatrix} \mathbf{c}^1\mathbf{A}^{\sigma_1} \\ \mathbf{c}^2\mathbf{A}^{\sigma_2} \\ \vdots \\ \mathbf{c}^m\mathbf{A}^{\sigma_m} \end{bmatrix}.$$

Для дальнейших математических выкладок оказываются справедливыми положения следующей теоремы [1].

Теорема 1. Закон управления вида (6) существует для развязки системы (5) тогда и только тогда, когда матрица $\mathbf{B}^*$ невырождена. Если данное условие выполнено, то при выборе

$$\mathbf{K} = (\mathbf{B}^*)^{-1}\mathbf{C}^*,\ \mathbf{F} = (\mathbf{B}^*)^{-1}$$

получается замкнутая система с передаточной матрицей в диагональной форме, представленной следующим образом:

$$\mathbf{W}(\mathrm{s}) = \mathrm{diag}\{s^{-\sigma_1}, s^{-\sigma_2}, \ldots, s^{-\sigma_m}\}. \quad (8)$$

Из выражения (8) следует, что все нули передаточной функции, находящиеся в правой полуплоскости, которые имеются в исходной системе с передаточной функцией вида (3) изначально (если они присутствуют), исключаются, а закон развязки преобразует систему в набор передаточных функций с независимыми интеграторами.

Однако, чтобы обеспечить наличие у развязанной системы заранее заданных желаемых полюсов, матрицы $\mathbf{K}$ и $\mathbf{F}$, в соответствии с методом Фальба–Воловича [4], подбираются специальным образом. Обозначим $\bar{n} = \sum_{i=1}^{m} \sigma_i$ общее количество полюсов, подлежащих назначению для замкнутой системы с обратной связью. Эти полюса делятся на m подмножеств λ_{ip}, где $p = 1, 2, \ldots, \sigma_i$ для каждого $i = 1, 2, \ldots, m$. Для каждого i определим характеристический полином $\varphi_i(s)$ следующим образом:

$$\varphi_i(s) = \sum_{p=1}^{\sigma_i} (s - \lambda_{ip}),\ i = 1, 2, \ldots, m \quad (9)$$

или в виде:

$$\varphi_i(s) = q_{i0} + q_{i1}s + q_{i2}s^2 + \ldots + s^{\sigma_i}.$$

Согласно разработанному методу, управляемая система должна иметь диагональную структуру, как показано в выражении (4), при этом ее элементы определяются передаточными функциями вида:

$$W_{ii}(s) = \frac{q_{i0}}{q_{i0} + q_{i1}s + q_{i2}s^2 + \ldots + s^{\sigma_i}}.$$

Обозначив

$$\mathbf{C}^{**} = \begin{bmatrix} \mathbf{c}^1\mathbf{A}^{\sigma_1} + \sum_{j=0}^{\sigma_1-1} q_{1j}\mathbf{c}^1\mathbf{A}^j \\ \mathbf{c}^m\mathbf{A}^{\sigma_m} + \sum_{j=0}^{\sigma_m-1} q_{mj}\mathbf{c}^m\mathbf{A}^j \end{bmatrix},$$

перейдем к формулировке следующей теоремы [1].

Теорема 2. Если система (5) допускает развязку, то существует закон управления вида (6), где

$$\mathbf{K} = (\mathbf{B}^*)^{-1}\mathbf{C}^{**} \text{ и } \mathbf{F} = (\mathbf{B}^*)^{-1} \begin{bmatrix} q_{10} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & q_{m0} \end{bmatrix}.$$

При использовании данного закона управления передаточные нули автоматически исключаются, что позволяет разработчику свободно выбирать собственные значения замкнутой системы на основе выражения (9).

Доказательство приведено в работе [1].

Замечание 1. Система (1), после развязки с использованием закона управления, сформированного на основе матриц $\mathbf{K}$ и $\mathbf{F}$, как указано в теореме 2, эквивалентна m независимым одноканальным системам и может быть представлена в следующей форме в пространстве состояний:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}_d\mathbf{x}(t) + \mathbf{B}_d\mathbf{u}_\tau + \mathbf{D}\mathbf{v}(t), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{R}\mathbf{v}(t), \end{cases} \quad (10)$$

где $\mathbf{A}_d = \mathbf{A} - \mathbf{B}\mathbf{K}$ и $\mathbf{B}_d = \mathbf{B}\mathbf{F}$ — матрицы системы, полученной после развязки, причем $\mathbf{A}_d$ является матрицей Гурвица.

Синтез наблюдателя внешнего возмущения

Внешнее возмущение представим в виде выхода следующего автономного линейного генератора [16]:

$$\dot{\xi}(t) = \mathbf{G}\xi(t) + \mathbf{L}\mathbf{v}(t),\ \mathbf{v}(t) = \boldsymbol{\theta}^T\xi(t) \tag{11}$$

или

$$\dot{\xi}(t) = (\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)\xi(t),\ \mathbf{v}(t) = \boldsymbol{\theta}^T\xi(t), \tag{12}$$

где $\xi \in \mathbb{R}^q$ — вектор состояния внешнего возмущения; $\mathbf{G} = \mathrm{blkdiag}\{\mathbf{G}_\alpha\}$, $\mathbf{G}_\alpha \in \mathbb{R}^{q_\alpha \times q_\alpha}$ — матрицы Гурвица, где $\mathrm{blkdiag}\{*\}$ — блочная диагональная матрица с диагональными элементами «*»; $\mathbf{L} = \mathrm{blkdiag}\{\mathbf{L}_\alpha\}$, $\mathbf{L}_\alpha \in \mathbb{R}^{q_\alpha}$ — постоянные векторы; $\boldsymbol{\theta}^T \in \mathbb{R}^{k \times q}$ — неизвестная матрица. Пары матриц $(\mathbf{G}_\alpha, \mathbf{L}_\alpha)$ выбираются произвольно таким образом, чтобы каждая пара была полностью управляема, $\alpha = 1, 2, \ldots, k$ и $q = \sum_{\alpha=1}^{k} q_\alpha$.

Замечание 2. Применим фундаментальное решение уравнения (12) для прогнозирования значения $\xi(t)$, которое может быть представлено следующим образом:

$$\xi(t) = \boldsymbol{\Phi}_i\xi(t - \tau_i),$$

где $\boldsymbol{\Phi}_i = e^{(\mathbf{G}+\mathbf{L}\boldsymbol{\theta}^T)\tau_i} \in \mathbb{R}^{q \times q}$.

На основе (11) разработаем наблюдатель возмущений [23], используемый для оценки состояния.

Лемма 1. Построим наблюдатель состояния внешнего возмущения для развязанной системы (10) вида

$$\begin{cases} \dot{\boldsymbol{\varphi}}(t) = \mathbf{G}\boldsymbol{\varphi}(t) + (\mathbf{G}\mathbf{N} + \mathbf{N}\mathbf{A}_d)\mathbf{x}(t) - \mathbf{N}\mathbf{B}_d\mathbf{u}_\tau, \\ \hat{\xi}(t) = \boldsymbol{\varphi} + \mathbf{N}\mathbf{x}(t), \end{cases} \tag{13}$$

где $\hat{\xi}(t) \in \mathbb{R}^q$ — оценка вектора состояния $\xi(t)$; $\boldsymbol{\varphi}(t) \in \mathbb{R}^q$ — вспомогательный вектор наблюдателя, а для матрицы $\mathbf{N} = [\mathbf{L}_{01}, \mathbf{L}_{02}, \ldots, \mathbf{L}_{0k}]^T \in \mathbb{R}^{q \times n}$ выполняется равенство:

$$\mathbf{N}_\alpha\mathbf{D} = \mathbf{L}_{0\alpha},$$

где $\mathbf{N}_\alpha \in \mathbb{R}^{q_\alpha \times n}$, а матрица $\mathbf{L}_{0\alpha}$ такая что

$$\mathbf{L}_{0\alpha} = [\mathbf{0}_{q_\alpha}, \ldots, \mathbf{0}_{q_\alpha}, \mathbf{L}_\alpha, \mathbf{0}_{q_\alpha}, \ldots, \mathbf{0}_{q_\alpha}],$$

где вектор $\mathbf{L}_\alpha$ α-й столбец; $\mathbf{0}_{q_\alpha}$ — нулевой вектор размерности q_α. Тогда $\xi(t) = \hat{\xi}(t) + \boldsymbol{\epsilon}$, где $\boldsymbol{\epsilon}$ экспоненциально затухает.

Доказательство приведено в работе [16].

Синтез закона управления и алгоритма адаптации

Чтобы преобразовать модель системы (10) в такую форму, при которой внешнее возмущение согласовано с управляющим воздействием $\mathbf{u}_\tau$, используем матричное уравнение регулятора [16, 17].

Лемма 2. При выполнении допущений 1 и 2, для системы (10) существует пара решений $\mathbf{M} \in \mathbb{R}^{n \times q}$ и $\boldsymbol{\psi} \in \mathbb{R}^{q \times m}$, удовлетворяющая следующим матричным уравнениям:

$$\mathbf{B}_d\boldsymbol{\psi}^T = \mathbf{A}_d\mathbf{M} - \mathbf{M}(\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T) + \mathbf{D}\boldsymbol{\theta}^T,\ \mathbf{C}\mathbf{M} = -\mathbf{R}\boldsymbol{\theta}^T.$$

Подставим $\mathbf{v}(t) = \boldsymbol{\theta}^T\xi(t)$ в (10), вводя ошибку состояния $\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{M}\xi(t)$, и, вычисляя ее производную с учетом (10) и (12).

В результате получим:

$$\dot{\mathbf{e}}(t) = \mathbf{A}_d\mathbf{e}(t) + \mathbf{B}_d\mathbf{u}_\tau + (\mathbf{A}_d\mathbf{M} - \mathbf{M}(\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T) + \mathbf{D}\boldsymbol{\theta}^T)\xi(t).$$

Доказательство приведено в работе [16].

Учитывая матричные уравнения из леммы 2, перепишем систему в виде:

$$\begin{cases} \dot{\mathbf{e}}(t) = \mathbf{A}_d\mathbf{e}(t) + \mathbf{B}_d(\boldsymbol{\psi}^T\xi(t) + \mathbf{u}_\tau), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{e}(t), \end{cases} \tag{14}$$

где матрица $\boldsymbol{\psi}^T$ неизвестна.

Поскольку матрица передаточной функции $\mathbf{W}(s)$ диагональная, управляющий сигнал $\mathbf{u}_\tau$ и состояние ξ экзосистемы имеют одинаковые задержки. Используя замечание 2, запишем:

$$\boldsymbol{\psi}^T\xi(t) = \begin{bmatrix} \boldsymbol{\psi}_1^T\xi(t) \\ \vdots \\ \boldsymbol{\psi}_m^T\xi(t) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\psi}_1^T\boldsymbol{\Phi}_1\xi(t - \tau_1) \\ \vdots \\ \boldsymbol{\psi}_m^T\boldsymbol{\Phi}_m\xi(t - \tau_m) \end{bmatrix} = \boldsymbol{\eta}^T\xi(t - \tau_i),$$

где $\boldsymbol{\eta}^T = [\boldsymbol{\psi}_1^T\boldsymbol{\Phi}_1, \boldsymbol{\psi}_2^T\boldsymbol{\Phi}_2, \ldots, \boldsymbol{\psi}_m^T\boldsymbol{\Phi}_m]^T$ — новая матрица неизвестных параметров, при этом $\boldsymbol{\psi}_i^T$ представляет собой i-ю строку матрицы $\boldsymbol{\psi}^T$, а τ_i — значение задержки соответствующего канала управления.

Основываясь на лемме 1, неизмеримое состояние экзосистемы ξ заменим его оценкой $\hat{\xi}$, в то время как экспоненциально убывающий компонент опустим, поскольку он не влияет на устойчивость замкнутой системы.

После замены переменных в выражении (14) получим

$$\begin{cases} \dot{\mathbf{e}}(t) = \mathbf{A}_d\mathbf{e}(t) + \mathbf{B}_d(\boldsymbol{\eta}^T\hat{\xi}(t - \tau_i) + \mathbf{u}_\tau), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{e}(t). \end{cases} \tag{15}$$

Поскольку матрица $\mathbf{A}_d$ является гурвицевой, можно синтезировать следующий закон управления для системы (15) с различными задержками на входе:

$$\mathbf{u} = -\hat{\boldsymbol{\eta}}^T\hat{\xi}(t), \tag{16}$$

где $\hat{\boldsymbol{\eta}}^T \in \mathbb{R}^{m \times q}$ — матрица коэффициентов, определяемая адаптивным алгоритмом.

Подставим (16) в (15):

$$\begin{cases} \dot{\mathbf{e}}(t) = \mathbf{A}_d\mathbf{e}(t) + \mathbf{B}_d\tilde{\boldsymbol{\eta}}_\tau^T\hat{\xi}(t - \tau_i), \\ \mathbf{y}(t) = \mathbf{C}\mathbf{e}(t), \end{cases} \tag{17}$$

где $\tilde{\boldsymbol{\eta}}^T = \boldsymbol{\eta}^T - \hat{\boldsymbol{\eta}}_\tau^T$ — матрица параметрической ошибки с учетом множественных задержек; $\hat{\boldsymbol{\eta}}_\tau^T = [\hat{\boldsymbol{\eta}}_1^T(t - \tau_1), \hat{\boldsymbol{\eta}}_1^T(t - \tau_2), \ldots \hat{\boldsymbol{\eta}}_1^T(t - \tau_m)]^T$ — матрица настраиваемых параметров с задержкой.

Из (17) следует, что отрицательное влияние запаздывания не было полностью устранено. Для решения данной проблемы построим расширенный вектор ошибки [24]

$$\hat{\mathbf{e}} = \mathbf{e} + \boldsymbol{\chi}, \tag{18}$$

где сигнал $\boldsymbol{\chi}$ определяется как

$$\dot{\chi} = \mathbf{A}_d\chi + \mathbf{B}_d(\hat{\eta}_\tau^T - \hat{\eta}^T)\hat{\xi}(t - \tau_i). \quad (19)$$

Продифференцируем выражение (18) с учетом (17) и (19):

$$\begin{cases} \dot{\hat{\mathbf{e}}}(t) = \mathbf{A}_d\hat{\mathbf{e}} + \mathbf{B}_d\tilde{\eta}^T\hat{\xi}(t - \tau_i), \\ \varepsilon(t) = \mathbf{C}\hat{\mathbf{e}}. \end{cases} \quad (20)$$

Перепишем (20) в виде:

$$\varepsilon = \mathbf{W}(s)[\tilde{\eta}^T\hat{\xi}(t - \tau_i)]. \quad (21)$$

Для применения алгоритма адаптации на основании (21), расширенная ошибка может быть определена в соответствии со схемой подстановки ([16, стр. 337, лемма D.4]) следующим образом:

$$\bar{\varepsilon} = \varepsilon + \mathbf{W}(s)[\hat{\eta}^T\hat{\xi}(t - \tau_i)] - \mathbf{W}(s)[\hat{\xi}_\tau^T]\hat{\eta}. \quad (22)$$

Тогда справедливо равенство:

$$\bar{\varepsilon} = \mathbf{W}(s)[\hat{\xi}_\tau^T]\hat{\eta}, \quad (23)$$

где

$$\mathbf{W}(s)[\hat{\xi}_\tau^T] = \\ = \begin{bmatrix} \mathbf{W}_{11}(s)[(\hat{\xi}^T(t - \tau_1))] & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{W}_{mm}(s)[(\hat{\xi}^T(t - \tau_m))] \end{bmatrix}.$$

Модель ошибки (23) хорошо известна в теории адаптивного управления и позволяет использовать градиентный алгоритм адаптации. Однако данный алгоритм адаптации может демонстрировать очень низкую скорость сходимости в некоторых случаях [18]. Чтобы улучшить сходимость всех сигналов в системе, построим алгоритм адаптации с расширенным регрессором, как предложено в [16, 22]. Основываясь на (23), рассмотрим новую (расширенную) переменную

$$\hat{\varepsilon} = \bar{\varepsilon} + \mathbf{W}(s)[\hat{\xi}_\tau^T]\hat{\eta}. \quad (24)$$

С учетом (15), (21) и (22), уравнение (24) перепишем в следующем виде:

$$\hat{\varepsilon} = \mathbf{y} - \mathbf{W}(s)[\mathbf{u}_\tau]$$

или

$$\hat{\varepsilon} = \mathbf{W}(s)[\hat{\xi}_\tau^T]\eta. \quad (25)$$

Обозначим $\Delta = \mathbf{W}(s)[\hat{\xi}_\tau^T]$. Умножая обе части уравнения (25) на Δ, получим

$$\Delta\hat{\varepsilon} = \Delta\Delta^T\eta. \quad (26)$$

Фильтруя обе части выражения (26) с помощью апериодического звена первого порядка $\mathbf{H}(s) = \dfrac{1}{\beta s + 1}$, где коэффициент $\beta > 0$, запишем

$$\underbrace{\mathbf{H}(s)[\Delta\hat{\varepsilon}]}_{\mathbf{Y}} = \underbrace{\mathbf{H}(s)[\Delta\Delta^T]}_{\Omega}\eta,$$

$$\mathbf{Y} = \Omega\eta. \quad (27)$$

На основе уравнения (27) алгоритм адаптации примет вид

$$\dot{\hat{\eta}} = \mu(\mathbf{Y} - \Omega\hat{\eta}), \quad (28)$$

где μ — коэффициент адаптации.

С алгоритмом адаптации (28) замкнутая система удовлетворяет следующим свойствам [22].

Лемма 3. При выполнении допущений 1 и 2, алгоритм адаптации (28) совместно с наблюдателем возмущений (13), законом управления (16) и схемой замены (22), при применении к объекту управления (10), приводит к следующим результатам:

— ограниченность $\|\bar{\varepsilon}\|$, $\|\hat{\eta}\|$;
— асимптотическая сходимость $\|\tilde{\eta}^T\hat{\xi}\| \to 0$, $\|\hat{\eta}_\tau^T\hat{\xi}_\tau^T\| \to 0$ при $t \to \infty$;
— если собственное значение $\lambda_K(t) \notin \mathcal{L}_1$ матрицы Ω минимальное, тогда $\|\tilde{\eta}\| \to 0$ при $t \to \infty$, где $\mathcal{L}_1$ — пространство абсолютно интегрируемых векторных функций;
— если $\Delta \in$ PE, тогда $\|\tilde{\eta}\|$ стремится к нулю экспоненциально, и эта сходимость может быть ускорена за счет увеличения параметра μ, где $\Delta \in$ PE — векторная функция $\Delta(t)$, которая удовлетворяет условию постоянного возбуждения (PE) [18].

Следовательно, можно сделать вывод, что предложенный метод позволяет достичь цели управления (2).

Теорема 3. При выполнении допущений 1 и 2, закон управления (16) совместно с внешним наблюдателем возмущений (13), схемой замены (22) и алгоритмом адаптации (28), при применении к системе (10), которая получена путем декомпозиции системы (1) с использованием закона управления (6) и матриц, определенных в теореме 2, обеспечивает ограниченность всех сигналов в замкнутой системе и достижение цели управления (2).

Численный пример

Рассмотрим систему вида (1) с задержками в двух сигналах управлений $\tau_1 = 5$ с и $\tau_2 = 3$ с. Зададим матрицы системы:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 \\ 0{,}1 & -0{,}9 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \mathbf{D} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, \mathbf{R} = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}.$$

Пусть внешние возмущения заданы в виде:

$$\mathbf{v}(t) = \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} 5\sin(2t) \\ 4 + 7\sin(3t) \end{bmatrix}.$$

Выберем следующие матрицы для внешнего возмущения:

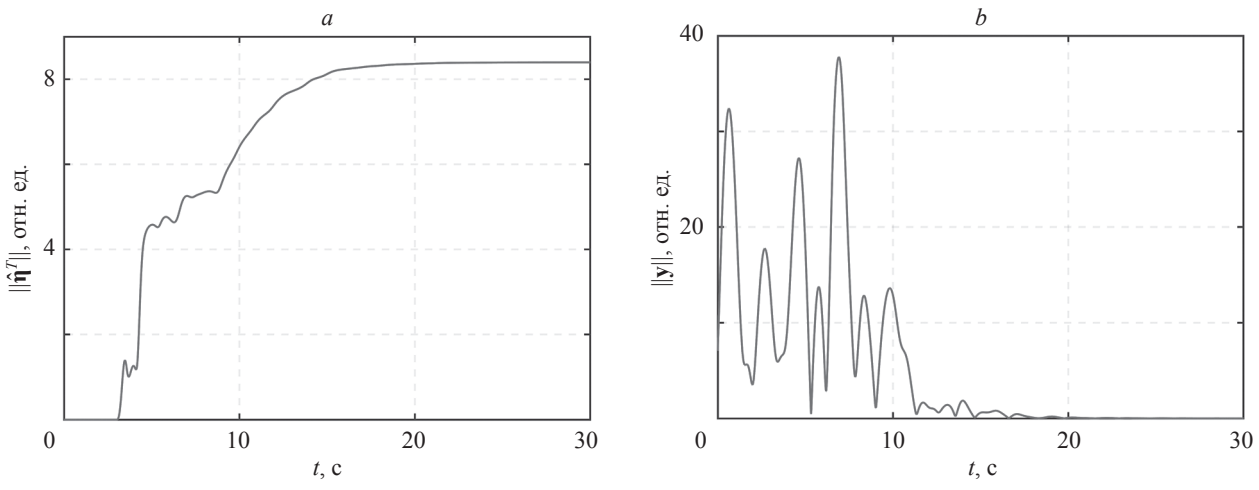

Рисунок. Переходные процессы в замкнутой системе при μ = 3: настраиваемого параметра $\hat{\boldsymbol{\eta}}^T$ (*a*); выходного сигнала системы **y** (*b*)

Figure. Transients in the closed-loop control system with μ = 3: transient process of the adjusted parameter $\hat{\boldsymbol{\eta}}^T$ (*a*); Transient process of the system output **y** (*b*)

$$\mathbf{G}_1 = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}, \mathbf{L}_1 = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

$$\mathbf{G}_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix}, \mathbf{L}_2 = \begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}.$$

Из выражения (7) получим: $\sigma_1 = 2$ и $\sigma_2 = 1$ для соответствующих выходов. Назначим полюса для первого канала такими, что $\lambda_{11} = -6$ и $\lambda_{12} = -2$, а для второго — $\lambda_{21} = -3$. В этом случае матрицы **F** и **K**, используемые для синтеза управления, примут вид

$$\mathbf{F} = \begin{bmatrix} 12 & 0 \\ 0 & 3 \end{bmatrix}, \mathbf{K} = \begin{bmatrix} 12{,}1 & 7{,}1 & 0 \\ 3 & 1 & 2 \end{bmatrix}.$$

Для синтеза закона управления и алгоритма адаптации сконструируем наблюдатель возмущений в соответствии с уравнением (13) с матрицами:

$$\mathbf{G} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -3 & -4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -6 & -11 & -6 \end{bmatrix},$$

$$\mathbf{N} = \begin{bmatrix} 0 & 0 & 0 \\ 0{,}6667 & -0{,}6667 & 1{,}3333 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 2 & 4 & -2 \end{bmatrix}.$$

При выборе коэффициента адаптации μ = 3 и фильтра $\mathbf{H}(s) = \dfrac{1}{s+1}$, а также при моделировании с начальным условием $\mathbf{x}(t) = [-1;\ 1;\ 0]$, получим результаты, представленные на рисунке.

Результаты моделирования показали работоспособность предложенного подхода к синтезу регулятора, который позволяет обеспечить ограниченность всех сигналов в замкнутой системе и выполнение целевого условия (2).

Заключение

Рассмотрена задача адаптивной компенсации для линейных стационарных систем с произвольными запаздываниями в канале управления и неизвестными внешними возмущениями. Предложенный метод обеспечивает эффективную компенсацию возмущений даже в случае наличия различных величин запаздывания в каналах управления. Результаты показывают, что данный метод может быть применен к неустойчивым системам без необходимости дополнительного формирования стабилизирующей компоненты управляющего сигнала. Однако метод реализуем для квадратных систем, имеющих равное количество входов и выходов.

Представленное решение может быть расширено на случай проектирования закона управления для класса систем с неизмеряемым вектором состояния, что является направлением для будущих исследований.

Литература

1. Wang Q-G. Decoupling Control. Springer, 2003. 369 p. https://doi.org/10.1007/3-540-46151-5
2. Morgan B.S. The synthesis of linear multivariable systems by state-variable feedback // IEEE Transactions on Automatic Control. 1964. V. 9. N 4. P. 405–411. https://doi.org/10.1109/TAC.1964.1105733
3. Gilbert E.G. The decoupling of multivariable systems by state feedback // SIAM Journal on Control. 1969. V. 7. N 1. P. 50–63. https://doi.org/10.1137/0307004
4. Falb P., Wolovich W. Decoupling in the design and synthesis of multivariable control systems // IEEE Transactions on Automatic Control. 1967. V. 12. N 6. P. 651–659. https://doi.org/10.1109/TAC.1967.1098737
5. Chu D., Malabre M. Numerically reliable design for proportional and derivative state-feedback decoupling controller // Automatica. 2002. V. 38. N 12. P. 2121–2125. https://doi.org/10.1016/S0005-1098(02)00138-3
6. Nijmeijer H., Respondek W. Dynamic input-output decoupling of nonlinear control systems // IEEE Transactions on Automatic Control. 1988. V. 33. N 11. P. 1065–1070. https://doi.org/10.1109/9.14420
7. Mertzios B.G., Christodoulou M.A. Decoupling and pole-zero assignment of singular systems with dynamic state feedback // Circuits, Systems and Signal Processing. 1986. V. 5. N 1. P. 49–68. https://doi.org/10.1007/bf01600186
8. Estrada M.B., Malabre M. Proportional and derivative state-feedback decoupling of linear systems // IEEE Transactions on Automatic Control. 2000. V. 45. N 4. P. 730–733. https://doi.org/10.1109/9.847111
9. Angelico B., Barbosa F.S., Toriumi F. State feedback decoupling control of a control moment gyroscope // Journal of Control, Automation and Electrical Systems. 2017. V. 28. N 1. P. 26–35. https://doi.org/10.1007/s40313-016-0277-8
10. Francis B.A., Wonham W.M. The internal model principle for linear multivariable regulators // Applied Mathematics and Optimization. 1975. V. 2. N 2. P. 170–194. https://doi.org/10.1007/bf01447855
11. Davison E. The robust control of a servomechanism problem for linear time-invariant multivariable systems // IEEE Transactions on Automatic Control. 1976. V. 21. N 1. P. 25–34. https://doi.org/10.1109/tac.1976.1101137
12. Johnson C. Accomodation of external disturbances in linear regulator and servomechanism problems // IEEE Transactions on Automatic Control. 1971. V. 16. N 6. P. 635–644. https://doi.org/10.1109/tac.1971.1099830
13. Nikiforov V.O. Nonlinear servocompensation of unknown external disturbances // Automatica. 2001. V. 37. N 10. P. 1647–1653. https://doi.org/10.1016/s0005-1098(01)00117-0
14. Nikiforov V.O. Adaptive non-linear tracking with complete compensation of unknown disturbances // European Journal of Control. 1998. V. 4. N 2. P. 132–139. https://doi.org/10.1016/s0947-3580(98)70107-4
15. Nikiforov V.O. Adaptive servomechanism controller with an implicit reference model // International Journal of Control. 1997. V. 68. N 2. P. 277–286. https://doi.org/10.1080/002071797223604
16. Nikiforov V.O., Gerasimov D.N. Adaptive Regulation: Reference Tracking and Disturbance Rejection. Springer, 2022. 376 p. https://doi.org/10.1007/978-3-030-96091-9
17. Isidori A. Lectures in Feedback Design for Multivariable Systems. Springer, 2017. 424 p. https://doi.org/10.1007/978-3-319-42031-8
18. Narendra K.S., Annaswamy A.M. Stable Adaptive Systems. PrenticeHall, 1989. 494 p.
19. Gerasimov D.N., Nikiforov V.O., Paramonov A.V. Adaptive disturbance compensation in delayed linear systems: Internal model approach // Proc. of the IEEE Conference on Control Applications (CCA). 2015. P. 1692–1696. https://doi.org/10.1109/cca.2015.7320853
20. Gerasimov D.N., Paramonov A.V., Nikiforov V.O. Algorithms of adaptive disturbance compensation in linear systems with arbitrary input delay // International Journal of Control. 2020. V. 93. N 7. P. 1596–1604. https://doi.org/10.1080/00207179.2018.1521527
21. Gerasimov D.N., Paramonov A.V., Nikiforov V.O. Algorithms of fast adaptive compensation of disturbance in linear systems with arbitrary input delay // IFAC-PapersOnLine. 2017. V. 50. N 1. P. 12892–12897. https://doi.org/10.1016/j.ifacol.2017.08.1942
22. Nikiforov V.O., Paramonov A.V., Gerasimov D.N., Pashenko A.V. Adaptive compensation of unmatched disturbances in MIMO LTI

References

1. Wang Q-G. *Decoupling Control*. Springer, 2003, 369 p. https://doi.org/10.1007/3-540-46151-5
2. Morgan B.S. The synthesis of linear multivariable systems by state-variable feedback. *IEEE Transactions on Automatic Control*, 1964, vol. 9, no. 4, pp. 405–411. https://doi.org/10.1109/TAC.1964.1105733
3. Gilbert E.G. The decoupling of multivariable systems by state feedback. *SIAM Journal on Control*, 1969, vol. 7, no. 1, pp. 50–63. https://doi.org/10.1137/0307004
4. Falb P., Wolovich W. Decoupling in the design and synthesis of multivariable control systems. *IEEE Transactions on Automatic Control*, 1967, vol. 12, no. 6, pp. 651–659. https://doi.org/10.1109/TAC.1967.1098737
5. Chu D., Malabre M. Numerically reliable design for proportional and derivative state-feedback decoupling controller. *Automatica*, 2002, vol. 38, no. 12, pp. 2121–2125. https://doi.org/10.1016/S0005-1098(02)00138-3
6. Nijmeijer H., Respondek W. Dynamic input-output decoupling of nonlinear control systems. *IEEE Transactions on Automatic Control*, 1988, vol. 33, no. 11, pp. 1065–1070. https://doi.org/10.1109/9.14420
7. Mertzios B.G., Christodoulou M.A. Decoupling and pole-zero assignment of singular systems with dynamic state feedback. *Circuits, Systems and Signal Processing*, 1986, vol. 5, no. 1, pp. 49–68. https://doi.org/10.1007/bf01600186
8. Estrada M.B., Malabre M. Proportional and derivative state-feedback decoupling of linear systems. *IEEE Transactions on Automatic Control*, 2000, vol. 45, no. 4, pp. 730–733. https://doi.org/10.1109/9.847111
9. Angelico B., Barbosa F.S., Toriumi F. State feedback decoupling control of a control moment gyroscope. *Journal of Control, Automation and Electrical Systems*, 2017, vol. 28, no. 1, pp. 26–35. https://doi.org/10.1007/s40313-016-0277-8
10. Francis B.A., Wonham W.M. The internal model principle for linear multivariable regulators. *Applied Mathematics and Optimization*, 1975, vol. 2, no. 2, pp. 170–194. https://doi.org/10.1007/bf01447855
11. Davison E. The robust control of a servomechanism problem for linear time-invariant multivariable systems. *IEEE Transactions on Automatic Control*, 1976, vol. 21, no. 1, pp. 25–34. https://doi.org/10.1109/tac.1976.1101137
12. Johnson C. Accomodation of external disturbances in linear regulator and servomechanism problems. *IEEE Transactions on Automatic Control*, 1971, vol. 16, no. 6, pp. 635–644. https://doi.org/10.1109/tac.1971.1099830
13. Nikiforov V.O. Nonlinear servocompensation of unknown external disturbances. *Automatica*, 2001, vol. 37, no. 10, pp. 1647–1653. https://doi.org/10.1016/s0005-1098(01)00117-0
14. Nikiforov V.O. Adaptive non-linear tracking with complete compensation of unknown disturbances. *European Journal of Control*, 1998, vol. 4, no. 2. pp. 132–139. https://doi.org/10.1016/s0947-3580(98)70107-4
15. Nikiforov V.O. Adaptive servomechanism controller with an implicit reference model. *International Journal of Control*, 1997, vol. 68, no. 2, pp. 277–286. https://doi.org/10.1080/002071797223604
16. Nikiforov V.O., Gerasimov D.N. *Adaptive Regulation: Reference Tracking and Disturbance Rejection*. Springer, 2022, 376 p. https://doi.org/10.1007/978-3-030-96091-9
17. Isidori A. *Lectures in Feedback Design for Multivariable Systems*. Springer, 2017, 424 p. https://doi.org/10.1007/978-3-319-42031-8
18. Narendra K.S., Annaswamy A.M. *Stable Adaptive Systems*. PrenticeHall, 1989, 494 p.
19. Gerasimov D.N., Nikiforov V.O., Paramonov A.V. Adaptive disturbance compensation in delayed linear systems: Internal model approach. *Proc. of the IEEE Conference on Control Applications (CCA)*, 2015, pp. 1692–1696. https://doi.org/10.1109/cca.2015.7320853
20. Gerasimov D.N., Paramonov A.V., Nikiforov V.O. Algorithms of adaptive disturbance compensation in linear systems with arbitrary input delay. *International Journal of Control*, 2020, vol. 93, no. 7, pp. 1596–1604. https://doi.org/10.1080/00207179.2018.1521527
21. Gerasimov D.N., Paramonov A.V., Nikiforov V.O. Algorithms of fast adaptive compensation of disturbance in linear systems with arbitrary input delay. *IFAC-PapersOnLine*, 2017, vol. 50, no. 1, pp. 12892–12897. https://doi.org/10.1016/j.ifacol.2017.08.1942
22. Nikiforov V.O., Paramonov A.V., Gerasimov D.N., Pashenko A.V. Adaptive compensation of unmatched disturbances in MIMO LTI

plants with input delay // Proc. of the 2021 American Control Conference (ACC). 2021. P. 2430–2435. https://doi.org/10.23919/acc50511.2021.9482716
23. Nikiforov V.O., Paramonov A.V., Gerasimov D.N. Adaptive compensation of unmatched disturbances in unstable MIMO LTI plants with distinct input delays // IFAC-PapersOnLine. 2023. V. 56. N 2. P. 9179–9184. https://doi.org/10.1016/j.ifacol.2023.10.159
24. Герасимов Д.Н., Парамонов А.В., Никифоров В.О. Алгоритм компенсации мультигармонических возмущений в линейных системах с произвольным запаздыванием: метод внутренней модели // Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16. № 6. С. 1023–1030. https://doi.org/10.17586/2226-1494-2016-16-6-1023-1030

plants with input delay. *Proc. of the 2021 American Control Conference (ACC)*, 2021, pp. 2430–2435. https://doi.org/10.23919/acc50511.2021.9482716
23. Nikiforov V.O., Paramonov A.V., Gerasimov D.N. Adaptive compensation of unmatched disturbances in unstable MIMO LTI plants with distinct input delays. *IFAC-PapersOnLine*, 2023, vol. 56, no. 2, pp. 9179–9184. https://doi.org/10.1016/j.ifacol.2023.10.159
24. Gerasimov D.N., Paramonov A.V., Nikiforov V.O. Algorithm of multiharmonic disturbance compensation in linear systems with arbitrary delay: internal model approach. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2016, vol. 16, no. 6, pp. 1023–1030. (in Russian). https://doi.org/10.17586/2226-1494-2016-16-6-1023-1030

Авторы

Ты Конг Винь — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0005-5106-8560, congvinhvkdn@gmail.com
Дударенко Наталия Александровна — кандидат технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, SC 36662185600, https://orcid.org/0000-0002-3553-0584, dudarenko@itmo.ru

Authors

Cong Vinh Tu — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0005-5106-8560, congvinhvkdn@gmail.com
Natalia A. Dudarenko — PhD, Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, SC 36662185600, https://orcid.org/0000-0002-3553-0584, dudarenko@itmo.ru

Статья поступила в редакцию 02.06.2025
Одобрена после рецензирования 12.09.2025
Принята к печати 17.11.2025

Received 02.06.2025
Approved after reviewing 12.09.2025
Accepted 17.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1089-1097
УДК 004.942

Разработка квазиоптимального алгоритма переключения мобильного робота

Дмитрий Николаевич Захаров[1], Александр Дмитриевич Панин[2], Андрей Михайлович Яременко[3]✉, Дмитрий Романович Алиев[4], Максим Игоревич Дербин[5], Олег Игоревич Борисов[6]

[1,2,3,4,5,6] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] dnzakharov@itmo.ru, https://orcid.org/0000-0001-9017-1505
[2] sashapanin220103@gmail.com, https://orcid.org/0009-0005-0209-878X
[3] amyaremenko@itmo.ru✉, https://orcid.org/0009-0007-7220-3957
[4] dmitrii_aliev@itmo.ru, https://orcid.org/0009-0000-4057-965X
[5] matius_d@list.ru, https://orcid.org/0009-0005-6451-0485
[6] borisov@itmo.ru, https://orcid.org/0000-0003-2861-746X

Аннотация
Введение. Всенаправленные мобильные платформы, известные своей маневренностью в ограниченных пространствах, часто сталкиваются не только с проблемами энергоэффективности из-за конструкции роликонесущих колес, но и с ограничениями при движении в реальных условиях эксплуатации, такими как перепады высот и неровности рельефа. Для преодоления ограничений мобильных платформ необходимо обеспечить возможность переключения между всенаправленным и классическим режимами движения с помощью адаптивного переключения режимов движения. Такой подход позволяет сочетать маневренность при навигации в тесных условиях с повышенной проходимостью и энергоэффективностью на неровных поверхностях и уклонах. **Метод.** В работе предложен алгоритм адаптивного переключения режимов движения, обеспечивающий переход от всенаправленной к классической кинематической схеме и обратно с помощью разработанного компактного механизма переключения. Для этого использованы усовершенствованные кинематические, динамические и энергетические модели в сочетании с лабораторными экспериментами с реконфигурируемой платформой. Предложенные усовершенствования позволяют осуществить простой и быстрый переход из одной кинематической схемы в другую с помощью разработанного компактного механизма переключения. **Основные результаты.** Экспериментальные исследования выполнены в лабораторных условиях на ровной бетонной поверхности, где робот выполнял движение по замкнутой траектории. В ходе эксперимента фиксировались энергопотребление и ошибки отслеживания траектории для голономного, неголономного и реконфигурируемого режимов движения. Сравнительный анализ показал, что применение предложенного алгоритма переключения позволяет снизить энергозатраты в среднем на 8 % при сохранении маневренности. Для более крупных роботов с массой, значительно превышающей массу механизма реконфигурации, экономия энергии в реальных сценариях может быть больше за счет оптимизации системой использования энергии и выбора наиболее эффективной конфигурации для различных сегментов траектории. Система сохраняет высокую маневренность и обеспечивает эффективную навигацию в сложных условиях. **Обсуждение.** Представленный алгоритм позволяет платформе достичь важнейшего баланса между мобильностью, эффективностью и точностью управления. Появляется возможность практического внедрения реконфигурируемых роботов в реальных сервисных приложениях. Полученные результаты имеют практическое значение для проектирования адаптивных механических и управляющих систем с повышенной эксплуатационной гибкостью мобильных платформ в условиях ограниченных ресурсов.

Ключевые слова
мобильный робот, роликонесущее колесо, всенаправленный робот, гибридная голономность, квазиоптимальный алгоритм

Благодарности
Работа выполнена при финансовой поддержке Российского научного фонда (проект № 25-29-00713).

Ссылка для цитирования: Захаров Д.Н., Панин А.Д., Яременко А.М., Алиев Д.Р., Дербин М.И., Борисов О.И. Разработка квазиоптимального алгоритма переключения мобильного робота // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1089–1097. doi: 10.17586/2226-1494-2025-25-6-1089-1097

© Захаров Д.Н., Панин А.Д., Яременко А.М., Алиев Д.Р., Дербин М.И., Борисов О.И., 2025

Experimental study of a quasi-optimal mobile robot switching algorithm

Dmitry N. Zakharov[1], Alexandr D. Panin[2], Andrey M. Iaremenko[3]✉, Dmitry R. Aliev[4], Maksim I. Derbin[5], Oleg I. Borisov[6]

[1,2,3,4,5,6] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] dnzakharov@itmo.ru, https://orcid.org/0000-0001-9017-1505
[2] sashapanin220103@gmail.com, https://orcid.org/0009-0005-0209-878X
[3] amyaremenko@itmo.ru✉, https://orcid.org/0009-0007-7220-3957
[4] dmitrii_aliev@itmo.ru, https://orcid.org/0009-0000-4057-965X
[5] matius_d@list.ru, https://orcid.org/0009-0005-6451-0485
[6] borisov@itmo.ru, https://orcid.org/0000-0003-2861-746X

Abstract

Omnidirectional mobile platforms, known for their exceptional maneuverability in confined spaces, often encounter not only energy efficiency challenges due to the design of roller-bearing wheels but also operational limitations in real-world environments such as height differences and uneven terrain. To overcome these limitations, it is necessary to enable switching between omnidirectional and conventional driving modes through adaptive motion mode switching. This approach combines the maneuverability required for navigation in tight spaces with improved off-road capability and energy efficiency on uneven surfaces and slopes. This study proposes an algorithm for adaptive motion mode switching, providing transitions from an omnidirectional to a classical kinematic scheme and back via a specially developed compact switching mechanism. To achieve this, enhanced kinematic, dynamic, and energy models were utilized in combination with laboratory experiments conducted on a reconfigurable platform. The proposed improvements make it possible to perform a simple and rapid transition between kinematic configurations using the compact switching mechanism. Experimental studies were carried out under laboratory conditions on a flat concrete surface where the robot followed a closed trajectory. During the experiments, energy consumption and trajectory-tracking errors were recorded for holonomic, nonholonomic, and reconfigurable motion modes. Comparative analysis demonstrated that the proposed switching algorithm reduces energy consumption by an average of 8 % while maintaining maneuverability. For larger robots whose total mass significantly exceeds that of the reconfiguration mechanism energy savings in real-world scenarios can be even greater due to the system ability to optimize energy usage and select the most efficient configuration for different trajectory segments. The system retains high maneuverability and ensures efficient navigation in complex environments. The presented algorithm enables the platform to achieve a crucial balance between mobility, efficiency, and control accuracy. This opens the possibility for the practical implementation of reconfigurable robots in real-world service applications. The obtained results have practical significance for the design of adaptive mechanical and control systems that enhance the operational flexibility of mobile platforms under resource-constrained conditions.

Keywords

mobile robot, wheel with rollers, omnidirectional robot, hybrid holonomy, quasi-optimal algorithm

Acknowledgments

This work was supported by the Russian Science Foundation, project No. 25-29-00713.

For citation: Zakharov D.N., Panin A.D., Iaremenko A.M., Aliev D.R., Derbin M.I., Borisov O.I. Experimental study of a quasi-optimal mobile robot switching algorithm. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1089–1097 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1089-1097

Введение

Сервисная робототехника — быстро развивающаяся область, которая имеет множество применений в различных сферах человеческой деятельности [1]. Мобильные платформы используются для решения задач, связанных с уборкой, доставкой [2], инспектированием, мониторингом и спасательными операциями [3]. Разработка всенаправленных платформ считается одним из наиболее перспективных направлений в наземной робототехнике [4–6]. Эти роботизированные платформы обладают высокой маневренностью в ограниченных пространствах, что позволяет использовать их в сложных условиях. Однако всенаправленные платформы имеют значительные ограничения, связанные с их конструкцией. Классические всенаправленные платформы оснащены колесами с роликами, которые свободно вращаются вокруг своей оси, что приводит к высокому энергопотреблению и снижению безопасности и являются критическими факторами для мобильных сервисных роботов. В работе [7] выполнен анализ различных конструкций мобильных роботов, включая всенаправленные роботы, и рассмотрено их энергопотребление и безопасность. В отечественном промышленном секторе остаются многочисленные предприятия, которые при проектировании не предусматривали роботизацию. Тем не менее именно они сегодня требуют внедрения мобильных роботов с реконфигурируемыми механизмами, способными адаптироваться к особенностям существующей инфраструктуры, такими как неравномерность напольного покрытия.

Существует ряд исследований, в которых ученые применяют адаптивные алгоритмы для улучшения характеристик таких роботов. В работе [8] представлен алгоритм адаптивного нечеткого динамического управления поверхностью для отслеживания траектории мобильного робота с четырьмя всенаправленными колесами, предназначенный для адаптивной настройки контроллера в ответ на изменения в динамике робота

или наличие неопределенностей. В [9] рассмотрены методы и алгоритмы большого роя роботов, обеспечивающие их управляемость в динамически изменяющихся условиях. Методы синтеза интеллектуальных систем управления для мобильных роботов, способных адаптироваться к изменяющимся условиям окружающей среды, обсуждаются в работе [10]. В [11] разработан механизм, позволяющий всенаправленному роботу с роликонесущими колесами реконфигурировать в обычного четырехколесного робота.

В настоящей работе представлена разработка алгоритма энергоэффективного управления системой реконфигурации, основанный на кинематическом, динамическом и энергетическом анализах ранее разработанного робота [11]. Данный алгоритм, использующий базовый контроллер отслеживания траектории, переключает механизм реконфигурации между обычным и всенаправленным режимами, минимизируя потребление энергии и погрешность отслеживания, одновременно позволяя роботу следовать по заранее заданным траекториям с помощью простых стратегий управления.

Описание предлагаемой методики

Методика основана на сравнительном анализе энергопотребления и точности отслеживания траектории в различных режимах движения. Для этого использовались усовершенствованные кинематические, динамические и энергетические модели, а также лабораторные эксперименты с реконфигурируемой платформой. Общая схема управления показана на рис. 1.

Рассматривается робот с двойным режимом мобильности как объект управления с переменной структурой, режим которого (голономный или неголономный) может быть изменен с помощью представляемого механизма реконфигурации. Предложенные в настоящей работе математические модели (кинематическая, динамическая и энергетическая) позволили синтезировать алгоритм переключения механизма реконфигурации. Полученный алгоритм обеспечивает баланс между энергоэффективностью и точностью при движении по траектории. Алгоритм принимает желаемую траекторию в качестве входных данных и вычисляет квазиоптимальный режим для каждого из ее сегментов на основе данных моделирования. Отслеживание траектории выполняется с применением базового контроллера.

Центральным устройством в системе, осуществляющим переход между режимами, является устройство реконфигурации (рис. 2).

Сжатый воздух подается во внутреннюю камеру *5* через поворотный быстросъемный фитинг *1*, расположенный на оси вращения колеса, посредством ниппеля *3*. Ниппель *3* является частью внутренней камеры (как у велосипедов) и ввинчен в вал *2*. Внутренняя камера *5* надувается, равномерно расширяясь и воздействуя на ролики колеса *6*. Прижимная пластина *7* обеспечивает улучшенное сцепление камеры *5* с роликом *6*.

Математическая модель робота с устройством реконфигурации

Кинематическая модель. Кинематическая модель робота с устройством реконфигурации может быть представлена следующим образом:

$$[\omega_1 \quad \omega_2 \quad \omega_3 \quad \omega_4]^T = \mathbf{J_h}[V_x \quad V_y \quad \omega_z],$$

где ω_i — угловая скорость i-го колеса; V_x, V_y, ω_z — линейные и угловая скорости в локальной системе координат. Согласно [12],

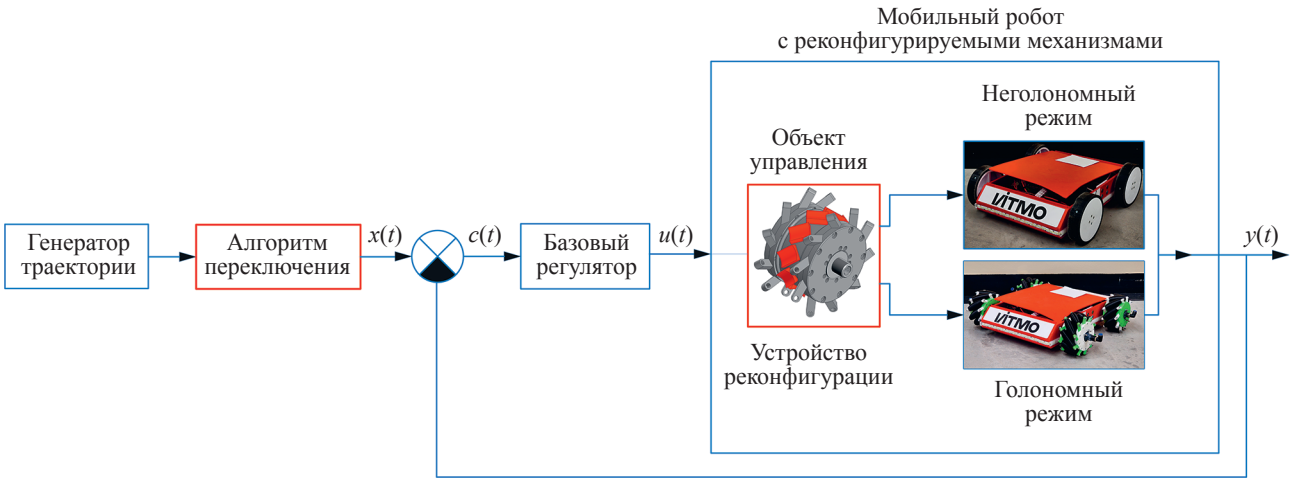

Рис. 1. Схема управления мобильным роботом с реконфигурируемыми механизмами.

$x(t)$ — желаемое положение; $e(t)$ — ошибка рассогласования между реальным положением и желаемым; $u(t)$ — управляющее воздействие; $y(t)$ — действительное положение

Fig. 1. Control scheme of a mobile robot with reconfigurable mechanisms.

$x(t)$ — desired position, $e(t)$ — error between the actual position and the desired position, $u(t)$ — control input, and $y(t)$ — actual position

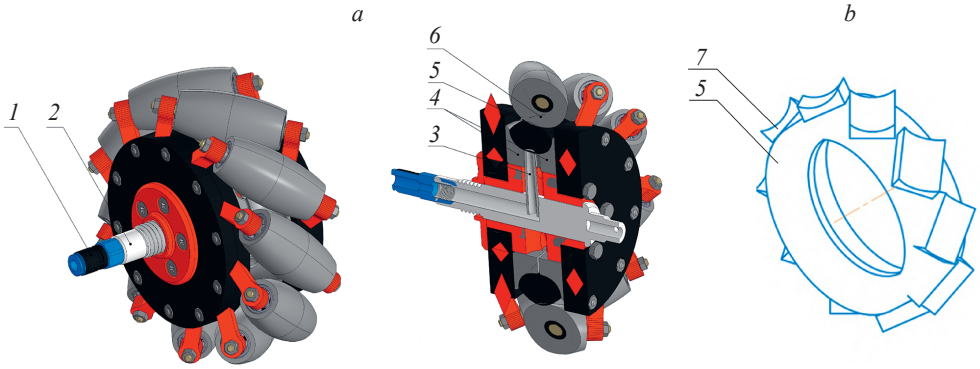

Рис. 2. Устройство механизма реконфигурации: колесо в сборе (*a*), конструкция блокирующего устройства (*b*).

1 — поворотный вставной фитинг; *2* — вал с отверстием для подачи и отвода воздуха; *3* — ниппель; *4* — ограничитель для укладки внутренней трубки; *5* — пневматическая камера; *6* — колесный ролик; *7* — прижимная пластина

Fig. 2. Reconfiguration mechanism device: assembled wheel (*a*); design of the locking device (*b*).

1 — rotary push-in fitting; *2* — shaft with an air supply and exhaust hole; *3* — nipple; *4* — retainer for inner tube placement; *5* — pneumatic chamber; *6* — wheel roller; *7* — pressure plate

$$\mathbf{J_h} = \begin{cases} \mathbf{J_{nh}} = \dfrac{1}{r}\begin{bmatrix} -1 & 0 & -L & -W \\ -1 & 0 & -L & -W \\ 1 & 0 & -L & -W \\ 1 & 1 & -L & -W \end{bmatrix}, \\ \text{если ролики заблокированы,} \\ \mathbf{J_m} = \dfrac{1}{r}\begin{bmatrix} -1 & -1 & -L & -W \\ -1 & 1 & -L & -W \\ 1 & -1 & -L & -W \\ 1 & 0 & -L & -W \end{bmatrix}, \\ \text{если ролики не заблокированы,} \end{cases}$$

где $\mathbf{J_h}$, $\mathbf{J_{nh}}$ и $\mathbf{J_m}$ — матрицы Якоби для гибридного, неголономного и голономного роботов; r — радиус колеса; L и W — расстояния до колес по осям x и y.

Динамическая модель. Рассмотрим динамическую модель гибридного робота, используя полученную кинематическую модель:

$$\mathbf{M_h}(q)\ddot{q} = \mathbf{J_h^1}(\boldsymbol{\tau} + \mathbf{F}_f r),$$

где $\mathbf{M_h}$ — матрица массы и инерции; $\boldsymbol{\tau}$ — вектор моментов колес робота; $\mathbf{F}_f$ — сила трения; $\mathbf{J_h^1}$ — псевдобратная матрица Якоби.

При этом матрица $\mathbf{M_h}$:

$$\mathbf{M_h} = \begin{cases} \mathbf{M_{nh}} = \begin{bmatrix} \rho_1\cos^2(\theta) & \dfrac{\rho_1}{2}\sin(2\theta) & 0 \\ \dfrac{\rho_1}{2}\sin(2\theta) & \rho_1\sin^2(\theta) & 0 \\ 0 & 0 & \rho_1 \end{bmatrix}, \\ \text{если ролики заблокированы,} \\ \mathbf{M_{nh}} = \begin{bmatrix} \rho_2 & 0 & 0 \\ 0 & \rho_2 & 0 \\ 0 & 0 & \dfrac{I_w(L+W)^2 + I_r r^2}{(L+W)^2} \end{bmatrix}, \\ \text{если ролики не заблокированы,} \end{cases}$$

где $\rho_1 = \dfrac{1}{2}m_r r^2 + I_w + \dfrac{I_r r^2}{4(L+W)^2}$, $\rho_2 = I_w + \dfrac{m_r r^2}{2}$. Здесь I_w — момент инерции колеса относительно его оси; I_r — момент инерции робота относительно вертикальной оси; m_r — масса робота с учетом колес.

Для полученной динамической модели роботов применим модель трения всенаправленного робота, предложенную в работе [13]:

$$\mathbf{f}_{i,s} = \mu_s \mathbf{N}_i \operatorname{sgn}(\mathbf{V}_{i,p})\hat{\mathbf{u}}_{i,p},$$

$$\mathbf{f}_{i,r} = \mu_r \mathbf{N}_i \operatorname{sgn}(\mathbf{V}_{i,p})\hat{\mathbf{u}}_{i,p},$$

$$\mathbf{f}_{i,v} = \mu_v \mathbf{N}_i \mathbf{V}_{i,p},$$

где $\mathbf{f}_{i,s}$ — сила трения скольжения; $\mathbf{f}_{i,r}$ — сила трения качения; $\mathbf{f}_{i,v}$ — сила вязкого трения; $\mathbf{V}_{i,p}$ — скорость контактного ролика i-го колеса; $\hat{\mathbf{u}}_{i,p}$ — единичный вектор; μ_s — коэффициент трения скольжения; μ_r — коэффициент трения качения; μ_v — коэффициент вязкого трения; $\mathbf{N}_i$ — реактивная сила i-го колеса.

Энергетическая модель. Согласно [13], общее энергопотребление робота при движении по траектории выразим в виде:

$$E_r = E_k + E_f + E_e + E_m + E_d =$$
$$= \int_{t_i}^{t_f} T dt + \int_{t_i}^{t_f} \mu_s N \nu + (\mu_v N \nu)\nu dt + R_a \int_{t_i}^{t_f} I_m{}^T \mathbf{I}_m dt +$$
$$+ M_R \int_{t_i}^{t_f} \dot{\theta} dt + \int_{t_i}^{t_f} P_d dt,$$

где E_k — потребление энергии для перемещения платформы; E_f — энергия, рассеиваемая на трение; E_e — энергия на преодоление момента трения; E_m — энергия, рассеиваемая в приводах; E_d — энергопотребление на оборудование робота; $[t_i, t_f]$ — временной интервал; T — кинетическая энергия; ν — линейная скорость точки контакта колеса с поверхностью; R_a — сопротивление цепи якоря; M_R — момент трения; P_d — общая мощность всех устройств; $\mathbf{I}_m$ — вектор тока электродвигателей:

$$\mathbf{I}_m = [I_{m1} \quad I_{m2} \quad I_{m3} \quad I_{m4}]^T.$$

При этом, помимо энергетических параметров, зависящих от времени и состояния, необходимо учитывать мгновенную энергию, затрачиваемую на переключение из голономного в неголономный режим и обратно. Для перехода из голономного в неголономный режим необходимо активировать устройство реконфигурации (заблокировать ролики колеса). Это требует нагнетания воздуха в камеру запорного устройства с помощью компрессора, что потребляет энергию следующим образом:

$$E_c = \frac{\frac{\gamma}{\gamma - 1} P_1 K \left[\left(\frac{P_2}{P_1} \right)^{\frac{\gamma - 1}{\gamma}} - 1 \right]}{\eta},$$

где E_c — энергия, потребляемая компрессором; γ — адиабатический показатель; P_1 — начальное давление в системе; P_2 — конечное давление; K — общий объем камер запорного устройства; η — коэффициент полезного действия компрессора. Кроме того, энергия должна быть затрачена на подачу сжатого воздуха к запорному механизму путем активации реле.

Таким образом, получим выражение энергии, потребляемой для активации устройства блокировки роликов:

$$E_l = P_r t_l + E_c,$$

где P_r — мощность реле; t_l — время накачивания камер.

Для перехода из неголономного режима в голономный необходимо открыть реле, сбрасывая давление воздуха из камеры наружу. Рассчитаем энергию, необходимую для обратного перехода:

$$E_o = P_r t_o + E_c,$$

где t_o — время спуска воздуха из камер.

Проектирование закона управления

Базовый регулятор. Для решения задачи слежения за траекторией был синтезирован регулятор на основе наблюдателя с высоким коэффициентом усиления, описанный в работе [14].

Алгоритм переключения. Учитывая траекторию, цель состоит в минимизации следующего функционала:

$$Q = \int_0^T (E + H\|q_e\|)dt,$$

где E — значение энергии; q_e — ошибка отслеживания; H — весовой коэффициент.

Введение функционала Q мотивировано компромиссом между энергоэффективностью и точностью позиционирования мобильных платформ. В частности, платформа с роликонесущими колесами обычно демонстрирует меньшую энергоэффективность по сравнению с платформой на обычных колесах. Однако кинематические ограничения, накладываемые в неголономном режиме движения, снижают точность позиционирования, особенно при резких маневрах. Функционал Q балансирует два данных аспекта с помощью весового коэффициента H.

Алгоритм принимает на вход траекторию и весовой коэффициент H. На выходе он формирует план переключения, представленный в виде временного массива моментов переключения. Для этого траектория разбивается на прямолинейные и криволинейные участки на основе заранее заданного порога кривизны. Далее моделируется энергопотребление для каждого режима движения вдоль траектории, и значение функционала Q вычисляется для каждого сегмента.

Строится план переключения по следующей схеме.

Даны: массивы $\mathbf{Q}$: $\mathbf{Q}^0 = [q_1^0, q_2^0, \ldots, q_n^0]$ (состояние механического робота); $\mathbf{Q}^1 = [q_1^1, q_2^1, \ldots, q_n^1]$ (состояние неголономного робота) и затраты на переход: $c_{0\to1} = \text{const}_1$, $c_{1\to0} = \text{const}_2$.

Цель — найти последовательность состояний $\mathbf{S} = [s_1, s_2, \ldots, s_n]$, которая минимизирует:

$$Q_{total} = \sum_{i=1}^{n} q_i^{s_i} + \sum_{\text{переходы}} c_{S_i \to S_{i+1}}.$$

Для достижения цели рассматриваются:

— $dp_0[i]$ — минимальная энергия до этапа i в состоянии 0, $dp_0[1] = e_1^0$;

— $dp_1[i]$ — минимальная энергия до этапа i в состоянии 1, $dp_1[i] = e_1^1$.

Вводится рекуррентное соотношение для ($2 \le i \le n$):

$$dp_0[i] = q_i^0 + \min(dp_0[i-1], dp_1[i-1] + c_{1\to0}),$$

$$dp_1[i] = q_1^0 + \min(dp_1[i-1], dp_0[i-1] + c_{0\to1}).$$

Определяется конечное состояние:

$$s_n = \begin{cases} 0, \text{ если } dp_0[n] \le dp_1[n] \\ 1, \text{ иначе} \end{cases}.$$

Используя обратное распространение, определяются состояния на каждом сегменте $i = (n-1, n-2, \ldots, 1)$:

— если $s_{i+1} = 0$:

$$s_i = \begin{cases} 0, \text{ если } dp_0[i+1] - q_{i+1}^0 = dp_0[i] \\ 1, \text{ иначе} \end{cases},$$

если $s_{i+1} = 1$:

$$s_n = \begin{cases} 0, \text{ если } dp_0[n] \le dp_1[n] \\ 1, \text{ иначе} \end{cases},$$

$$s_i = \begin{cases} 1, \text{ если } dp_1[i+1] - q_{i+1}^0 = dp_1[i] \\ 0, \text{ иначе} \end{cases}.$$

На основе массива состояний определяются временные границы сегментов, на которых происходит переключение.

Описание эксперимента

В качестве объекта исследования использовался мобильный робот с возможностью реконфигурации колесной базы. Характеристики испытуемого робота представлены в таблице.

Таблица. Параметры испытуемого робота

Table. Parameters of the test robot

Параметр	Значение
Габариты робота (длина × ширина × × высота), мм	240 × 290 × 120
Масса робота, кг	10
Радиус колеса, мм	56
Рабочее напряжение, В	12
Ширина колеса, мм	70,5
Масса колеса, кг	0,8
Максимальный радиус ролика, мм	14,5
Рабочее давление, МПа	0,6

Робот включает в себя пластиковый корпус, изготовленный из полилактида, одноплатный компьютер верхнего уровня, плату управления двигателями, контроллеры двигателей постоянного тока, рассчитанные на управление 12 В двигателями с рабочим током до 10 А. Пневматическая часть включает в себя стационарный компрессор с производительностью на выход не менее 30 л/мин, фильтр-регулятор со степенью очистки 40 мкм и блок электромагнитных клапанов, управляемых напряжением 12 В. Блок клапанов состоит из 8 электромагнитных соленоидов. В момент переключения однократно включаются четыре клапана на подачу воздуха, пневмокамеры расширяются, блокируя движение роликов. Для перехода в исходное состояние активируются оставшиеся четыре клапана, воздух стравливается в атмосферу и камеры сужаются, освобождая ролики. Эксперименты проводились в лабораторных условиях: бетонный и ровный пол, камеры робота и ролики для роликонесущих колес были изготовлены из термопластичного полиуретана. Аккумулятор для каждого эксперимента был полностью заряжен. Эксперимент проводился по замкнутой траектории (рис. 3). На траектории были определены точки переключения для реконфигурируемого робота. Энергопотребление и ошибки отслеживания вдоль траектории измерялись для каждого типа мобильной платформы. Данные представлены на рис. 4 и 5. Для замкнутой траектории (рис. 4) неголономный робот израсходовал 805 Дж энергии, в то время как голономный — 885 Дж.

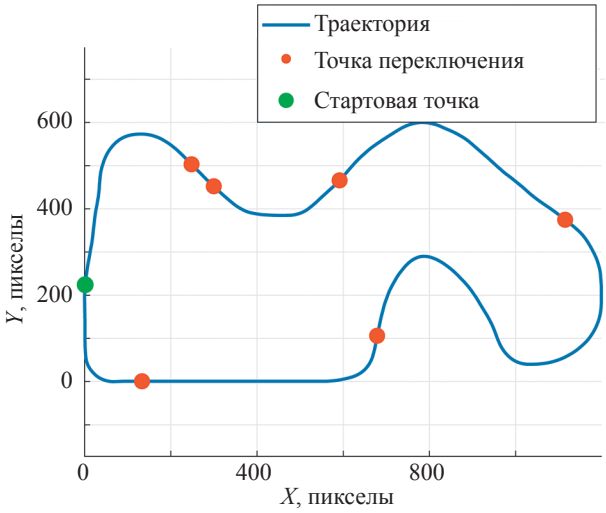

Рис. 3. Исходная траектория

Fig. 3. Initial trajectory

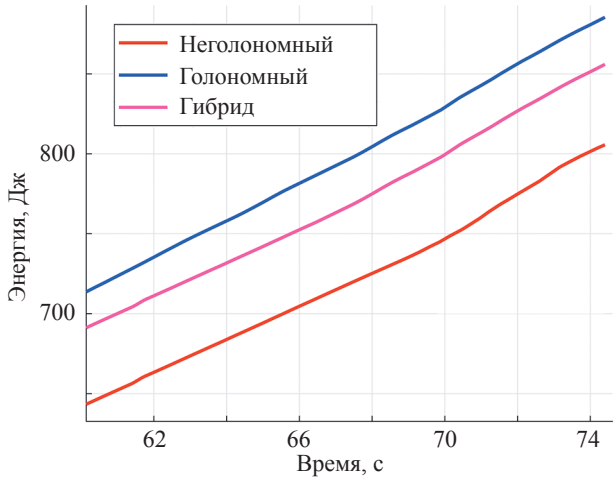

Рис. 4. Потребление энергии за одно прохождение по замкнутой траектории

Fig. 4. Energy consumption per one pass along the closed trajectory

Для оценки качества отслеживания были зафиксированы ошибки позиционирования робота относительно заданной траектории по осям x и y, а также вычислена норма ошибки слежения (рис. 5).

Результаты расчета функционала Q с учетом точек переключения показаны на рис. 6. В конечный момент времени значение функционала Q для реконфигурируемого (гибрид) робота, переключающегося, согласно предложенному алгоритму, составило 998 условных единиц (усл. ед.), в то время как для неголономного и голономного роботов — 1009 и 1003 усл. ед. соответственно.

Обсуждение результатов

Результаты экспериментов соответствуют исходным предположениям. Реконфигурируемый робот демонстрирует более низкую энергоэффективность на прямолинейных траекториях по сравнению с неголономным роботом. Полученная разница может быть обусловлена тем, что реконфигурируемые колеса не достигают идеально круглой формы, что приводит к неоптимальному использованию энергии. При этом, за счет алгоритма адаптивного переключения режима движения, реконфигурируемый робот способен экономить больше энергии по сравнению с голономным роботом. Предполагается, что в сценариях, более близких к реальным приложениям, мобильный реконфигурируемый робот может обеспечить более высокую энергоэффективность. Это ожидается из-за увеличения сил трения при элементарных движениях в неидеальных условиях, что, вероятно, приведет к более выраженным различиям в энергопотреблении.

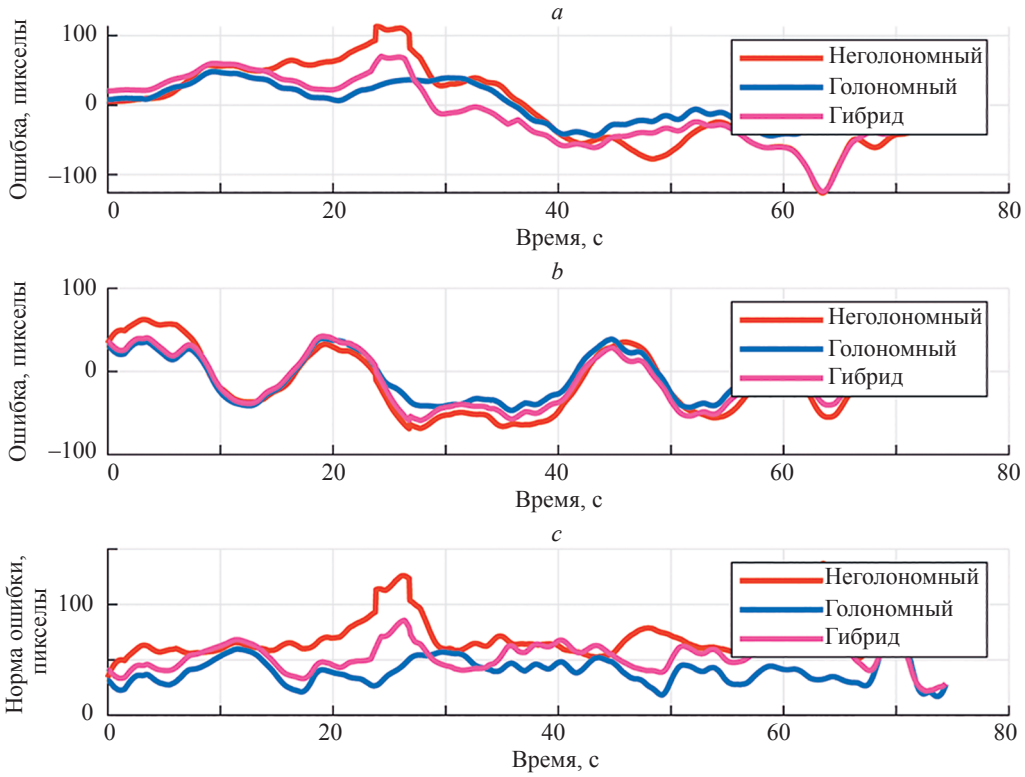

Рис. 5. Ошибки слежения: по осям *x* (*a*) и *y* (*b*); норма ошибки (*c*)

Fig. 5. Tracking errors: on the *x*-axis (*a*); on the *y*-axis (*b*); error norm (*c*)

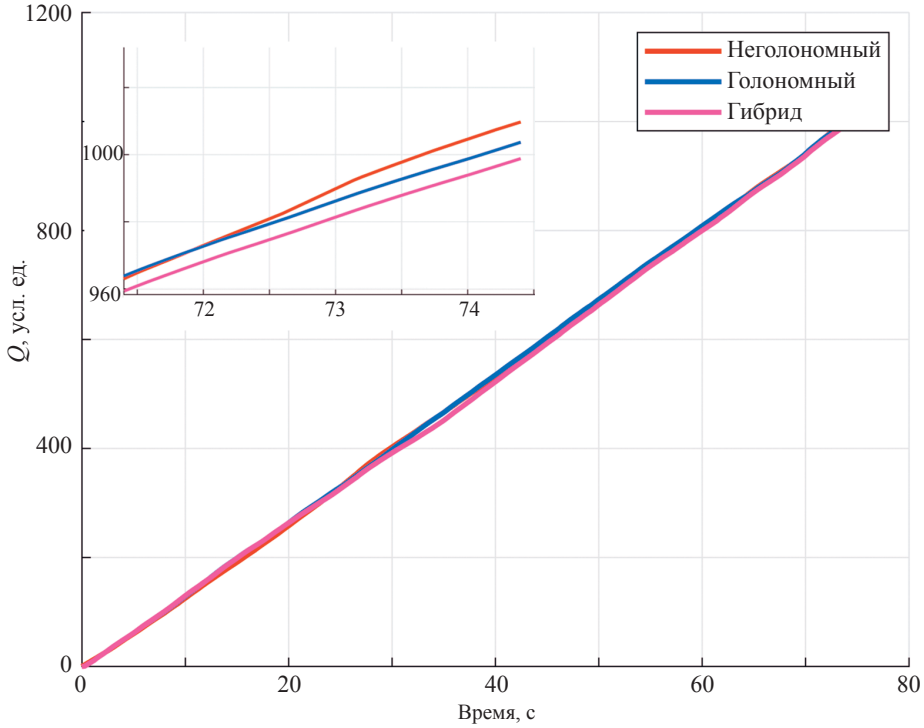

Рис. 6. График зависимости функционала *Q* от времени для всей траектории.

На вставке показан фрагмент кривых в диапазоне времени 71,5–74,5 с, демонстрирующий различие в значениях функционала между тремя типами роботов (неголономным, голономным и гибридным)

Fig. 6. Graph of the functional *Q* vs. time for the entire trajectory.

The inset shows a fragment of the curves in the time range 71.5–74.5 s, demonstrating the difference in functional values between three types of robots (nonholonomic, holonomic, and hybrid)

Также важно подчеркнуть адаптивные возможности платформы с двойным режимом передвижения. Например, в средах с присутствием людей или динамическими препятствиями платформа может отдавать приоритет маневренности для эффективной навигации. Напротив, в статичных средах, таких как пустые ангары или склады, платформа может использовать неголономный режим для быстрого и энергоэффективного передвижения. Эта динамическая адаптивность подчеркивает потенциальную полезность платформы в различных операционных контекстах.

Заключение

Представлен алгоритм квазиоптимального переключения режимов движения для реконфигурируемых роботов, который оптимизирует энергоэффективность при сохранении маневренности. Путем доработки кинематических, динамических и энергетических моделей был создан комплексный подход для анализа компромиссов между мобильностью, эффективностью и точностью управления. Результаты экспериментов подтвердили, что предложенный подход снижает энергопотребление на 8 % на тестовой траектории. Однако в реальных сценариях, особенно для более крупных роботов, где масса платформы значительно превышает массу механизма реконфигурации, экономия энергии может достигать больших значений. Несмотря на относительно небольшую экономию энергии в идеальных условиях, система обеспечивает сохранение маневренности там, где это свойство критично для задач слежения за траекторией. Полученные результаты демонстрируют перспективность применения гибридных стратегий передвижения для мобильных роботов, функционирующих в условиях ограниченных энергетических ресурсов. Будущие работы будут сосредоточены на дальнейшем совершенствовании стратегии переключения режимов, исследовании адаптивности в реальном времени и расширении экспериментальной валидации на более сложные сценарии и практические приложения.

Литература

1. Chen W., Wu P., Gong Y., Zhang Z., Wang K. The role of energy consumption in robotic mobile fulfillment systems: Performance evaluation and operating policies with dynamic priority // Omega. 2025. V. 130. P. 103168. https://doi.org/10.1016/j.omega.2024.103168
2. Kim T., Kang G., Lee D., Shim D.H. Development of an indoor delivery mobile robot for a multi-floor environment // IEEE Access. 2024. V. 12. P. 45202–45215. https://doi.org/10.1109/access.2024.3381489
3. Kebede G., Gelaw A., Andualem H., Hailu A.T. Review of the characteristics of mobile robots for health care application // International Journal of Intelligent Robotics and Applications. 2024. V 8. N 2. P. 480–502. https://doi.org/10.1007/s41315-024-00324-3
4. Giurgiu T., Bârsan G., Virca I., Pupăză C. Mecanum wheeled platforms for special applications // International conference KNOWLEDGE-BASED ORGANIZATION. 2022. V. 28. N 3. P. 44–51. https://doi.org/10.2478/kbo-2022-0086
5. Zhewen Z., Hongliu Y., Chengjia W., Pu H., Jiangui W. A comprehensive study on mecanum wheel-based mobility and suspension solutions for intelligent nursing wheelchairs // Scientific Reports. 2024. V. 14. N 1. P. 20644. https://doi.org/10.1038/s41598-024-71459-3
6. Qian J., Zi B., Wang D., Ma Y., Zhang D. The design and development of an omni-directional mobile robot oriented to an intelligent manufacturing system // Sensors. 2017. V. 17. N 9. P. 2073. https://doi.org/10.3390/s17092073
7. Vestman R. A Comparative Study of Omnidirectional and Differential Drive Systems for Mobile Manipulator Robots: A Performance Review of Strengths and Weaknesses. Dissertation. 2023. 60 p.
8. Minh D.N., Do Quang H., Phuong N.D., Manh T.N., Bui D.N. An adaptive fuzzy dynamic surface control tracking algorithm for mecanum wheeled mobile robot // International Journal of Mechanical Engineering and Robotics Research. 2023. V. 12. N 6. P. 354–361. https://doi.org/10.18178/ijmerr.12.6.354-361
9. Yang X., Hu Y., Gao H., Ding K., Li Z., Zhu P., Sun Y., Liu C. Risk-aware non-myopic motion planner for large-scale robotic swarm using cvar constraints // Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2024. P. 5784–5790. https://doi.org/10.1109/iros58592.2024.10802191
10. Zhu K., Zhang T. Deep reinforcement learning based mobile robot navigation: a review // Tsinghua Science and Technology. 2021. V. 26. N 5. P. 674–691. https://doi.org/10.26599/tst.2021.9010012
11. Zakharov D.N., Iaremenko A.M., Kurovskii D.M., Kurovskii A.M., Borisov O.I., Zhang B. Development of a mobile reconfigurable mecanum robot with a locking device of rollers // Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2024. P. 3553–3558. https://doi.org/10.1109/iros58592.2024.10801549

References

1. Chen W., Wu P., Gong Y., Zhang Z., Wang K. The role of energy consumption in robotic mobile fulfillment systems: Performance evaluation and operating policies with dynamic priority. *Omega*, 2025, vol. 130, pp. 103168. https://doi.org/10.1016/j.omega.2024.103168
2. Kim T., Kang G., Lee D., Shim D.H. Development of an indoor delivery mobile robot for a multi-floor environment. *IEEE Access*, 2024, vol. 12, pp. 45202–45215.https://doi.org/10.1109/access.2024.3381489
3. Kebede G., Gelaw A., Andualem H., Hailu A.T. Review of the characteristics of mobile robots for health care application. *International Journal of Intelligent Robotics and Applications*, 2024, vol 8, no. 2, pp. 480–502. https://doi.org/10.1007/s41315-024-00324-3
4. Giurgiu T., Bârsan G., Virca I., Pupăză C. Mecanum wheeled platforms for special applications. *International conference KNOWLEDGE-BASED ORGANIZATION*, 2022, vol. 28, no. 3, pp. 44–51. https://doi.org/10.2478/kbo-2022-0086
5. Zhewen Z., Hongliu Y., Chengjia W., Pu H., Jiangui W. A comprehensive study on mecanum wheel-based mobility and suspension solutions for intelligent nursing wheelchairs. *Scientific Reports*, 2024, vol. 14, no. 1, pp. 20644. https://doi.org/10.1038/s41598-024-71459-3
6. Qian J., Zi B., Wang D., Ma Y., Zhang D. The design and development of an omni-directional mobile robot oriented to an intelligent manufacturing system. *Sensors*, 2017, vol. 17, no. 9, pp. 2073. https://doi.org/10.3390/s17092073
7. Vestman R. A *Comparative Study of Omnidirectional and Differential Drive Systems for Mobile Manipulator Robots: A Performance Review of Strengths and Weaknesses*. Dissertation. 2023, 60 p.
8. Minh D.N., Do Quang H., Phuong N.D., Manh T.N., Bui D.N. An adaptive fuzzy dynamic surface control tracking algorithm for mecanum wheeled mobile robot. *International Journal of Mechanical Engineering and Robotics Research*, 2023, vol. 12, no. 6, pp. 354–361. https://doi.org/10.18178/ijmerr.12.6.354-361
9. Yang X., Hu Y., Gao H., Ding K., Li Z., Zhu P., Sun Y., Liu C. Risk-aware non-myopic motion planner for large-scale robotic swarm using cvar constraints. *Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2024, pp. 5784–5790. https://doi.org/10.1109/iros58592.2024.10802191
10. Zhu K., Zhang T. Deep reinforcement learning based mobile robot navigation: a review. *Tsinghua Science and Technology*, 2021, vol. 26, no. 5, pp. 674–691. https://doi.org/10.26599/tst.2021.9010012
11. *Zakharov D.N., Iaremenko A.M., Kurovskii D.M., Kurovskii A.M., Borisov O.I., Zhang B. Development of a mobile reconfigurable mecanum robot with a locking device of rollers. Proc. of the IEEE/RSJ International Conference on Intelligent Robots and Systems*

12. Abd Mutalib M.A., Azlan N.Z. Prototype development of mecanum wheels mobile robot: a review // Applied Research and Smart Technology (ARSTech). 2020. V. 1. N 2. P. 71–82. https://doi.org/10.23917/arstech.v1i2.39
13. Xie L., Herberger W., Xu W., Stol K.A. Experimental validation of energy consumption model for the four-wheeled omnidirectional mecanum robots for energy-optimal motion control // Proc. of the IEEE 14th International Workshop on Advanced Motion Control (AMC). 2016. P. 565–572. https://doi.org/10.1109/amc.2016.7496410
14. Zakharov D.N., Kurovskii D.M., Kurovskii A.M., Zhivitskii A.Y., Borisov O.I., Fakhro K., Pyrkin AA. Experimental study of trajectory control algorithms under incomplete information on parameters and state variables of a nonlinear model of mobile robot motion // Proc. of the 10th International Conference on Control, Decision and Information Technologies (CoDIT). 2024. P. 1584–1588. https://doi.org/10.1109/codit62066.2024.10708441

(IROS), 2024, pp. 3553–3558. https://doi.org/10.1109/iros58592.2024.10801549
12. Abd Mutalib M.A., Azlan N.Z. Prototype development of mecanum wheels mobile robot: a review. *Applied Research and Smart Technology (ARSTech)*, 2020, vol. 1, no. 2, pp. 71–82. https://doi.org/10.23917/arstech.v1i2.39
13. Xie L., Herberger W., Xu W., Stol K.A. Experimental validation of energy consumption model for the four-wheeled omnidirectional mecanum robots for energy-optimal motion control. *Proc. of the IEEE 14th International Workshop on Advanced Motion Control (AMC)*, 2016, pp. 565–572. https://doi.org/10.1109/amc.2016.7496410
14. Zakharov D.N., Kurovskii D.M., Kurovskii A.M., Zhivitskii A.Y., Borisov O.I., Fakhro K., Pyrkin AA. Experimental study of trajectory control algorithms under incomplete information on parameters and state variables of a nonlinear model of mobile robot motion. *Proc. of the 10th International Conference on Control, Decision and Information Technologies (CoDIT)*, 2024, pp. 1584–1588. https://doi.org/10.1109/codit62066.2024.10708441

Авторы

Захаров Дмитрий Николаевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0000-0001-9017-1505, dnzakharov@itmo.ru
Панин Александр Дмитриевич — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0005-0209-878X, sashapanin220103@gmail.com
Яременко Андрей Михайлович — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0007-7220-3957, amyaremenko@itmo.ru
Алиев Дмитрий Романович — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0000-4057-965X, dmitrii_aliev@itmo.ru
Дербин Максим Игоревич — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0005-6451-0485, matius_d@list.ru
Борисов Олег Игоревич — кандидат технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, SC 55858200900, https://orcid.org/0000-0003-2861-746X, borisov@itmo.ru

Authors

Dmitry N. Zakharov — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0000-0001-9017-1505, dnzakharov@itmo.ru
Alexandr D. Panin — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0005-0209-878X, sashapanin220103@gmail.com
Andrey M. Iaremenko — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0007-7220-3957, amyaremenko@itmo.ru
Dmitry R. Aliev — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0000-4057-965X, dmitrii_aliev@itmo.ru
Maksim I. Derbin — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0005-6451-0485, matius_d@list.ru
Oleg I. Borisov — PhD, Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, SC 55858200900, https://orcid.org/0000-0003-2861-746X, borisov@itmo.ru

Статья поступила в редакцию 13.08.2025
Одобрена после рецензирования 11.10.2025
Принята к печати 22.11.2025

Received 13.08.2025
Approved after reviewing 11.10.2025
Accepted 22.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1098-1106
УДК 681.513

Управление с предиктором по выходу для неустойчивых линейных систем с входным запаздыванием

Антон Александрович Пыркин[1]✉, Константин Юрьевич Калинин[2], Хан Конг Чан[3]

[1,2,3] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] pyrkin@itmo.ru✉, https://orcid.org/0000-0001-8806-4057
[2] kgkalinin@gmail.com, https://orcid.org/0009-0009-3092-6254
[3] werkaiaye@gmail.com, https://orcid.org/0009-0007-1116-7876

Аннотация
Введение. Рассмотрена классическая задача управления по выходу для линейной системы с входным запаздыванием и постоянными известными параметрами. Объект управления может быть неустойчивым, из-за чего большинство известных методов становятся неэффективными или неконструктивными. **Метод.** Предложен новый алгоритм управления на основе наблюдателя Люенбергера и предиктора Смита с добавлением корректирующих членов, определяемых простыми выражениями, не требующими сложных вычислений. Регулятор имеет линейную структуру, однако, в корректирующем члене предусмотрен периодический сброс до нуля значений интеграторов. **Основные результаты.** Аналитически показано, что замкнутая система объекта управления с входным запаздыванием и модифицированным предиктором Смита является глобально экспоненциально устойчивой. **Обсуждение.** Предложенный метод управления системами с входным запаздыванием превосходит все известные аналоги по простоте реализации и эффективной работоспособности для неустойчивыых систем. В дальнейших исследованиях представленный подход будет направлен на нелинейные и параметрически неопределенные системы с входным запаздыванием.

Ключевые слова
входное запаздывание, предиктор, управление по выходу, неустойчивые системы

Благодарности
Работа выполнена при финансовой поддержке Российского научного фонда (проект № 25-29-00896).

Ссылка для цитирования: Пыркин А.А., Калинин К.Ю., Чан Х.К. Управление с предиктором по выходу для неустойчивых линейных систем с входным запаздыванием // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1098–1106. doi: 10.17586/2226-1494-2025-25-6-1098-1106

Output predictor control for unstable linear systems with input delay

Anton A. Pyrkin[1]✉, Konstantin Yu. Kalinin[2], Han Cong Tran[3]

[1,2,3] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] pyrkin@itmo.ru✉, https://orcid.org/0000-0001-8806-4057
[2] kgkalinin@gmail.com, https://orcid.org/0009-0009-3092-6254
[3] werkaiaye@gmail.com, https://orcid.org/0009-0007-1116-7876

Abstract
The classical output control problem for a linear system with an input delay and constant known parameters is considered. The plant may be unstable, making most of the known methods ineffective or unconstructive. A new control algorithm based on the Luenberger observer and the Smith predictor is proposed, incorporating correction terms defined by simple expressions that eliminate the need for complex calculations. The resulting regulator has a linear structure; however, the correction term provides for a periodic reset of the corresponding regulator variable. It is analytically proven that a closed system of a plant with an input delay and a modified Smith predictor is globally exponentially stable. The resulting method for controlling systems with an input delay surpasses all analogues known to the authors in terms of simplicity

© Пыркин А.А., Калинин К.Ю., Чан Х.К., 2025

of implementation and effective performance for unstable systems. In future works, this approach will be extended to nonlinear and parametrically uncertain systems with an input delay.

Keywords
input delay, predictor, output control, unstable systems

Acknowledgements
This work was supported by the Russian Foundation grant (project No. 25-29-00896).

For citation: Pyrkin A.A., Kalinin K.Yu., Tran H.C. Output predictor control for unstable linear systems with input delay. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1098–1106 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1098-1106

Введение

Задача управления системами с входным запаздыванием не теряет своей актуальности, и даже приобретает дополнительный интерес с развитием новых методов управления сложными динамическими процессами. Такие методы управления развиваются в областях, где скорость принятия решения имеет критический характер, а небольшие задержки в контуре управления не позволяют достигнуть желаемого результата.

В классической теории управления для класса линейных систем с известными параметрами в работе [1] впервые было показано условие, при каких допустимых значения входного запаздывания замкнутая система может сохранять устойчивость. Для произвольной величины запаздывания, но для класса устойчивых линейных систем был предложен предиктор Смита [2, 3], получивший широкое применение как в теоретических исследованиях, так и на практике. Суть метода заключается в построении дополнительного управляющего контура, который может полностью скомпенсировать негативный эффект входного запаздывания, обеспечивая глобальную устойчивость замкнутой системы, если параметры модели объекта точно известны.

Развитием предиктора Смита стали работа [4], известная как предиктор Манитиуса–Олброта, и его модификации [5, 6], где были получены алгоритмы управления для формирования стабилизирующей обратной связи для неустойчивых систем. Долгое время этот подход был и остается базовым для синтеза предикторов для систем: с неизвестными параметрами, нелинейных систем, систем с распределенными параметрами, моделируемых уравнениями в частных производных [7, 8]. Отметим, что данный результат имеет теоретический характер, а при технической его реализации остаются открытыми вопросы устойчивости замкнутой системы при наличии помех или ошибок при вычислениях, несмотря на большое число научных работ, посвященных развитию подхода [4]. Например, с использованием предиктора Манитиуса–Олброта [4] были решены задачи управления неустойчивыми системами с запаздыванием и параметрически неопределенными возмущениями: для синусоидальных воздействий синтезированы регуляторы по состоянию [9] и по выходу [10], для смещенного синусоидального воздействия, действующего на вход и выход объекта [11] и более общий случай регулятора по выходу для компенсации смещенного мультисинусоидального возмущения, действующего на вход и выход системы с запаздывающим управлением [12]. Однако результаты [9–12] не лишены недостатков, присущих предиктору [4]. Позднее предпринимались попытки улучшить робастные свойства предиктора Манитиуса–Олброта в задачах стабилизации неустойчивых систем с запаздыванием и возмущающими воздействиями [13, 14].

Особое внимание заслуживает работа [15], где была изучена проблема скрытой неустойчивой динамики при использовании предикторов типа [4], синтезирован наблюдатель этой нуль-динамики и получен закон управления, позволяющий в отличие от аналогов робастно стабилизировать неустойчивые объекты с запаздывающим управлением, обеспечивая глобальную асимптотическую устойчивость замкнутой системы. С теоретической точки зрения решена важная и долгое время открытая проблема неустойчивой скрытой динамики в предикторах для неустойчивых систем. Однако в этом алгоритме необходимо выполнять неочевидные для инженера вычисления и переключения, что может вызывать трудности при реализации.

В работе [16] предложен новый алгоритм управления, отличающийся от [15] тем, что он не является модификацией [4–6], а служит модификацией предиктора Смита с добавлением корректирующего члена, позволяющего стабилизировать замкнутую систему, в том числе и скрытую нуль-динамику. Как и в работе [15] в новом регуляторе используется прием сброса значений интегратора, однако, техническая реализация закона управления существенно проще.

В настоящей работе представлено развитие алгоритма [16] для управления по выходу неустойчивыми линейными системами с входным запаздыванием с добавлением наблюдателя Люенбергера. При очень простой структуре регулятора представлено строгое аналитическое доказательство устойчивости замкнутой системы. Разработанный подход в будущем может быть развит и для нелинейных систем, а также для систем с неизвестными параметрами.

Постановка задачи

Рассмотрим линейный объект управления с входным запаздыванием и постоянными параметрами

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t - D), \tag{1}$$

$$y(t) = \mathbf{C}\mathbf{x}(t), \tag{2}$$

где $\mathbf{x}(t) \in R^n$ — неизмеряемый вектор переменных состояния; $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ — полностью управляемая и наблюдаемая тройка матриц с известными элементами; $D \geq 0$ — известное постоянное запаздывание; $y(t) \in R$ — регулируемая и измеряемая выходная переменная.

Требуется синтезировать закон управления $u(t)$ по измерениям выходной переменной $y(t)$, обеспечивающий асимптотическую устойчивость нулевого положения равновесия объекта $\mathbf{x} = 0$ и ограниченность всех переменных состояния замкнутой системы.

Синтез наблюдателя состояния

В задачах управления по выходу, как правило, используются динамические фильтры, которые позволяют оценивать вектор переменных состояния или вектор производных выходной регулируемой переменной до некоторого порядка. В этом случае становится возможным синтез стабилизирующего управления на основе восстановленных оценок. Для объекта управления (1), (2) хорошо известен способ оценивания переменных состояния с помощью наблюдателя Люенбергера вида

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}u(t-D) + \mathbf{L}(y(t) - \mathbf{C}\hat{\mathbf{x}}), \qquad (3)$$

где $\hat{\mathbf{x}}(t)$ — вектор оценок переменных состояния, такой, что невязка $\tilde{\mathbf{x}}(t) = \mathbf{x}(t) - \hat{\mathbf{x}}(t)$ экспоненциально стремится к нулю, если матрица соответствующего размера $\mathbf{L}$ выбрана из условия гурвицевости матрицы $\mathbf{H} = \mathbf{A} - \mathbf{LC}$. Практически для невязки $\tilde{\mathbf{x}}(t)$ справедлива динамическая модель

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{H}\tilde{\mathbf{x}}(t), \qquad (4)$$

откуда следует экспоненциальная устойчивость положения равновесия $\tilde{\mathbf{x}}(t) = 0$. Далее при синтезе закона управления используется вектор оценок переменных состояния $\hat{\mathbf{x}}(t)$.

Синтез закона управления по выходу с предиктором

Следуя подходу модифицированного предиктора Смита [10], выберем закон управления по выходу в виде:

$$u(t) = \mathbf{K}\hat{\mathbf{x}}(t) + \mathbf{K}\boldsymbol{\psi}(t), \qquad (5)$$

$$\dot{\boldsymbol{\psi}}(t) = \mathbf{A}\boldsymbol{\psi}(t) + \mathbf{B}u(t) - \mathbf{B}u(t-D) - \mathbf{BKz}(t) + \mathbf{BK}\boldsymbol{\varepsilon}(t), \qquad (6)$$

с добавлением корректирующих членов:

$$\mathbf{z}(t) = \mathbf{e}^{\mathbf{A}D}[\hat{\mathbf{x}}(t) - \hat{\mathbf{x}}(t-D) - \boldsymbol{\psi}(t-D) + \boldsymbol{\varepsilon}(t-D)], \qquad (7)$$

$$\dot{\boldsymbol{\varepsilon}}(t) = \mathbf{F}\boldsymbol{\varepsilon}(t) - \mathbf{BKz}(t),\ \boldsymbol{\varepsilon}(mT) = 0,\ m = 0, 1, 2, 3, \ldots, \qquad (8)$$

где $\mathbf{K}$ — вектор такой, что матрица $\mathbf{F} = \mathbf{A} + \mathbf{BK}$ гурвицева; $\hat{\mathbf{x}}(t)$ определен в (3), $\boldsymbol{\psi}(t) \in R^n$ — вектор переменных состояния фильтра (6); $\mathbf{z}(t) \in R^n$ — вспомогательная векторная переменная, вычисляемая по формуле (7); $\mathbf{e}^{\mathbf{A}D}$ — матричная экспонента; $\boldsymbol{\varepsilon}(t) \in R^n$ — вектор переменных состояния динамического корректирующего члена со сбросом значений в ноль по правилу (8) с периодом $T > 0$, значение которого выбирается из условия устойчивости замкнутой системы.

Запишем модель замкнутой системы, вычислив производную переменной $\mathbf{z}(t)$:

$$\begin{aligned}
\dot{\mathbf{z}}(t) &= \mathbf{e}^{\mathbf{A}D}[\dot{\hat{\mathbf{x}}}(t) - \dot{\hat{\mathbf{x}}}(t-D) - \dot{\boldsymbol{\psi}}(t-D) + \dot{\boldsymbol{\varepsilon}}(t-D)] = \\
&= \mathbf{e}^{\mathbf{A}D}[\mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}u(t-D) + \mathbf{LC}(\mathbf{x}(t) - \hat{\mathbf{x}}(t))] - \\
&- \mathbf{e}^{\mathbf{A}D}[\mathbf{A}\hat{\mathbf{x}}(t-D) + \mathbf{B}u(t-2D) + \mathbf{LC}(\mathbf{x}(t-D) - \hat{\mathbf{x}}(t-D))] - \\
&- \mathbf{e}^{\mathbf{A}D}[\mathbf{A}\boldsymbol{\psi}(t-D) + \mathbf{B}u(t-D) - \mathbf{B}u(t-2D) - \\
&- \mathbf{BKz}(t-D) + \mathbf{BK}\boldsymbol{\varepsilon}(t-D)] + \\
&+ \mathbf{e}^{\mathbf{A}D}[\mathbf{A}\boldsymbol{\varepsilon}(t-D) + \mathbf{BK}\boldsymbol{\varepsilon}(t-D) - \mathbf{BKz}(t-D)] = \\
&= \mathbf{A}\mathbf{e}^{\mathbf{A}D}[\hat{\mathbf{x}}(t) - \hat{\mathbf{x}}(t-D) - \boldsymbol{\psi}(t-D) + \boldsymbol{\varepsilon}(t-D)] + \\
&+ \mathbf{e}^{\mathbf{A}D}\mathbf{LC}[\tilde{\mathbf{x}}(t) - \tilde{\mathbf{x}}(t-D)] = \\
&= \mathbf{Az}(t) + \mathbf{e}^{\mathbf{A}D}\mathbf{LC}[\tilde{\mathbf{x}}(t) - \tilde{\mathbf{x}}(t-D)]. \qquad (9)
\end{aligned}$$

На основе (5) и (7) получим выражение для запаздывающего управления

$$\begin{aligned}
u(t-D) &= \mathbf{K}\hat{\mathbf{x}}(t-D) + \mathbf{K}\boldsymbol{\psi}(t-D) = \\
&= \mathbf{K}\hat{\mathbf{x}}(t) + \mathbf{K}\boldsymbol{\varepsilon}(t-D) - \mathbf{K}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) = \\
&= \mathbf{Kx}(t) + \mathbf{K}\boldsymbol{\varepsilon}(t-D) - \mathbf{K}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) - \mathbf{K}\tilde{\mathbf{x}}(t)
\end{aligned}$$

и перепишем модель замкнутой системы (1)–(3), (5)–(8) с учетом вычислений (4), (9) в компактном виде:

$$\dot{\mathbf{x}}(t) = \mathbf{Fx}(t) + \mathbf{BK}\boldsymbol{\varepsilon}(t-D) - \mathbf{BK}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) - \mathbf{BK}\tilde{\mathbf{x}}(t), \qquad (10)$$

$$\begin{aligned}
\dot{\boldsymbol{\psi}}(t) &= \mathbf{F}\boldsymbol{\psi}(t) + \mathbf{BK}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) - \mathbf{BKz}(t) - \\
&- \mathbf{BK}\boldsymbol{\varepsilon}(t-D) + \mathbf{BK}\boldsymbol{\varepsilon}(t), \qquad (11)
\end{aligned}$$

$$\dot{\boldsymbol{\varepsilon}}(t) = \mathbf{F}\boldsymbol{\varepsilon}(t) - \mathbf{BKz}(t), \qquad (12)$$

$$\dot{\mathbf{z}}(t) = \mathbf{Az}(t) + \mathbf{e}^{\mathbf{A}D}\mathbf{LC}[\tilde{\mathbf{x}}(t) - \tilde{\mathbf{x}}(t-D)], \qquad (13)$$

$$\dot{\tilde{\mathbf{x}}}(t) = \mathbf{H}\tilde{\mathbf{x}}(t), \qquad (14)$$

где переменные $\mathbf{x}(t)$, $\boldsymbol{\psi}(t)$ и $\tilde{\mathbf{x}}(t)$ являются непрерывными, а $\boldsymbol{\varepsilon}(t)$ и $\mathbf{z}(t)$ меняют скачкообразно свои значения согласно (8) для $\boldsymbol{\varepsilon}(t)$ в моменты времени mT и (7) для $\mathbf{z}(t)$ в моменты времени $mT + D$ соответственно.

Заметим, что подсистема (13) может быть неустойчива в силу свойств матрицы $\mathbf{A}$, которая по постановке задачи не обязательно гурвицева. Если не использовать корректировку переменной $\boldsymbol{\varepsilon}(t)$, то замкнутая система может быть неустойчивой, поскольку на вход устойчивых по входу подсистем (10)–(12) попадает потенциально неограниченный сигнал $\mathbf{z}(t)$. Далее рассмотрим замкнутую систему с принудительным обнулением переменной $\boldsymbol{\varepsilon}(t)$ с периодом T и покажем, при каких условиях сброса переменной $\boldsymbol{\varepsilon}(t)$ замкнутая система является асимптотически устойчивой.

Основной результат

Интегрируя модель (14) для невязки $\tilde{\mathbf{x}}(t)$ получим выражение

$$\tilde{\mathbf{x}}(t) = \mathbf{e}^{\mathbf{H}t}\tilde{\mathbf{x}}(0),$$

где $\tilde{\mathbf{x}}(0) = \mathbf{x}(0) - \hat{\mathbf{x}}(0)$ — некоторый вектор с постоянными элементами, зависящими от начальных условий, а все элементы матричной функции $\mathbf{e}^{\mathbf{H}t}$ экспоненциально стремятся к нулю из-за гурвицевости матрицы $\mathbf{H}$.

Для запаздывающего сигнала $\tilde{\mathbf{x}}(t-D)$ на интервале $0 \le t < D$ имеем

$$\tilde{\mathbf{x}}(t-D) = \mathbf{x}(t-D),$$

что соответствует свободному движению объекта до момента включения регулятора, в том числе и наблюдателя, до момента времени $t = 0$. Предположим, что на этом конечном интервале $-D \le t < 0$ состояние объекта ограничено.

Для $t \ge D$ имеем

$$\tilde{\mathbf{x}}(t-D) = \mathbf{e}^{-\mathbf{H}D}\mathbf{e}^{\mathbf{H}t}\tilde{\mathbf{x}}(0) = \mathbf{e}^{-\mathbf{H}D}\tilde{\mathbf{x}}(t).$$

В силу определения (7), (8) функция $\mathbf{z}(t)$ меняет скачкообразно значение в каждый момент времени $t_m = mT + D$ для $m \in Z_{\ge 0}$, а в соответствии с (13) для переменной $\mathbf{z}(t)$ между переключениями в моменты времени (t_m, t_{m+1}) справедливо выражение

$$\mathbf{z}(t) = \mathbf{e}^{\mathbf{A}(t-t_m)}\mathbf{z}(t_m) + \int_{t_m}^{t} \mathbf{e}^{\mathbf{A}(t-s)}\,\mathbf{e}^{\mathbf{A}D}\mathbf{LC}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(s)ds = \tag{15}$$
$$= \mathbf{e}^{\mathbf{A}(t-t_m)}\mathbf{z}(t_m) + \mathbf{e}^{\mathbf{A}t}\mathbf{e}^{\mathbf{A}D}\left(\int_{t_m}^{t} \mathbf{e}^{-\mathbf{A}s}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds\right)[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0),$$

где $\mathbf{I}$ — единичная матрица порядка n; s — переменная интегрирования.

Для вычисления интеграла в выражении (15) рассмотрим лемму с учетом обозначения $\mathbf{H} = \mathbf{A} - \mathbf{LC}$.

Лемма 1 [16]. Справедливо соотношение для интеграла

$$\int \mathbf{e}^{-\mathbf{A}s}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds = -\mathbf{e}^{-\mathbf{A}s}\mathbf{e}^{\mathbf{H}s} + \mathbf{const}.$$

Доказательство леммы 1.

Дифференцируя функцию $-\mathbf{e}^{-\mathbf{A}s}\mathbf{e}^{\mathbf{H}s}$, получим

$$\frac{d}{dt}(-\mathbf{e}^{-\mathbf{A}s}\mathbf{e}^{\mathbf{H}s}) = \mathbf{e}^{-\mathbf{A}s}\mathbf{A}\mathbf{e}^{\mathbf{H}s} - \mathbf{e}^{-\mathbf{A}s}\mathbf{e}^{\mathbf{H}s}\mathbf{H} =$$
$$= \mathbf{e}^{-\mathbf{A}s}\mathbf{A}\mathbf{e}^{\mathbf{H}s} - \mathbf{e}^{-\mathbf{A}s}(\mathbf{A} - \mathbf{LC})\mathbf{e}^{\mathbf{H}s} = \mathbf{e}^{-\mathbf{A}s}\,\mathbf{LC}\mathbf{e}^{\mathbf{H}s},$$

что соответствует подынтегральному выражению, что и требовалось доказать. ■

Продолжим вычисления в (15), получим выражение для переменной $\mathbf{z}(t)$ как кусочно-непрерывной функции времени между переключениями

$$\mathbf{z}(t) = \mathbf{e}^{\mathbf{A}(t-t_m)}\mathbf{z}(t_m) + \mathbf{e}^{\mathbf{A}t}\mathbf{e}^{\mathbf{A}D}(\mathbf{e}^{-\mathbf{A}t_m}\mathbf{e}^{\mathbf{H}t_m} - \mathbf{e}^{-\mathbf{A}t}\mathbf{e}^{\mathbf{H}t}) \times$$
$$\times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) = \mathbf{e}^{\mathbf{A}(t-t_m)}\mathbf{z}(t_m) + \mathbf{e}^{\mathbf{A}D}(\mathbf{e}^{\mathbf{A}(t-t_m)}\mathbf{e}^{\mathbf{H}t_m} - \mathbf{e}^{\mathbf{H}t}) \times$$
$$\times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0). \tag{16}$$

Исследуем последовательность значений функции $\mathbf{z}(t)$ в моменты времени $t = mT + D$ для $m \in Z_{\ge 0}$

$$\mathbf{z}(mT + D) = \mathbf{e}^{\mathbf{A}D}[\hat{\mathbf{x}}(mT + D) - \hat{\mathbf{x}}(mT) - \boldsymbol{\psi}(mT)], \tag{17}$$

где уже учтено тождество $\boldsymbol{\varepsilon}(t - D) = 0$, $\forall t = mT + D$.

Для этого рассмотрим вспомогательную непрерывную функцию времени

$$\boldsymbol{\xi}(t) = \hat{\mathbf{x}}(t) - \hat{\mathbf{x}}(t-D) - \boldsymbol{\psi}(t-D) = \mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) - \boldsymbol{\varepsilon}(t-D) \tag{18}$$

и ее значения в моменты времени $t = mT + D$ для $m \in Z_{\ge 0}$, связанные с искомой последовательностью $\mathbf{z}(mT + D)$ соотношением $\mathbf{z}(mT + D) = \mathbf{e}^{\mathbf{A}D}\boldsymbol{\xi}(mT + D)$. Продифференцировав (18), запишем:

$$\dot{\boldsymbol{\xi}}(t) = \mathbf{e}^{-\mathbf{A}D}\dot{\mathbf{z}}(t) - \dot{\boldsymbol{\varepsilon}}(t-D) =$$
$$= \mathbf{A}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) + \mathbf{LC}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\mathbf{e}^{\mathbf{H}t}\tilde{\mathbf{x}}(0) - \mathbf{F}\boldsymbol{\varepsilon}(t-D) +$$
$$+ \mathbf{BKz}(t-D) = \mathbf{A}\boldsymbol{\xi}(t) + \mathbf{BKz}(t-D) - \mathbf{BK}\boldsymbol{\varepsilon}(t-D) +$$
$$+ \mathbf{LC}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\mathbf{e}^{\mathbf{H}t}\tilde{\mathbf{x}}(0) = \mathbf{A}\boldsymbol{\xi}(t) + \mathbf{BKz}(t-D) -$$
$$- \mathbf{BK}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) + \mathbf{BK}\boldsymbol{\xi}(t) + \mathbf{LC}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\mathbf{e}^{\mathbf{H}t}\tilde{\mathbf{x}}(0) =$$
$$= \mathbf{F}\boldsymbol{\xi}(t) - \mathbf{BK}\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) + \mathbf{BKz}(t-D) +$$
$$+ \mathbf{LC}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\mathbf{e}^{\mathbf{H}t}\tilde{\mathbf{x}}(0). \tag{19}$$

Проинтегрировав (19), получим

$$\boldsymbol{\xi}(t) = \mathbf{e}^{\mathbf{F}t}\boldsymbol{\xi}(0) - \int_0^t \mathbf{e}^{\mathbf{F}(t-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds +$$
$$+ \int_0^t \mathbf{e}^{\mathbf{F}(t-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0). \tag{20}$$

На основе (20) запишем выражение для последовательности $\boldsymbol{\xi}_m = \boldsymbol{\xi}(mT + D)$:

$$\boldsymbol{\xi}_m = \mathbf{e}^{\mathbf{F}(mT+D)}\boldsymbol{\xi}(0) -$$
$$- \int_0^{mT+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds +$$
$$+ \int_0^{mT+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0). \tag{21}$$

Вычислим значение последовательности (21)

$$\boldsymbol{\xi}_{m+1} = \mathbf{e}^{\mathbf{F}(mT+T+D)}\boldsymbol{\xi}(0) -$$
$$- \int_0^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+T+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds +$$
$$+ \int_0^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+T+D-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0).$$

Подставим выражение $\mathbf{e}^{\mathbf{F}(mT+D)}\boldsymbol{\xi}(0)$ из (21):

$$\boldsymbol{\xi}_{m+1} = \mathbf{e}^{\mathbf{F}T}\boldsymbol{\xi}_m + \mathbf{e}^{\mathbf{F}T}\int_0^{mT+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds -$$
$$- \mathbf{e}^{\mathbf{F}T}\int_0^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds -$$
$$- \mathbf{e}^{\mathbf{F}T}\int_0^{mT+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) +$$
$$+ \mathbf{e}^{\mathbf{F}T}\int_0^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) =$$
$$= \mathbf{e}^{\mathbf{F}T}\boldsymbol{\xi}_m - \mathbf{e}^{\mathbf{F}T}\int_{mT+D}^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds +$$
$$+ \mathbf{e}^{\mathbf{F}T}\int_{mT+D}^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) =$$
$$= \mathbf{e}^{\mathbf{F}T}\boldsymbol{\xi}_m - \mathbf{e}^{\mathbf{F}T}\int_{mT+D}^{mT+2D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds -$$
$$- \mathbf{e}^{\mathbf{F}T}\int_{mT+2D}^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(s) - \mathbf{z}(s-D))ds +$$
$$+ \mathbf{e}^{\mathbf{F}T}\int_{mT+D}^{mT+T+D} \mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{LC}\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0). \tag{22}$$

Отметим, что в моменты времени $t = mT + 2D$ скачком меняется значение функции $\mathbf{z}(t - D)$, поэтому со-

ответствующий интеграл был разбит на два интервала $[mT + D; mT + 2D)$ и $[mT + 2D; mT + D + T)$. Выполним преобразования для первых двух интегралов в (22).

В моменты времени $[mT + D; mT + D + T)$ функция $\mathbf{z}(t)$ имеет вид

$$\mathbf{z}(t) = \mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{z}(mT + D) + \\ + \mathbf{e}^{\mathbf{A}D}(\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{H}t})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) = \\ = \mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{z}(mT + D) + \\ + (\mathbf{e}^{\mathbf{A}(t-mT)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0).$$

Для первого интервала $[mT + D; mT + 2D)$ значение функции $\mathbf{z}(t - D)$ равно

$$\mathbf{z}(t - D) = \mathbf{e}^{\mathbf{A}(t-(m-1)T-2D)}\mathbf{z}((m - 1)T + D) + \\ + (\mathbf{e}^{\mathbf{A}(t-(m-1)T-D)}\mathbf{e}^{\mathbf{H}((m-1)T+D)} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t-\mathbf{H}D})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0),$$

что соответствует непрерывному изменению с момента коррекции $\boldsymbol{\varepsilon}(t)$ на шаге $m - 1$. Тогда для первого интервала имеем

$$\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) - \mathbf{z}(t - D) = \mathbf{e}^{-\mathbf{A}D}\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{z}(mT + D) - \\ - \mathbf{e}^{\mathbf{A}(t-(m-1)T-2D)}\mathbf{z}((m - 1)T + D) + \\ + \mathbf{e}^{\mathbf{A}D}(\mathbf{e}^{\mathbf{A}(t-mT)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) - \\ - (\mathbf{e}^{\mathbf{A}(t-(m-1)T-D)}\,\mathbf{e}^{\mathbf{H}((m-1)T+D)} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t-\mathbf{H}D})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) = \\ = \mathbf{e}^{\mathbf{A}(t-mT-2D)}[\mathbf{z}(t - mT + D) - \mathbf{e}^{\mathbf{A}T}\mathbf{z}((m - 1)T + D)] + \\ + (\mathbf{e}^{-\mathbf{A}D}\mathbf{e}^{\mathbf{A}(t-mT)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{A}(t-(m-1)T-D)}\,\mathbf{e}^{\mathbf{H}((m-1)T+D)}) \times \\ \times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) - (\mathbf{e}^{\mathbf{H}t} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t-\mathbf{H}D})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) = \\ = \mathbf{e}^{\mathbf{A}(t-mT-2D)}[\mathbf{z}(mT + D) - \mathbf{e}^{\mathbf{A}T}\mathbf{z}((m - 1)T + D)] + \\ + \mathbf{e}^{\mathbf{A}t}(\mathbf{e}^{-\mathbf{A}(mT+D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{-\mathbf{A}((m-1)T+D)}\mathbf{e}^{\mathbf{H}((m-1)T+D)}) \times \\ \times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) + (\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} - \mathbf{I})\mathbf{e}^{\mathbf{H}t}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0). \quad (23)$$

В момент времени $mT + 2D$ функция $\mathbf{z}(t - D)$ меняет скачком свое значение, что соответствует шагу m. На втором интервале $[mT + 2D; mT + D + T)$ получим

$$\mathbf{z}(t - D) = \mathbf{e}^{\mathbf{A}(t-mT-2D)}\mathbf{z}(mT + D) + \\ + (\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t-\mathbf{H}D})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0)$$

и

$$\mathbf{e}^{-\mathbf{A}D}\mathbf{z}(t) - \mathbf{z}(t - D) = \mathbf{e}^{-\mathbf{A}D}\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{z}(mT + D) - \\ - \mathbf{e}^{\mathbf{A}(t-mT-2D)}\mathbf{z}(mT + D) + \\ + (\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{H}t})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) - \\ - (\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t-\mathbf{H}D})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) = \\ = (\mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{\mathbf{A}(t-mT-D)}\mathbf{e}^{\mathbf{H}(mT+D)}) \times \\ \times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) + (\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{\mathbf{H}t-\mathbf{H}D} - \mathbf{e}^{\mathbf{H}t})[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) = \\ = (\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} - \mathbf{I})\mathbf{e}^{\mathbf{H}t}[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0). \quad (24)$$

Подставляя (23) и (24) в соответствующие подынтегральные выражения в (22), запишем

$$\xi_{m+1} = \mathbf{e}^{\mathbf{F}T}\xi_m - \mathbf{e}^{\mathbf{F}T}\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BKe}^{\mathbf{A}(s-mT-2D)}ds \times \\ \times [\mathbf{z}(mT + D) - \mathbf{e}^{\mathbf{A}T}\mathbf{z}((m - 1)T + D)] - \\ - \mathbf{e}^{\mathbf{F}T}\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BKe}^{\mathbf{A}s}ds \times \\ \times (\mathbf{e}^{-\mathbf{A}(mT+D)}\,\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{-\mathbf{A}((m-1)T+D)}\mathbf{e}^{\mathbf{H}((m-1)T+D)}) \times \\ \times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) - \mathbf{e}^{\mathbf{F}T}\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{BK}(\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} - \mathbf{I})\mathbf{e}^{\mathbf{H}s}ds \times \\ \times [\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) - \mathbf{e}^{\mathbf{F}T}\int\limits_{mT+2D}^{mT+T+D}\mathbf{e}^{\mathbf{F}(mT+D-s)} \times \\ \times \mathbf{BK}(\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} - \mathbf{I})\mathbf{e}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0) + \\ + \mathbf{e}^{\mathbf{F}T}\int\limits_{mT+D}^{mT+T+D}\mathbf{e}^{\mathbf{F}(mT+D-s)}\mathbf{LCe}^{\mathbf{H}s}ds[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0).$$

Заметим, что на основании выражений (17) и (18), справедливы соотношения

$$\mathbf{z}(mT + D) = \mathbf{e}^{\mathbf{A}D}\xi_m,\ \mathbf{z}(mT + D - T) = \mathbf{e}^{\mathbf{A}D}\xi_{m-1},$$

тогда, объединяя подобные члены и вынося за пределы интеграла выражения, не зависящие от переменной интегрирования s, получим выражение

$$\xi_{m+1} = \mathbf{e}^{\mathbf{F}T}\xi_m - \mathbf{e}^{\mathbf{F}(mT+T+D)}\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{-\mathbf{F}s} \times \\ \times \mathbf{BKe}^{\mathbf{A}s}ds\mathbf{e}^{-\mathbf{A}(mT+D)}[\xi_m - \mathbf{e}^{\mathbf{A}T}\xi_{m-1}] + \mathbf{g}_m, \quad (25)$$

где введено обозначение $\mathbf{g}_m$ для следующей последовательности

$$\mathbf{g}_m = \mathbf{Q}_m[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0), \quad (26)$$

зависящей от матрицы

$$\mathbf{Q}_m = -\mathbf{e}^{\mathbf{F}(mT+T+D)}\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{-\mathbf{F}s}\mathbf{BKe}^{\mathbf{A}s}ds \times \\ \times (\mathbf{e}^{-\mathbf{A}(mT+D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{-\mathbf{A}((m-1)T+D)}\mathbf{e}^{\mathbf{H}((m-1)T+D)}) + \\ + \mathbf{e}^{\mathbf{F}(mT+T+D)} \times \\ \times \int\limits_{mT+D}^{mT+T+D}\mathbf{e}^{-\mathbf{F}s}(\mathbf{BK} - \mathbf{BKe}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} + \mathbf{LC})\mathbf{e}^{\mathbf{H}s}ds. \quad (27)$$

Лемма 2. Значения матрицы $\mathbf{Q}_m$ равномерно убывают к нулю с ростом m.

Доказательство леммы 2.

Воспользуемся результатом леммы 1 и запишем выражения для интегралов в (25) и (27):

$$\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{-\mathbf{F}s}\mathbf{BKe}^{\mathbf{A}s}ds = \\ = \mathbf{e}^{-\mathbf{F}(mT+D)}\mathbf{e}^{\mathbf{A}(mT+D)} - \mathbf{e}^{-\mathbf{F}(mT+2D)}\mathbf{e}^{\mathbf{A}(mT+2D)}. \quad (28)$$

Тогда для первого слагаемого в (27), получим аналитическое выражение

$$\mathbf{e}^{\mathbf{F}(mT+T+D)}\int\limits_{mT+D}^{mT+2D}\mathbf{e}^{-\mathbf{F}s}\mathbf{BKe}^{\mathbf{A}s}ds \times \\ \times (\mathbf{e}^{-\mathbf{A}(mT+D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{-\mathbf{A}((m-1)T+D)}\mathbf{e}^{\mathbf{H}((m-1)T+D)}) = \\ = \mathbf{e}^{\mathbf{F}(mT+T+D)}(\mathbf{e}^{-\mathbf{F}(mT+D)}\mathbf{e}^{\mathbf{A}(mT+D)} - \mathbf{e}^{-\mathbf{F}(mT+2D)}\mathbf{e}^{\mathbf{A}(mT+2D)}) \times \\ \times (\mathbf{e}^{-\mathbf{A}(mT+D)}\mathbf{e}^{\mathbf{H}(mT+D)} - \mathbf{e}^{-\mathbf{A}((m-1)T+D)}\mathbf{e}^{\mathbf{H}((m-1)T+D)}) = \\ = \mathbf{e}^{\mathbf{F}T}(\mathbf{I} - \mathbf{e}^{-\mathbf{F}D}\mathbf{e}^{\mathbf{A}D})\mathbf{e}^{\mathbf{H}(mT+D)} - \\ - \mathbf{e}^{\mathbf{F}T}(\mathbf{I} - \mathbf{e}^{-\mathbf{F}D}\mathbf{e}^{\mathbf{A}D})\mathbf{e}^{\mathbf{A}T}\mathbf{e}^{\mathbf{H}((m-1)T+D)}) = \\ = \mathbf{e}^{\mathbf{F}T}(\mathbf{I} - \mathbf{e}^{-\mathbf{F}D}\mathbf{e}^{\mathbf{A}D})(\mathbf{I} - \mathbf{e}^{\mathbf{A}T}\mathbf{e}^{-\mathbf{H}T})\mathbf{e}^{\mathbf{H}\,(mT+D)}. \quad (29)$$

На основе выражения (27) с использованием (29) запишем выражение для матрицы $\mathbf{Q}_{m+1}$

$$\mathbf{Q}_m = -\mathbf{e}^{\mathbf{F}T}(\mathbf{I} - \mathbf{e}^{-\mathbf{F}D}\mathbf{e}^{\mathbf{A}D})(\mathbf{I} - \mathbf{e}^{\mathbf{A}T}\mathbf{e}^{-\mathbf{H}T})\mathbf{e}^{\mathbf{H}((m+1)T+D)} + \\ + \mathbf{e}^{\mathbf{F}((m+1)T+T+D)}\int\limits_{(m+1)T+D}^{(m+1)T+T+D}\mathbf{e}^{-\mathbf{F}s} \times \\ \times (\mathbf{BK} - \mathbf{BK}\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} + \mathbf{LC})\mathbf{e}^{\mathbf{H}s}ds. \qquad (30)$$

Меняя в (30) переменную интегрирования $s = \bar{s} + T$, получим

$$\mathbf{Q}_{m+1} = -\mathbf{e}^{\mathbf{F}T}(\mathbf{I} - \mathbf{e}^{-\mathbf{F}D}\mathbf{e}^{\mathbf{A}D})(\mathbf{I} - \mathbf{e}^{\mathbf{A}T}\mathbf{e}^{-\mathbf{H}T})\mathbf{e}^{\mathbf{H}(mT+D)}\,\mathbf{e}^{\mathbf{H}T} + \\ + \mathbf{e}^{\mathbf{F}((m+1)T+T+D)}\int\limits_{mT+D}^{mT+T+D}\mathbf{e}^{-\mathbf{F}(\bar{s}+T)} \times \\ \times (\mathbf{BK} - \mathbf{BK}\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} + \mathbf{LC})\mathbf{e}^{\mathbf{H}(\bar{s}+T)}d\bar{s} = \\ = -\mathbf{e}^{\mathbf{F}T}(\mathbf{I} - \mathbf{e}^{-\mathbf{F}D}\mathbf{e}^{\mathbf{A}D})(\mathbf{I} - \mathbf{e}^{\mathbf{A}T}\mathbf{e}^{-\mathbf{H}T})\mathbf{e}^{\mathbf{H}(mT+D)}\mathbf{e}^{\mathbf{H}T} + \\ + \mathbf{e}^{\mathbf{F}(mT+T+D)}\int\limits_{mT+D}^{mT+T+D}\mathbf{e}^{-\mathbf{F}\bar{s}} \times \\ \times (\mathbf{BK} - \mathbf{BK}\mathbf{e}^{\mathbf{A}D}\mathbf{e}^{-\mathbf{H}D} + \mathbf{LC})\mathbf{e}^{\mathbf{H}\bar{s}}ds\mathbf{e}^{\mathbf{H}T} = \mathbf{Q}(m)\mathbf{e}^{\mathbf{H}T}.$$

В итоге, справедливо соотношение

$$\|\mathbf{Q}_{m+1}\| = \alpha_1\|\mathbf{Q}_m\|,$$

где $\alpha_1 = \|\mathbf{e}^{\mathbf{H}T}\| < 1$, $\forall T > 0$ из-за гурвицевости $\mathbf{H}$, что и требовалось доказать. ■

Таким образом, выражение (25) может быть представлено в виде:

$$\xi_{m+1} + \mathbf{G}_1\xi_m + \mathbf{G}_2\xi_{m-1} = \mathbf{g}_m, \qquad (31)$$

где матричные коэффициенты

$$\mathbf{G}_1 = -\mathbf{e}^{\mathbf{F}T} + \mathbf{e}^{\mathbf{F}(mT+T+D)}(\mathbf{e}^{-\mathbf{F}(mT+D)}\mathbf{e}^{\mathbf{A}(mT+D)} - \\ - \mathbf{e}^{-\mathbf{F}(mT+2D)}\mathbf{e}^{\mathbf{A}(mT+2D)})\mathbf{e}^{-\mathbf{A}(mT+D)} = \\ = -\mathbf{e}^{\mathbf{F}T} + \mathbf{e}^{\mathbf{F}T} - \mathbf{e}^{\mathbf{F}(T-D)}\mathbf{e}^{\mathbf{A}D} = -\mathbf{e}^{\mathbf{F}(T-D)}\mathbf{e}^{\mathbf{A}D},$$

$$\mathbf{G}_2 = -\mathbf{e}^{\mathbf{F}(mT+T+D)}(\mathbf{e}^{-\mathbf{F}(mT+D)}\mathbf{e}^{\mathbf{A}(mT+D)} - \\ - \mathbf{e}^{-\mathbf{F}(mT+2D)}\mathbf{e}^{\mathbf{A}(mT+2D)})\mathbf{e}^{-\mathbf{A}(mT+D)}\mathbf{e}^{\mathbf{A}T} = \\ = \mathbf{e}^{\mathbf{F}(T-D)}\mathbf{e}^{\mathbf{A}(T+D)} - \mathbf{e}^{\mathbf{F}T}\mathbf{e}^{\mathbf{A}T}$$

и последовательность $\mathbf{g}_m$, экспоненциально убывающая к нулю в силу (26) и леммы 2.

Теорема. Существует $T_0 > D$ такое, что для $\forall T \geq T_0$ последовательность ξ_m экспоненциально сходится к $\mathbf{0}$.

Доказательство теоремы.

Перепишем (31) в блочном матричном виде:

$$\begin{pmatrix}\xi_m \\ \xi_{m+1}\end{pmatrix} = \begin{pmatrix}\mathbf{0} & \mathbf{I} \\ -\mathbf{G}_2 & -\mathbf{G}_1\end{pmatrix}\begin{pmatrix}\xi_{m-1} \\ \xi_m\end{pmatrix} + \begin{pmatrix}\mathbf{0} \\ \mathbf{I}\end{pmatrix}\mathbf{g}_m.$$

Введем в рассмотрение вектор $\chi(m) = \mathrm{col}(\xi_{m-1}, \xi_m)$, матрицу $\mathbf{P} = \begin{pmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{I}\end{pmatrix}$ и функцию Ляпунова

$$V(m) = V_1(m) + V_2(m),$$

где $V_1(m) = \chi^{\mathrm{T}}(m)\mathbf{P}\chi(m)$ и $V_2(m) = \mathbf{g}_m^{\mathrm{T}}\mathbf{g}_m$. Рассмотрим сначала $V_1(m)$.

$$V_1(m) - V_1(m+1) = \chi^{\mathrm{T}}(m)\mathbf{P}\chi(m) - \chi^{\mathrm{T}}(m+1)\mathbf{P}\chi(m+1) = \\ = \chi^{\mathrm{T}}(m)\left[\begin{pmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{I}\end{pmatrix} - \begin{pmatrix}\mathbf{0} & -\mathbf{G}_2^{\mathrm{T}} \\ \mathbf{I} & -\mathbf{G}_1^{\mathrm{T}}\end{pmatrix}\begin{pmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & 2\mathbf{I}\end{pmatrix}\begin{pmatrix}\mathbf{0} & \mathbf{I} \\ -\mathbf{G}_2 & -\mathbf{G}_1\end{pmatrix}\right]\chi(m) = \\ = \chi^{\mathrm{T}}(m)\left[\begin{pmatrix}\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\end{pmatrix} - 2\begin{pmatrix}\mathbf{G}_2^{\mathrm{T}}\mathbf{G}_2 & \mathbf{G}_2^{\mathrm{T}}\mathbf{G}_1 \\ \mathbf{G}_1^{\mathrm{T}}\mathbf{G}_2 & \mathbf{G}_1^{\mathrm{T}}\mathbf{G}_1\end{pmatrix}\right]\chi(m).$$

Поскольку матрица $\mathbf{F} = \mathbf{A} + \mathbf{BK}$ гурвицева, исходя из соответствующего выбора $\mathbf{K}$, то

$$\lim_{T\to\infty}\mathbf{e}^{\mathbf{F}T} = 0.$$

Более того, коэффициенты $\mathbf{K}$ могут быть выбраны так, чтобы выполнялось

$$\mathbf{Re}\{\lambda_{\max}\{\mathbf{F}\} + \lambda_{\max}\{\mathbf{A}\}\} < 0,$$

где $\lambda_{\max}\{(\cdot)\}$ означает максимальное собственное число матрицы $(\cdot)$, что будет гарантировать затухание функции $\mathbf{e}^{\mathbf{F}T}$ быстрее возможного роста функции $\mathbf{e}^{\mathbf{A}T}$ для произвольной матрицы $\mathbf{A}$, обеспечивая

$$\lim_{T\to\infty}\mathbf{e}^{\mathbf{F}T}\mathbf{e}^{\mathbf{A}T} = 0.$$

Поскольку функция

$$\mathbf{N}(T) = 2\begin{pmatrix}\mathbf{G}_2^{\mathrm{T}}(T)\mathbf{G}_2(T) & \mathbf{G}_2^{\mathrm{T}}(T)\mathbf{G}_1(T) \\ \mathbf{G}_1^{\mathrm{T}}(T)\mathbf{G}_2(T) & \mathbf{G}_1^{\mathrm{T}}(T)\mathbf{G}_1(T)\end{pmatrix}$$

является непрерывной, ограниченной для $\forall T \geq 0$, а все ее элементы с ростом T стремятся к нулю $\lim\limits_{T\to\infty}\mathbf{N}(T) = 0$, то существует $T_0 > D$ такое, что $\forall T \geq T_0$ справедливо неравенство

$$\|\mathbf{N}(T)\| \leq \alpha_2\mathbf{I},\ 0 < \alpha_2 < 1.$$

Следовательно, при $\forall T \geq T_0$

$$V_1(m) - V_1(m+1) = \chi^{\mathrm{T}}(m)[\mathbf{I}_{2n} - \mathbf{N}(T)]\chi(m) > \\ > (1 - \alpha_2)\chi^{\mathrm{T}}(m)\chi(m) \geq \frac{1 - \alpha_2}{2}V_1(m),$$

откуда получим

$$V_1(m+1) < \frac{1 + \alpha_2}{2}V_1(m).$$

Для $V_2(m)$ можно заметить

$$V_2(m) = \mathbf{g}_m^{\mathrm{T}}\mathbf{g}_m = \|\mathbf{g}_m^{\mathrm{T}}\mathbf{g}_m\| = \|\tilde{\mathbf{x}}(0)[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\|\cdot\|\mathbf{Q}_m\|^2$$

и, соответственно,

$$V_2(m+1) = \|\tilde{\mathbf{x}}(0)[\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\|\cdot\|\mathbf{Q}_{m+1}\|^2 = \alpha_1^2 V_2(m).$$

Таким образом, при условии $T \geq T_0$

$$V(m+1) < \beta V(m),$$

откуда следует равномерная экспоненциальная сходимость последовательности $V(m)$ к нулю с коэффи-

циентом $\beta = \max\left\{\alpha_1^2;\ \dfrac{1+\alpha_2}{2}\right\} < 1$ и экспоненциальная сходимость последовательности ξ_m, что и требовалось доказать. ■

Напомним, что последовательность ξ_m представляет собой значения непрерывной функции $\xi(t)$ в моменты времени $t = mT + D$, и она сходится к нулю. Также заметим, что функция $\mathbf{z}(t)$ в моменты времени $t = mT + D$ связана со сходящейся последовательностью ξ_m соотношением $\mathbf{z}(mT + D) = \mathbf{e}^{\mathbf{A}D}\xi_m$, а на интервале $[mT + D;\ mT + D + T)$ изменяется согласно выражению (16). Максимальное значение $\mathbf{z}(t)$ на интервале между переключениями $[mT + D;\ mT + D + T)$ можно оценить:

$$\|\mathbf{z}(t)\| \le \max_{0\le s<T}\|\mathbf{e}^{\mathbf{A}s}\|\cdot\|\mathbf{z}(mT+D)\| +$$
$$+ \|\mathbf{e}^{\mathbf{A}D}\|\cdot\|\mathbf{I} - \mathbf{e}^{-\mathbf{H}D}]\tilde{\mathbf{x}}(0)\|\cdot \max_{0\le s<T}\|\mathbf{e}^{\mathbf{A}s} - \mathbf{e}^{\mathbf{H}s}\|\cdot\|\mathbf{e}^{\mathbf{H}(mT+D)}\|.$$

Так как последовательности $\mathbf{z}(mT + D)$ и $\mathbf{e}^{\mathbf{H}(mT+D)}$ экспоненциально стремятся к нулю с увеличением m, то можно показать экспоненциальную сходимость к нулю всех элементов вектора $\mathbf{z}(t)$, указав соответствующую мажоранту.

Утверждение. Объект управления с запаздывающим управлением (1), (2) и регулятором на основе наблюдателя Люенбергера (3) и предиктора Смита (5), (6) с корректирующими членами (7), (8) обеспечивает глобальную экспоненциальную устойчивость нулевого положения равновесия замкнутой системы.

Доказательство утверждения. Замкнутая система описывается моделью (10)–(14) с переключениями сигналов $\boldsymbol{\varepsilon}(t)$ по правилу (8) и сигнала $\mathbf{z}(t)$ в соответствии с (17). Все элементы вектора $\mathbf{z}(t)$ экспоненциально сходятся к нулю, поскольку норма этого вектора ограничена мажорирующей затухающей экспонентой. Видно, что сигнал $\boldsymbol{\varepsilon}(t)$ также сходится экспоненциально

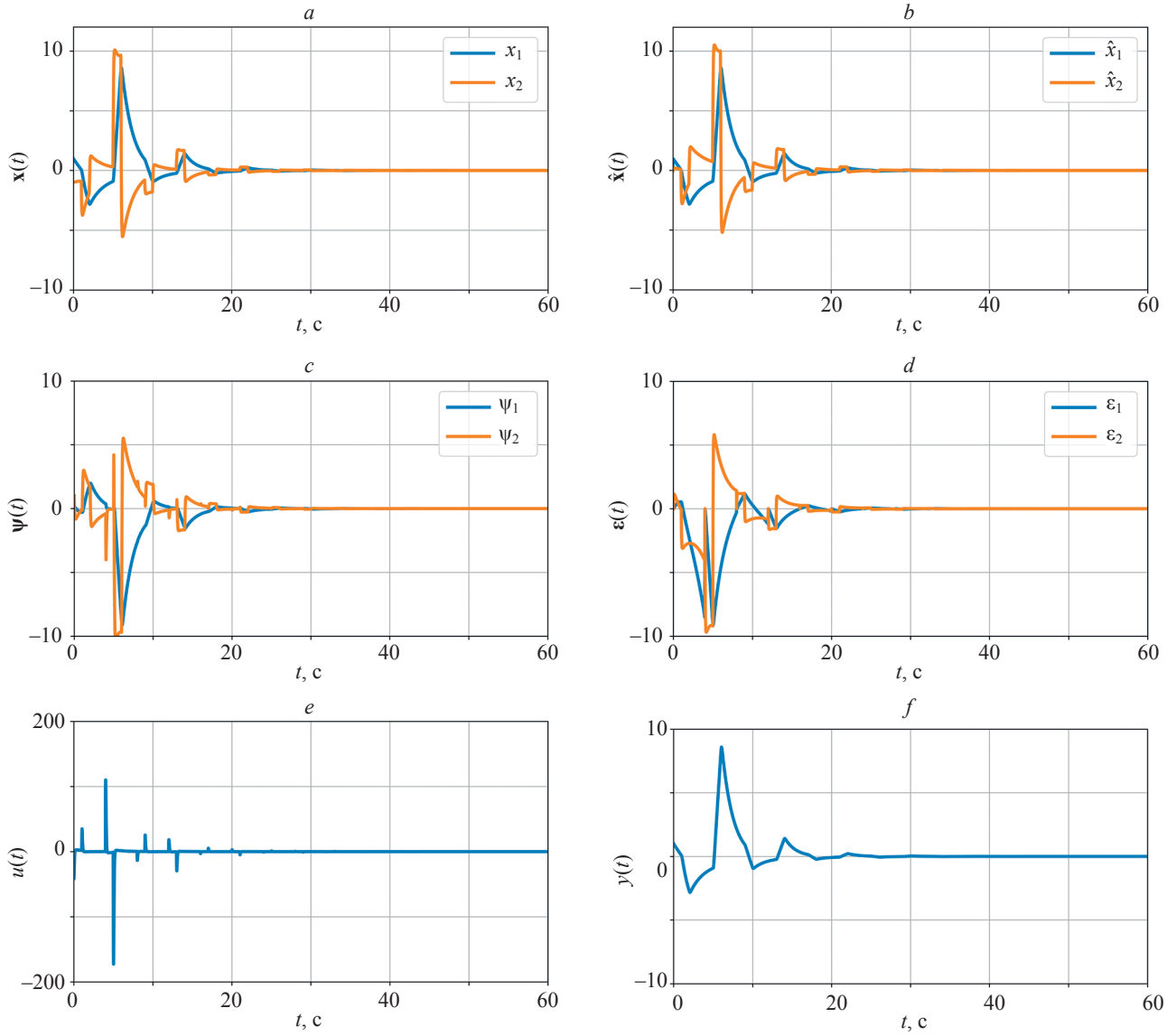

Рисунок. Результаты моделирования замкнутой системы: переменные $\mathbf{x}$ (*a*) и оценка $\hat{\mathbf{x}}$ (*b*) состояний; переменные регулятора $\boldsymbol{\psi}$ (*c*) и $\boldsymbol{\varepsilon}$ (*d*); сигнал управления u (*e*); выходная переменная y (*f*)

Figure. Results of closed-loop system simulation: state variables $\mathbf{x}$ (*a*); state estimates $\hat{\mathbf{x}}$ (*b*); regulator variables $\boldsymbol{\psi}$ (*c*) and $\boldsymbol{\varepsilon}$ (*d*); control signal u (*e*) and output variable y (*f*)

к нулю, учитывая модель (12). Аналогично можно показать ограниченность и экспоненциальную сходимость к нулю переменных состояния моделей (10) и (11), откуда следует экспоненциальная устойчивость нулевого положения равновесия замкнутой системы в целом. ■

Пример численного моделирования

Рассмотрим неустойчивый объект управления (1), (2) с параметрами $\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0{,}2 & 0 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $\mathbf{C} = [1 \quad 0]$, $\mathbf{x}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ и регулятор (3), (5)–(8) с параметрами $\mathbf{K} = [-20 \quad -30]$, $\mathbf{L} = \begin{bmatrix} 100 \\ 20 \end{bmatrix}$ и нулевыми начальными условиями $\hat{\mathbf{x}}(0)$, $\boldsymbol{\psi}(0)$, $\boldsymbol{\varepsilon}(0)$. На рисунке представлены результаты моделирования для запаздывания $D = 1$ и интервала сброса интегратора $T = 5$ с использованием метода интегрирования **ode1** (**Euler**) с фиксированным шагом 10^{-2} с.

Заключение

Предложено решение задачи управления линейной системой по выходу на основе предиктора для неустойчивых систем с входным запаздыванием, отличающееся более простой структурой в реализации и строгим аналитическим доказательством устойчивости. Разработанный подход открывает широкие возможности для дальнейшего развития теории автоматического управления для систем с запаздыванием, включая адаптивные и нелинейные системы.

Литература

1. Цыпкин Я.З. Устойчивость систем с запаздывающей обратной связью // Автоматика и телемеханика. 1946. Т. 7. № 2-3. С. 107–129.
2. Smith O.J.M. Closer control of loops with dead time // Chemical Engineering Progress. 1957. V. 53. N 5. P. 217–219.
3. Smith O.J.M. A controller to overcome dead time // ISA Journal. 1959. V. 6. P. 28–33.
4. Manitius A.Z., Olbrot A.W. Finite spectrum assignment problem for systems with delays // IEEE Transactions on Automatic Control. 1979. V. 24. N 4. P. 541–552. https://doi.org/10.1109/tac.1979.1102124
5. Kwon W.H., Pearson A.E. Feedback stabilization of linear systems with delayed control // IEEE Transactions on Automatic Control. 1980. V. 25. N 2. P. 266–269. https://doi.org/10.1109/tac.1980.1102288
6. Arstein Z. Linear systems with delayed controls: A reduction // IEEE Transactions on Automatic Control. 1982. V. 27. N 4. P. 869–879. https://doi.org/10.1109/tac.1982.1103023
7. Krstic M., Smyshlyaev A. Backstepping boundary control for first-order hyperbolic PDEs and application to systems with actuator and sensor delays // Systems and Control Letters. 2008. V. 57. N 9. P. 750–758. https://doi.org/10.1016/j.sysconle.2008.02.005
8. Kristic M. Delay Compensation for Nonlinear, Adaptive, and PDE Systems. Birkhäuser Boston, 2009. 466 p.
9. Pyrkin A., Smyshlyaev A., Bekiaris-Liberis N., Krstic M. Rejection of sinusoidal disturbance of unknown frequency for linear system with input delay // Proc. of the 2010 American Control Conference. 2010. P. 5688–5693. https://doi.org/10.1109/acc.2010.5531131
10. Pyrkin A., Smyshlyaev A., Bekiaris-Liberis N., Krstic M. Output control algorithm for unstable plant with input delay and cancellation of unknown biased harmonic disturbance // IFAC Proceedings Volumes. 2010. V. 43. N 2. P. 39–44. https://doi.org/10.3182/20100607-3-cz-4010.00009
11. Pyrkin A.A., Bobtsov A.A. Adaptive controller for linear system with input delay and output disturbance // IEEE Transactions on Automatic Control. 2016. V. 61. N 12. P. 4229–4234. https://doi.org/10.1109/tac.2015.2509428
12. Пыркин А.А., Бобцов А.А., Никифоров В.О., Колюбин С.А., Ведяков А.А., Борисов О.И., Громов В.С. Компенсация полигармонического возмущения, действующего на состояние и выход линейного объекта с запаздыванием в канале управления // Автоматика и телемеханика. 2015. № 12. С. 43–64.
13. Furtat I., Fridman E., Fradkov A. Disturbance compensation with finite spectrum assignment for plants with input delay // IEEE Transactions on Automatic Control. 2018. V. 63. N 1. P. 298–305. https://doi.org/10.1109/tac.2017.2732279
14. Furtat I., Gushchin P. Tracking control algorithms for plants with input time-delays based on state and disturbance predictors and sub-predictors // Journal of the Franklin Institute. 2019. V. 356. N 8. P. 4496–4512. https://doi.org/10.1016/j.jfranklin.2019.03.013

References

1. Tcypkin Ya.S. Stability of systems with retarding ferd-back. *Avtomatika i telemehanika*, 1946, vol. 7, no. 2-3, pp. 107–129. (in Russian)
2. Smith O.J.M. Closer control of loops with dead time. *Chemical Engineering Progress*, 1957, vol. 53, no. 5, pp. 217–219.
3. Smith O.J.M. A controller to overcome dead time. *ISA Journal*, 1959, vol. 6, pp. 28–33.
4. Manitius A.Z., Olbrot A.W. Finite spectrum assignment problem for systems with delays. *IEEE Transactions on Automatic Control*, 1979, vol. 24, no. 4, pp. 541–552. https://doi.org/10.1109/tac.1979.1102124
5. Kwon W.H., Pearson A.E. Feedback stabilization of linear systems with delayed control. *IEEE Transactions on Automatic Control*, 1980, vol. 25, no. 2, pp. 266–269. https://doi.org/10.1109/tac.1980.1102288
6. Arstein Z. Linear systems with delayed controls: A reduction. *IEEE Transactions on Automatic Control*, 1982, vol. 27, no. 4, pp. 869–879. https://doi.org/10.1109/tac.1982.1103023
7. Krstic M., Smyshlyaev A. Backstepping boundary control for first-order hyperbolic PDEs and application to systems with actuator and sensor delays. *Systems and Control Letters*, 2008, vol. 57, no. 9, pp. 750–758. https://doi.org/10.1016/j.sysconle.2008.02.005
8. Kristic M. *Delay Compensation for Nonlinear, Adaptive, and PDE Systems*. Birkhäuser Boston, 2009, 466 p.
9. Pyrkin A., Smyshlyaev A., Bekiaris-Liberis N., Krstic M. Rejection of sinusoidal disturbance of unknown frequency for linear system with input delay. *Proc. of the 2010 American Control Conference*, 2010, pp. 5688–5693. https://doi.org/10.1109/acc.2010.5531131
10. Pyrkin A., Smyshlyaev A., Bekiaris-Liberis N., Krstic M. Output control algorithm for unstable plant with input delay and cancellation of unknown biased harmonic disturbance. *IFAC Proceedings Volumes*, 2010, vol. 43, no. 2, pp. 39–44. https://doi.org/10.3182/20100607-3-cz-4010.00009
11. Pyrkin A.A., Bobtsov A.A. Adaptive controller for linear system with input delay and output disturbance. *IEEE Transactions on Automatic Control*, 2016, vol. 61, no. 12, pp. 4229–4234. https://doi.org/10.1109/tac.2015.2509428
12. Pyrkin A.A., Bobtsov A.A., Nikiforov V.O., Kolyubin S.A., Vedyakov A.A., Borisov O.I., Gromov V.S. Compensation of polyharmonic disturbance of state and output of a linear plant with delay in the control channel. *Automation and Remote Control*, 2015, vol. 76, no. 12, pp. 2124–2142. https://doi.org/10.1134/s0005117915120036
13. Furtat I., Fridman E., Fradkov A. Disturbance compensation with finite spectrum assignment for plants with input delay. *IEEE Transactions on Automatic Control,* 2018, vol. 63, no. 1, pp. 298–305. https://doi.org/10.1109/tac.2017.2732279
14. Furtat I., Gushchin P. Tracking control algorithms for plants with input time-delays based on state and disturbance predictors and sub-predictors. *Journal of the Franklin Institute*, 2019, vol. 356, no. 8. pp. 4496–4512. https://doi.org/10.1016/j.jfranklin.2019.03.013
15. Nikiforov V.O., Gerasimov D.N. Robust closed-loop state predictor for unstable systems with input delay. *Proc. of the 62nd IEEE*

15. Nikiforov V.O., Gerasimov D.N. Robust closed-loop state predictor for unstable systems with input delay // Proc. of the 62nd IEEE Conference on Decision and Control (CDC). 2023. P. 5708–5713. https://doi.org/10.1109/cdc49753.2023.10383221
16. Пыркин А.А., Калинин К.Ю. Модифицированный предиктор Смита для неустойчивых линейных систем // Известия высших учебных заведений. Приборостроение. 2025. Т. 68. № 9. С. 753–761. https://doi.org/10.17586/0021-3454-2025-68-9-753-761

Conference on Decision and Control (CDC), 2023, pp. 5708–5713. https://doi.org/10.1109/cdc49753.2023.10383221
16. Pyrkin A.A., Kalinin K.Yu. Modified Smith predictor for unstable linear systems. *Journal of Instrument Engineering*, 2025, vol. 68, no. 9, pp. 753–761. (in Russian). https://doi.org/10.17586/0021-3454-2025-68-9-753-761

Авторы

Пыркин Антон Александрович — доктор технических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 26656070700, https://orcid.org/0000-0001-8806-4057, pyrkin@itmo.ru
Калинин Константин Юрьевич – аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0009-3092-6254, kgkalinin@gmail.com
Чан Хан Конг — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0007-1116-7876, werkaiaye@gmail.com

Authors

Anton A. Pyrkin — D.Sc., Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 26656070700, https://orcid.org/0000-0001-8806-4057, pyrkin@itmo.ru

Konstantin Yu. Kalinin — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0009-3092-6254, kgkalinin@gmail.com
Han Cong Tran — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0007-1116-7876, werkaiaye@gmail.com

Статья поступила в редакцию 27.09.2025
Одобрена после рецензирования 12.10.2025
Принята к печати 22.11.2025

Received 27.09.2025
Approved after reviewing 12.10.2025
Accepted 22.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ
ноябрь–декабрь 2025 Том 25 № 6 http://ntv.ifmo.ru/
SCIENTIFIC AND TECHNICAL JOURNAL OF INFORMATION TECHNOLOGIES, MECHANICS AND OPTICS
November–December 2025 Vol. 25 No 6 http://ntv.ifmo.ru/en/
ISSN 2226-1494 (print) ISSN 2500-0373 (online)

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

COMPUTER SCIENCE

doi: 10.17586/2226-1494-2025-25-6-1107-1116
УДК 004.934.2

Многомодальный подход к определению депрессии с использованием полуавтоматической разметки данных и детерминированных методов машинного обучения

Алёна Николаевна Величко[1], Алексей Анатольевич Карпов[2]✉

[1,2] Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация

[1] velichko.a@iias.spb.su, https://orcid.org/0000-0002-8503-8512
[2] karpov@iias.spb.su✉, https://orcid.org/0000-0003-3424-652X

Аннотация
Введение. Исследована актуальная задача автоматического определения психоэмоциональных состояний человека. Научный интерес к исследованиям автоматического многомодального определения депрессии объясняется распространением тревожно-депрессивных расстройств и повышенной, в связи с этим, нагрузкой на первичное звено здравоохранения. Специфичность задачи обусловлена ее комплексностью, недостаточным объемом или неточностью исходных данных, дисбалансом классов. Сравнительные исследования показывают, что результаты классификации с полуавтоматической разметкой данных часто выше, чем при ее полной автоматизации. **Метод.** Предложенный многомодальный подход к определению депрессии сочетает полуавтоматическую разметку данных и детерминированные методы машинного обучения с использованием нескольких наборов признаков. Для обучения моделей применен многомодальный корпус Extended Distress Analysis Interview Corpus (E-DAIC), содержащий аудиозаписи, автоматически полученные из этих аудиозаписей тексты и вычисленные из видеозаписей видеопризнаки, а также аннотации с результатами тестов Patient Health Questionnaire-8 для каждой записи. Полуавтоматическая разметка позволяет получать точные временные метки и тексты высказываний, для снижения зашумленности данных, используемых при обучении моделей. В предложенном подходе применяется несколько наборов признаков, извлеченных из трех модальностей (акустические экспертные признаки eGeMAPS и нейросетевые акустические признаки DenseNet, визуальные экспертные признаки OpenFace и текстовые признаки Word2Vec). Их комплексная обработка минимизирует влияние дисбаланса классов в данных на результат классификации. **Основные результаты.** Экспериментальные исследования с использованием преимущественно экспертных признаков (DenseNet, OpenFace, Word2Vec) и детерминированных методов классификации (Catboost), обладающих свойством интерпретируемости результатов, на корпусе E-DAIC позволили получить значения показателей, сопоставимые с современными международными исследованиями (68,0 % и 64,3 % по показателям взвешенной средней F1-меры (Weighted F1-measure) и невзвешенной средней полноты (Unweighted Average Recall) соответственно). **Обсуждение.** Применение полуавтоматического подхода к разметке данных и объединение модальностей позволило улучшить качество разметки и распознавание депрессии по сравнению с одномодальными подходами. Достигнуты более сбалансированные результаты классификации. Применение детерминированных методов классификации на основе деревьев решений позволит в дальнейшем провести анализ результатов классификации за счет интерпретируемости их решений. Для этой цели можно также задействовать другие методы интерпретации, например, SHapley Additive exPlanations и Local Interpretable Model-agnostic Explanations.

Ключевые слова
компьютерная паралингвистика, многомодальное распознавание депрессии, деструктивные явления, разметка данных, машинное обучение

Благодарности
Работа выполнена при финансовой поддержке Российского научного фонда (проект № 25-11-00319).

Ссылка для цитирования: Величко А.Н., Карпов А.А. Многомодальный подход к определению депрессии с использованием полуавтоматической разметки данных и детерминированных методов машинного обучения // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1107–1116. doi: 10.17586/2226-1494-2025-25-6-1107-1116

© Величко А.Н., Карпов А.А., 2025

A multimodal approach for depression detection using semi-automatic data annotation and deterministic machine learning methods

Alena N. Velichko[1], Alexey A. Karpov[2]✉

[1,2] St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

[1] velichko.a@iias.spb.su, https://orcid.org/0000-0002-8503-8512
[2] karpov@iias.spb.su✉, https://orcid.org/0000-0003-3424-652X

Abstract
A trending task of automatic psycho-emotional human state detection was studied in this work. A scientific interest to researches devoted to the automatic multimodal depression detection can arise out of the widespread of anxiety-depressive disorders and difficulties of their detection in primary health care. A specificity of the task was caused by its complexity, lack of data, imbalance of classes and inaccuracies in it. Comparative researches show that classification results on semi-automatic annotated data are higher than ones on automatic-annotated data. The proposed approach for depression detection combines a semi-automatic data annotation and deterministic machine learning methods with the utilization of several feature sets. To build our models, we utilized the multimodal Extended Distress Analysis Interview Corpus (E-DAIC) which consists of audio recordings, automatically extracted from these audio recordings texts and video feature sets extracted from video recordings as well as annotation including Patient Health Questionnaire (PHQ-8) scale for each recording. A semi-automatic annotation makes it possible to get the exact time stamps and speech texts to reduce the noisiness in the training data. In the proposed approach we use several feature sets, extracted from each modality (acoustic expert feature set eGeMAPS and neural acoustic feature set DenseNet, visual expert feature set OpenFace and text feature set Word2Vec). A complex processing of these features minimizes the effect of class imbalance in the data on classification results. Experimental researches with the use of mostly expert features (DenseNet, OpenFace, Word2Vec) and deterministic machine learning classification methods (Catboost) which have the property of interpretability of classification results yielded the experimental results on the E-DAIC corpus which are comparable with the existing ones in the field (68.0 % and 64.3 % for Weighted F1-measure (WF1) and Unweighted Average Recall (UAR) accordingly). The usage of a semi-automatic annotation approach and modalities fusion improved both quality of annotation and depression detection comparing to the unimodal approaches. More balanced classification results are achieved. The usage of deterministic machine learning classification methods based on decision trees allows us to provide an interpretability analysis of the classification results in the future due to their interpretability feature. Other methods of results interpretation like SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) also can be used for this purpose.

Keywords
computational paralinguistics, multimodal depression detection, destructive behavior, data annotation, machine learning

Acknowledgements
The work is carried out within the framework of the Russian Science Foundation (project No. 25-11-00319).

Fot citation: Velichko A.N., Karpov A.A. A multimodal approach for depression detection using semi-automatic data annotation and deterministic machine learning methods. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1107–1116 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1107-1116

Введение

В последние годы в области компьютерной паралингвистики наблюдается возросший интерес к исследованиям психоэмоциональных состояний человека по его речи. Также отмечена необходимость автоматического определения речевых проявлений тревожно-депрессивных расстройств заболеваний, что является следствием как повсеместного распространения данного вида расстройств, так и повышенной нагрузки на первичное звено здравоохранения [1]. Большое депрессивное расстройство сопровождается рядом симптомов, значительно снижающих качество жизнедеятельности, а при отсутствии необходимого и своевременного лечения, может стать причиной инвалидности [2]. Кроме того, депрессия может являться как самостоятельным диагнозом, так и иметь корреляцию с другими состояниями, например болезнью Альцгеймера, деменцией [3, 4], тревожностью [5] и аффективными расстройствами (например, маниакальное состояние) [6].

На данный момент задача автоматического определения депрессии является одной из актуальных задач в мировом научном сообществе. Так, например, она была представлена на международных соревнованиях по аудиовизуальному определению эмоционального состояния человека Audio-Visual Emotional Challenge (AVEC) AVEC'19 [7]. Среди наиболее популярных и крупных корпусов, содержащих речь людей с депрессией, можно выделить Distress Analysis Interview Corpus Wizard-of-Oz (DAIC-WOZ) и Extended Distress Analysis Interview Corpus (E-DAIC) [8], AVEC'13 и AVEC'14, а также CONVERGE [9], Chinese Multimodal Depression Corpus (CDMC) [10], и другие. Ввиду специфичности задачи, многие существующие корпуса являются небольшими (относительно корпусов, использующихся для других задач) и в основном не сбалансированы по классам, однако в последние годы многие группы ученых собирают собственные корпусы (например, работы [11–13]). Среди последних работ преимущественно превалируют научные работы, анализирующие отдельные модальности (аудио (например [6, 11, 13, 14]), текст [3, 4, 15–17]), или работы, в которых рассматриваются многомодальные подходы (аудио, видео, текст) [12, 17–19]. Данная тенденция, вероятно, связана с тем, что в существующих корпусах видеомодальность отсутствует или представлена не исходными данными (виде-

озаписями), а визуальными признаками, вычисленными из видеозаписей (как, например, в корпусах DAIC-WOZ и E-DAIC), что объясняется как повышенной сложностью записи качественных видео, так и намерениями сохранить анонимность участников записи.

За последние годы среди методов машинного обучения для задачи определения депрессии по текстовым данным набирают популярность большие языковые модели (Llama-2, GPT-3.5, GPT-4) [18–20]. Также актуальными являются различные варианты модели Bidirectional Encoder Representations from Transformers (BERT: BERT, LongBERT, RoBERTa, DepRoBERTa), например, в работах [20–22]. Среди машинных классификаторов используются как нейросетевые подходы (Convolutional Neural Networks (CNN), Gated Recurrent Unit (GRU) и Long Short-Term Memory (LSTM)), например, в работах [23–25], так и детерминированные методы на основе деревьев решений (XGBoost, Random Forest) [5, 12, 18] и векторов (X-vectors, Support Vector Machine (SVM), Support Vector Regression (SVR)) [4, 5, 11, 12–14, 18, 21].

Существует также ряд отечественных исследований по определению депрессии, в основном по текстовой модальности. Имеются различные текстовые корпусы, как, например, RusNeuroPsych [26], корпус профилей социальной сети ВКонтакте [27] и корпус эссе [28], на основе которых проводятся как исследования с использованием современных языковых моделей [29, 30], так и исследования информативных параметров поведения, а также текстовых маркеров депрессии [31].

Вопрос интерпретируемости результатов классификации поднимался в работе [16], где был применен метод LIME, и [5, 12], если учесть использование методов машинного обучения с высокой степенью интерпретируемости (детерминированные методы на основе деревьев решений). Стоит отметить, что критерий интерпретируемости является одним из важнейших для внедрения и практического использования систем автоматического определения психоэмоциональных состояний человека, в том числе депрессии. В большинстве перечисленных работ корпуса DAIC-WOZ и E-DAIC были использованы в том виде, в котором они представлены авторами, несмотря на ряд неточностей в аннотации этих корпусов, которые потенциально могут повлиять на качество обучения машинных классификаторов. Настоящая работа опирается на работу [25], в которой корпус E-DAIC был предобработан с применением полуавтоматической экспертной разметки.

Описание предлагаемого подхода

Для распознавания депрессии (двух классов: наличие/отсутствие) предложен многомодальный подход, представленный на рисунке. Он является логическим продолжением опубликованной работы [25], описывающей процесс полуавтоматической разметки и экспериментальные исследования подхода для аудиомодальности, и использует результаты, полученные в ней, для предобработки данных других модальностей. Разработанный подход основан на комбинировании признаков, извлеченных из трех модальностей (акустические экспертные признаки eGeMAPS и нейросетевые акустические признаки DenseNet [32], визуальные экспертные признаки OpenFace [33] и текстовые признаки Word2Vec [34]), а также их комплексной обработке с целью минимизации влияния дисбаланса классов в данных на результат классификации.

В качестве экспериментальных данных в настоящей работе используется многомодальный корпус E-DAIC [8]. Данный корпус является расширенной версией корпуса DAIC-WOZ, и содержит 275 интервью, представленных в виде аудиозаписей, автоматически полученных из этих аудиозаписей текстов и видеопризнаков, вычисленных из видеозаписей, а также аннотации, где представлены результаты тестов PHQ-8 для каждой записи. Ввиду того, что в корпусе был обнаружен ряд неточностей, которые потенциально могут повлиять на качество распознавания депрессии, была проведена полуавтоматическая экспертная разметка данных корпуса, которая включала: использование программного инструментария Faster-Whisper[1] для автоматического распознавания англоязычной речи, ряда скриптов и программного инструментария Praat [35] для постобработки и проверки полученных орфографических транскрипций, выравнивание частоты дискретизации и количества каналов в аудиозаписях по всему корпусу, а также проверку бинарных значений в аннотации корпуса. Результатом полуавтоматической разметки является не только корректная разметка классов, но также временные метки и соответствующий им произнесенный текст. Кроме того, при последующих экспериментальных исследованиях по распознаванию депрессии на аудиоданных было выявлено, что, несмотря на понижение средних значений показателей качества классификации, был сокращен разрыв значений показателей для каждого класса за счет повышения качества распознавания депрессии [25, 36].

Видеоданные в обоих корпусах DAIC-WOZ и E-DAIC представлены в виде признаков, вычисленных из сырых видеозаписей: Visual Geometry Group (VGG), ResNet, OpenFace, мешок видео слов (Bag of Video Words, BOVW). Имеющиеся видеопризнаки, представленные в корпусе [25] в покадровом виде (2048 признаков для ResNet, 4096 — для VGG и 49 — для OpenFace с 30 кадрами/с во всех трех случаях), были прорежены (применен каждый 6-й кадр) с целью уменьшения размерности матрицы признаков и разделены с использованием скриптов, в которых использовались временные метки, полученные ранее при разметке. Таким образом, данные каждой из трех модальностей были предобработаны и представлены в виде наборов признаков согласно временным меткам, полученным ранее при разметке.

Полученный в процессе разметки произнесенный текст использован в качестве данных текстовой модальности. Он был нормализован с применением скриптов путем приведения к нижнему регистру, удаления пун-

[1] Программный инструментарий Faster-Whisper [Электронный ресурс]. Режим доступа: https://github.com/SYSTRAN/faster-whisper, свободный. Яз. англ. (дата обращения: 22.08.2025).

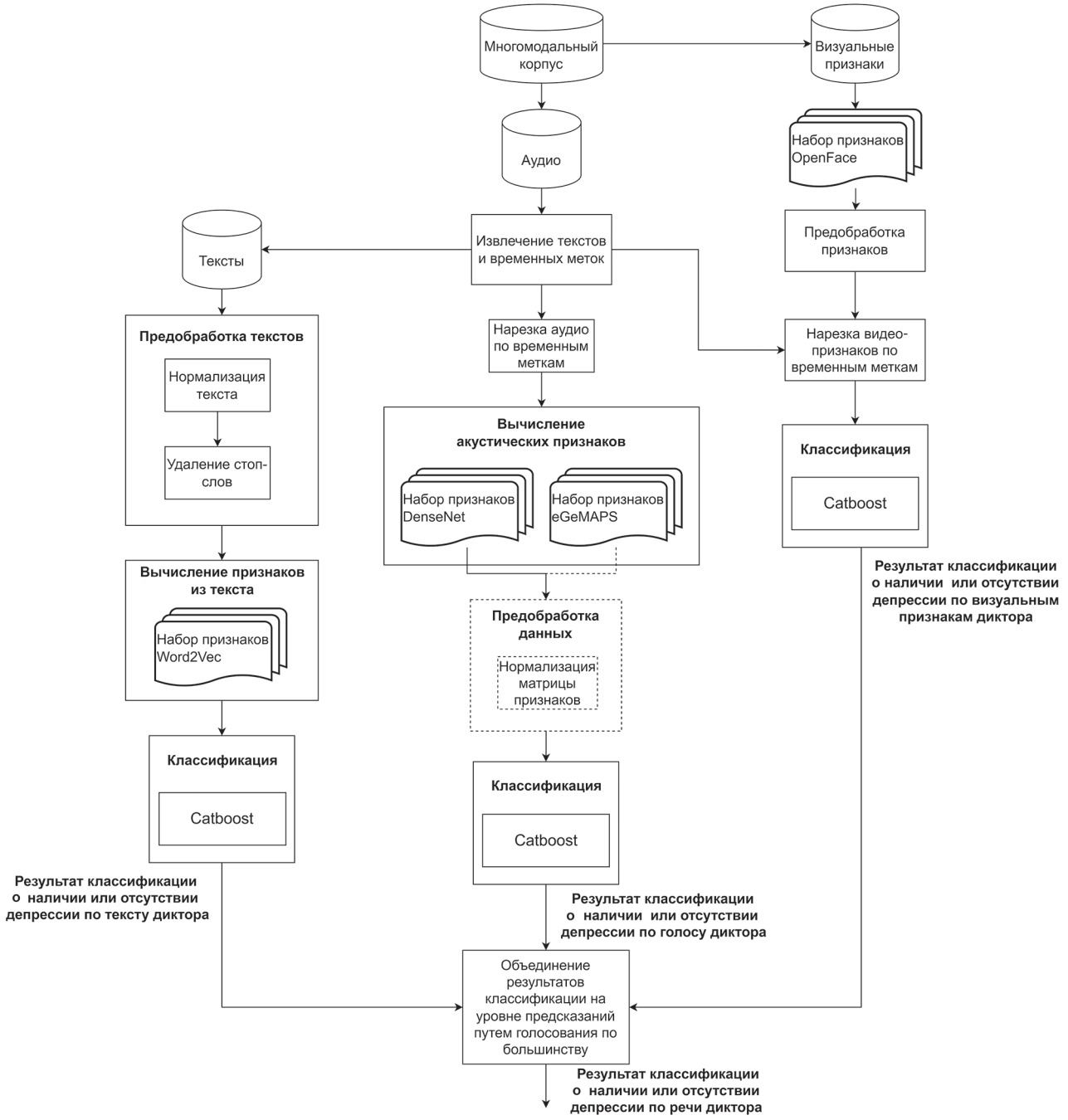

Рисунок. Схематичное изображение подхода к распознаванию депрессии для двух классов (наличие/отсутствие) с использованием многомодальных данных

Figure. A scheme of an approach for depression detection for 2 classes (presence/absence) using multimodal data

ктуации и стоп-слов, после чего с помощью методов векторизации Word2Vec и Term Frequency — Inverse Document Frequency (TF-IDF) вычислены векторные представления, которые являются входными данными классификатора.

Экспериментальное исследование предложенного подхода

Экспериментальные исследования были проведены с использованием разбиения данных, представленных в работе [25] (train, dev, test), для возможности последующего сравнения с аналогичными работами на данном корпусе. Результаты распознавания депрессии по видеоданным приведены в табл. 1, по текстовым данным — в табл. 2. В качестве показателей были выбраны: невзвешенная средняя полнота (Unweighted Average Recall, UAR); полнота для каждого класса (Recall); взвешенная средняя F1-мера (WF1); F-мера для каждого класса. Для бинарной классификации опробованы детерминированные методы машинного обучения: градиентный бустинг на основе деревьев решений (Catboost), метод случайного леса (Random Forest, RF), метод линейного дискриминантного анализа (Linear Discriminant

Таблица 1. Результаты распознавания депрессии на корпусе E-DAIC с использованием видеоданных, %
Table 1. Results of depression detection using E-DAIC corpus data for visual modality, %

Классификатор	UAR	Recall		WF1	F1-мера	
		(0)	(1)		(0)	(1)
VGG						
Catboost	49,1	71,8	27,3	52,8	65,9	31,2
RF	50,1	**93,9**	6,3	52,2	**75,9**	10,1
LDA	**51,4**	76,1	26,6	51,0	68,9	**33,8**
***k*-NN**	50,4	70,5	**30,3**	**54,0**	66,2	**33,8**
ResNet						
Catboost	46,1	60,8	**31,5**	49,5	60,1	**32,1**
RF	50,1	**97,0**	3,2	49,4	**75,9**	6,0
LDA	**53,6**	85,0	22,2	**55,9**	73,2	30,2
k-NN	50,6	83,2	18,0	53,7	71,4	24,7
OpenFace						
Catboost	52,8	56,7	**48,9**	54,2	59,9	**45,2**
RF	**55,6**	96,1	15,0	**56,5**	**76,8**	24,8
LDA	51,2	**97,0**	5,4	49,8	75,4	10,0
k-NN	52,6	66,8	38,4	55,3	64,8	40,3
BOVW						
Catboost	**52,3**	64,7	**39,4**	**55,6**	65,0	**39,7**
RF	50,2	70,4	30,0	54,3	66,7	33,1
LDA	49,1	**84,0**	14,2	52,6	**71,7**	20,1
k-NN	49,5	67,1	32,0	53,4	64,8	34,0

Analysis, LDA) и метод *k*-ближайших соседей (*k*-Nearest Neighbours, *k*-NN). Жирным выделены лучшие результаты, полученные с использованием классификаторов на различных наборах признаков по соответствующим показателям, а также классификаторы, с использованием которых удалось добиться лучших и/или наиболее сбалансированных результатов по каждому набору признаков. В круглых скобках указан номер класса: 0 — отсутствие и 1 — наличие депрессии. Результат многомодального распознавания депрессии был получен путем объединения модальностей посредством голосования по большинству между полученными предсказаниями, где на входе имеются три бинарных значения {0, 1}, полученных от трех классификаторов

Таблица 2. Результаты распознавания депрессии на корпусе E-DAIC с использованием текстовых данных, %
Table 2. Results of depression detection using E-DAIC corpus data for text modality, %

Классификатор	UAR	Recall		WF1	F1-мера	
		(0)	(1)		(0)	(1)
Word2Vec						
Catboost (базовые параметры)	51,4	74,3	28,6	55,3	**68,4**	33,3
RF	50,5	**91,4**	9,5	52,3	74,4	15,4
LDA	52,9	77,1	28,6	56,7	70,1	34,3
k-NN	47,6	85,7	9,5	50,0	71,4	14,3
Catboost (подобранные параметры)	**62,3**	62,9	**61,9**	**63,1**	67,7	**55,3**
TF-IDF						
Catboost (базовые параметры)	**59,5**	**100,0**	19,0	**62,3**	**80,5**	32,0
RF	50,0	**100,0**	0,0	48,0	77,0	0,0
LDA	50,0	**100,0**	0,0	48,0	77,0	0,0
k-NN	48,6	68,6	**28,6**	52,4	64,9	31,6
Catboost (подобранные параметры)	54,3	80,0	**28,6**	58,1	71,8	**35,3**

для каждой модальности, а на выходе единственное бинарное значение.

При исследованиях с видеомодальностью было выявлено, что признаки OpenFace в сочетании с методом Catboost оказались наиболее эффективными. В экспериментальных исследованиях с текстовыми данными метод Catboost с признаками Word2Vec показал лучшие результаты, поскольку обеспечил наиболее сбалансированные результаты распознавания между классами.

Обсуждение результатов

Предложенный подход распознавания депрессии с использованием многомодальных данных, основанный на классификаторе Catboost и нескольких наборах признаков, показал достаточно высокую точность WF1 = 68,0 % (табл. 3), что подтверждает эффективность объединения трех модальностей. Жирным выделены лучшие результаты, полученные при использовании подхода по соответствующим показателям. В круглых скобках указан номер класса: 0 — отсутствие и 1 — наличие депрессии. Полученные результаты сравнения с аналогичными работами показывают, что предложенный подход находится как минимум на мировом уровне, а в ряде случае превосходит аналогичные подходы, при этом являясь более быстрым и менее ресурсозатратным к оборудованию. Предложенный подход оказался эффективнее одномодальных методов для аудио-, видео- и многомодальных методов, представленных в работах [36, 37], однако по ряду показателей оказался менее точным, чем подходы, представленные в других работах, посвященных текстовой модальности и многомодальным подходам.

В ходе исследования были использованы данные многомодального корпуса E-DAIC для бинарного распознавания депрессии, проведено исследование эффективности одно- и многомодальных подходов, а также разработан подход многомодального анализа акустических (DenseNet), визуальных (OpenFace) и текстовых (Word2Vec) признаков. Использование полуавтоматической разметки данных позволило получить менее зашумленные данные. Основное внимание уделено поиску оптимальных параметров и типов признаков, а также сравнению классификаторов. Для классификации в работе были использованы детерминированные методы классификации, такие как Catboost, а итоговое решение принималось посредством голосования по большинству, что минимизировало влияние дисбаланса классов и повысило надежность получаемых результатов. Эксперименты с видеоданными показали, что признаки OpenFace в сочетании с деревьями решений, являются наиболее эффективным решением для видеомодальности. В экспериментах с текстовыми данными метод Catboost с признаками Word2Vec оказался наилучшим, поскольку обеспечил сбалансированные результаты распознавания. Итоговые результаты мно-

Таблица 3. Сравнение показателей эффективности подхода к бинарному распознаванию депрессии с современными подходами на корпусе E-DAIC, %

Table 3. A comparison of the proposed approach efficiency with modern approaches for binary depression detection using E-DAIC corpus data, %

Ссылка на источник подхода	UAR	Recall		WF1	F1-мера	
		(0)	(1)		(0)	(1)
Аудиомодальность						
[37]	52,0	—	—	—	—	—
[38]	41,0	—	—	—	—	—
Настоящая работа	**61,0**	74,3	47,6	63,9	72,2	50,0
Видеомодальность						
[37]	49,0	—	—	—	—	—
[38]	48,0	—	—	—	—	—
Настоящая работа	**52,8**	56,7	48,9	54,2	59,9	45,2
Текстовая модальность						
[17]	**73,9**	**84,6**	**63,3**	73,0	**74,6**	**71,7**
[15]	—	—	—	84,0	—	—
[16]	—	—	—	**86,0**	—	—
Настоящая работа	62,3	62,9	61,9	63,1	67,7	55,3
Многомодальное распознавание						
[39]	**76,9**	83,1	**70,9**	79,5	**84,6**	**68,2**
[37]	58,0	—	—	—	—	—
[38]	54,0	—	—	—	—	—
[23]	—	—	—	**92,2**	—	—
Настоящая работа	64,3	**85,7**	42,9	68,0	77,9	51,4

гомодального подхода составляют WF1 = 68,0 % и UAR = 64,3 % (табл. 3) и сопоставимы с современными международными исследованиями, при этом достигается сбалансированная классификация между классами (при отсутствии и наличии депрессии). К сожалению, не представляется возможным полностью исключить влияние данного фактора внутри одного корпуса, так как дисбаланс классов в корпусе E-DAIC достаточно высок и составляет разницу более, чем в два раза: записей, принадлежащих классу 0 (отсутствие депрессии) — 189, а классу 1 (наличие депрессии) — 86. В качестве возможного решения данной проблемы можно представить межкорпусные исследования, с целью повышения количества записей миноритарного класса. В качестве альтернативного решения представляется возможным опробовать такие техники борьбы с дисбалансом классов как синтетическое увеличение миноритарного класса, например, Synthetic Minority Oversampling Technique, успешно примененного в работе [40] для классификации эмоций в видеоданных. Повышение объяснимости представленного подхода и полученных результатов обуславливается тем, что были использованы методы машинной классификации, основанные на деревьях решений, а также преимущественно экспертные признаки, понятные человеку. Интерпретировать результаты, полученные деревьями решений, можно, например, построив гистограммы на основе вычисленных важностей признаков, поскольку современные методы машинного обучения, реализуемые на базе деревьев решений, имеют встроенную функцию вычисления важности признаков, кроме того, возможен вариант графического отображения дерева решений.

Заключение

В ходе данного исследования был разработан автоматический многомодальный подход к определению депрессии, основанный на экспертных признаках и детерминированных методах машинного обучения. В качестве данных для обучения и тестирования использовался многомодальный корпус Extended Distress Analysis Interview Corpus, данные которого были предобработаны полуавтоматическим путем. Полученные экспериментальные результаты показали, что полуавтоматическая разметка данных позволила не только избавиться от шумов в данных, влияющих на качество обучения машинных классификаторов, но и сократить дисбаланс в результатах распознавания между классами, а также повысить качество распознавания депрессии и получить результаты, сопоставимые с известными аналогами, а в ряде случаев превосходящие их.

В дальнейшем планируется углубленная работа с признаками для каждой модальности. Также будет уделено особое внимание интерпретируемости результатов как при помощи методов машинного обучения, способных предоставить интерпретируемые результаты классификации (что частично было реализовано в настоящей работе), так и методов типа SHapley Additive exPlanations и Local Interpretable Model-agnostic Explanations. Также одним из будущих направлений данной работы является повышение качества распознавания для текстовой модальности и многомодального подхода в том числе с использованием больших языковых моделей.

Системы автоматического многомодального определения депрессии потенциально могут быть использованы специалистами первичного звена здравоохранения в качестве экспертных систем при первичной диагностике тревожно-депрессивных расстройств, а также в ряде комплексных систем, целью которых является автоматическая оценка психоэмоционального состояния человека. Среди ограничений при внедрении предложенного подхода можно выделить: различия в качестве обучающих и данных реального мира (влияние данного фактора может быть уменьшено за счет составления более вариативных наборов обучающих данных, а также добавления методов фильтрации посторонних шумов), относительно низкое качество распознавания (может быть повышено путем добавления вариантов данных, а также комплексирования моделей и признаков). При этом, при внедрении таких систем стоит учитывать, что вне зависимости от качества распознавания, система является вспомогательным инструментом, а итоговое решение должен принимать специалист.

Литература

1. Ушаков И.Б., Бубеев Ю.А., Сыркин Л.Д., Карпов А.А., Поляков А.В., Иванов А.В., Усов В.М. Дистанционное телеконсультирование в первичном звене здравоохранения для скрининга тревожно-депрессивных расстройств с контуром обратной связи от пациента // Системный анализ и управление в биомедицинских системах. 2023. Т. 22. № 4. С. 140–153. https://doi.org/10.36622/VSTU.2023.22.4.022
2. Depressive disorder WHO (depression). WHO. 2023 [Электронный ресурс]. URL: https://www.who.int/news-room/fact-sheets/detail/depression (дата обращения: 22.08.2025)
3. Wu W., Zhang C., Woodland P.C. Confidence estimation for automatic detection of depression and Alzheimer's disease based on clinical interviews // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2024. P. 3160–3164. https://doi.org/10.21437/Interspeech.2024-546
4. Braun F., Bayerl S.P., Perez-Toro P.A., Hoenig F., Lehfeld H., Hillemacher T., Noeth E., Bocklet T., Riedhammer K. Classifying dementia in the presence of depression: a cross-corpus study // Proc.

References

1. Ushakov I.B., Bubeev Yu.A., Syrkin L.D., Karpov A.A., Polyakov A.V., Ivanov A.V., Usov V.M. Remote tele-counceling in primary healthcare for screening of anxiety-depressive disorders with a feedback loop from the patient. *System analysis and management in biomedical systems*, 2023, vol. 22, no. 4, pp. 140–153. (in Russian). https://doi.org/10.36622/VSTU.2023.22.4.022
2. *Depressive disorder WHO (depression)*. WHO. 2023. Available at: https://www.who.int/news-room/fact-sheets/detail/depression (accessed: 22.08.2025)
3. Wu W., Zhang C., Woodland P.C. Confidence estimation for automatic detection of depression and Alzheimer's disease based on clinical interviews. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2024, pp. 3160–3164. https://doi.org/10.21437/Interspeech.2024-546
4. Braun F., Bayerl S.P., Perez-Toro P.A., Hoenig F., Lehfeld H., Hillemacher T., Noeth E., Bocklet T., Riedhammer K. Classifying dementia in the presence of depression: a cross-corpus study. *Proc. of the Annual Conference of the International Speech Communication*

of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 2308–2312. https://doi.org/10.21437/Interspeech.2023-1997
5. Brueckner R., Kwon N., Subramanian V., Blaylock N., O'Connell H. Anxiety and Depression Detection using Vocal Biomarkers. Canaryspeech report. 2025. [Электронный ресурс]. URL: https://canaryspeech.com/blog/anxiety-and-depression-detection-using-vocal-biomarkers/ (дата обращения: 22.08.2025)
6. Ji J., Dong W., Li J., Peng J., Feng C., Liu R., Shi C., Ma Y. Depressive and mania mood state detection through voice as a biomarker using machine learning // Frontiers in Neurology. 2024. V. 15. P. 1394210. https://doi.org/10.3389/fneur.2024.1394210
7. Ringeval F., Schuller B., Valstar M., Cummins N., Cowie R., Tavabi L., et al. AVEC 2019 Workshop and Challenge: state-of-mind, detecting depression with AI, and cross-cultural affect recognition // Proc. of the 9th International on Audio/Visual Emotion Challenge and Workshop. 2019. P. 3–12. https://doi.org/10.1145/3347320.3357688
8. Gratch J., Artstein R., Lucas G., Stratou G., Scherer S., Nazarian A., et al. The Distress Analysis Interview Corpus of Human and Computer Interviews // Proc. of the 9th International Conference on Language Resources and Evaluation (LREC'14). 2014. P. 3123–3128.
9. Li Y., Shi S., Yang F., Gao J., Li Y., Tao M., et al. Patterns of co-morbidity with anxiety disorders in Chinese women with recurrent major depression // Psychological Medicine. 2012. V. 42. N 6. P. 1239–1248. https://doi.org/10.1017/s003329171100273x
10. Zou B., Han J., Wang Y., Liu R., Zhao S., Feng L., Lyu X., Ma H. Semi-structural interview-based Chinese multimodal depression corpus towards automatic preliminary screening of depressive disorders // IEEE Transactions on Affective Computing. 2022. V. 14. N 4. P. 2823–2838. https://doi.org/10.1109/TAFFC.2022.3181210
11. Campbell E.L., Dineley J., Conde P., Matcham F., White K.M., Oetzmann C., et al. The RADAR-CNS Consortium. Classifying depression symptom severity: Assessment of speech representations in personalized and generalized machine learning models // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 1738–1742. https://doi.org/10.21437/Interspeech.2023-1721
12. Fara S., Hickey O., Georgescu A., Goria S., Molimpakis E., Cummins N. Bayesian Networks for the robust and unbiased prediction of depression and its symptoms utilizing speech and multimodal data // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 1728–1732. https://doi.org/10.21437/Interspeech.2023-1709
13. Tao F., Esposito A., Vinciarelli A. The androids corpus: a new publicly available benchmark for speech based depression detection // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 4149–4153. https://doi.org/10.21437/Interspeech.2023-894
14. Phukan O.C., Jain S., Singh S., Singh M., Budaru A.B., Sarma R. ComFeAT: Combination of neural and spectral features for improved depression detection // arXiv. 2024. arXiv:2406.06774. https://doi.org/10.48550/arXiv.2406.06774
15. Burdisso S., Villatoro-Tello E., Madikeri S., Motlicek P. Node-weighted graph convolutional network for depression detection in transcribed clinical interviews // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 3617–3621. https://doi.org/10.21437/interspeech.2023-1923
16. Zhang X., Li C., Chen W., Zheng J., Li F. Optimizing depression detection in clinical doctor-patient interviews using a multi-instance learning framework // Scientific Reports. 2025. V. 15. N 1. P. 6637. https://doi.org/10.1038/s41598-025-90117-w
17. Tank C., Pol S., Katoch V., Meht S., Anand A., Shah R.R. Depression detection and analysis using large language models on textual and audio-visual modalities // arXiv. 2024. arXiv:2407.06125. https://doi.org/10.48550/arXiv.2407.06125
18. Zhang W., Mao K., Chen J. A multimodal approach for detection and assessment of depression using text, audio and video // Phenomics. 2024. V. 4. N 3. P. 234–249. https://doi.org/10.1007/s43657-023-00152-8
19. Zhang X., Liu H., Xu K., Zhang Q., Liu D., Ahmed B., Epps J. When LLMs meet acoustic landmarks: an efficient approach to integrate speech into large language models for depression detection // Proc. of the Conference on Empirical Methods in Natural Language Processing. 2024. P. 146–158. https://doi.org/10.18653/v1/2024.emnlp-main.8

Association Interspeech, 2023, pp. 2308–2312. https://doi.org/10.21437/Interspeech.2023-1997
5. Brueckner R., Kwon N., Subramanian V., Blaylock N., O'Connell H. *Anxiety and Depression Detection using Vocal Biomarkers*. Canaryspeech report. 2025. Available at https://canaryspeech.com/blog/anxiety-and-depression-detection-using-vocal-biomarkers/ (accessed: 22.08.2025)
6. Ji J., Dong W., Li J., Peng J., Feng C., Liu R., Shi C., Ma Y. Depressive and mania mood state detection through voice as a biomarker using machine learning. *Frontiers in Neurology*, 2024, vol. 15, pp. 1394210. https://doi.org/10.3389/fneur.2024.1394210
7. Ringeval F., Schuller B., Valstar M., Cummins N., Cowie R., Tavabi L., et al. AVEC 2019 Workshop and Challenge: state-of-mind, detecting depression with AI, and cross-cultural affect recognition. *Proc. of the 9th International on Audio/Visual Emotion Challenge and Workshop*, 2019, pp. 3–12. https://doi.org/10.1145/3347320.3357688
8. Gratch J., Artstein R., Lucas G., Stratou G., Scherer S., Nazarian A., et al. The Distress Analysis Interview Corpus of Human and *Computer Interviews. Proc. of the 9th International Conference on Language Resources and Evaluation (LREC'14)*, 2014, pp. 3123–3128.
9. Li Y., Shi S., Yang F., Gao J., Li Y., Tao M., et al. Patterns of co-morbidity with anxiety disorders in Chinese women with recurrent major depression. *Psychological Medicine*, 2012, vol. 42, no. 6, pp. 1239–1248. https://doi.org/10.1017/s003329171100273x
10. Zou B., Han J., Wang Y., Liu R., Zhao S., Feng L., Lyu X., Ma H. Semi-structural interview-based Chinese multimodal depression corpus towards automatic preliminary screening of depressive disorders. *IEEE Transactions on Affective Computing*, 2022, vol. 14, no. 4, pp. 2823–2838. https://doi.org/10.1109/TAFFC.2022.3181210
11. Campbell E.L., Dineley J., Conde P., Matcham F., White K.M., Oetzmann C., et al. The RADAR-CNS Consortium. Classifying depression symptom severity: Assessment of speech representations in personalized and generalized machine learning models. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 1738–1742. https://doi.org/10.21437/Interspeech.2023-1721
12. Fara S., Hickey O., Georgescu A., Goria S., Molimpakis E., Cummins N. Bayesian Networks for the robust and unbiased prediction of depression and its symptoms utilizing speech and multimodal data. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 1728–1732. https://doi.org/10.21437/Interspeech.2023-1709
13. Tao F., Esposito A., Vinciarelli A. The androids corpus: a new publicly available benchmark for speech based depression detection. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 4149–4153. https://doi.org/10.21437/Interspeech.2023-894
14. Phukan O.C., Jain S., Singh S., Singh M., Budaru A.B., Sarma R. ComFeAT: Combination of neural and spectral features for improved depression detection. *arXiv*, 2024, arXiv:2406.06774. https://doi.org/10.48550/arXiv.2406.06774
15. Burdisso S., Villatoro-Tello E., Madikeri S., Motlicek P. Node-weighted graph convolutional network for depression detection in transcribed clinical interviews. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 3617–3621. https://doi.org/10.21437/interspeech.2023-1923
16. Zhang X., Li C., Chen W., Zheng J., Li F. Optimizing depression detection in clinical doctor-patient interviews using a multi-instance learning framework. *Scientific Reports*, 2025, vol. 15, no. 1, pp. 6637. https://doi.org/10.1038/s41598-025-90117-w
17. Tank C., Pol S., Katoch V., Meht S., Anand A., Shah R.R. Depression detection and analysis using large language models on textual and audio-visual modalities. *arXiv*, 2024, arXiv:2407.06125. https://doi.org/10.48550/arXiv.2407.06125
18. Zhang W., Mao K., Chen J. A multimodal approach for detection and assessment of depression using text, audio and video. *Phenomics*, 2024, vol. 4, no. 3, pp. 234–249. https://doi.org/10.1007/s43657-023-00152-8
19. Zhang X., Liu H., Xu K., Zhang Q., Liu D., Ahmed B., Epps J. When LLMs meet acoustic landmarks: an efficient approach to integrate speech into large language models for depression detection. *Proc. of the Conference on Empirical Methods in Natural Language Processing*, 2024, pp. 146–158. https://doi.org/10.18653/v1/2024.emnlp-main.8

20. Dumpala S.H., Dikaios K., Nunes A., Rudzicz F., Uher R., Oore S. Self-supervised embeddings for detecting individual symptoms of depression // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2024. P. 1450–1454. https://doi.org/10.21437/Interspeech.2024-2344
21. Sadeghi M., Richer R., Egger B., Schindler-Gmelch L., Rupp L.H., Rahimi F., Berking M., Eskofier B.M. Harnessing multimodal approaches for depression detection using large language models and facial expressions // npj Mental Health Research. 2024. V. 3. N 1. P. 66. https://doi.org/10.1038/s44184-024-00112-8
22. Wang J., Ravi V., Flint J., Alwan A. Speechformer-CTC: Sequential modeling of depression detection with speech temporal classification // Speech Communication. 2024. V. 163. P. 103106. https://doi.org/10.1016/j.specom.2024.103106
23. Jin N., Ye R., Li P. Diagnosis of depression based on facial multimodal data // Frontiers in Psychiatry. 2025. V. 16. P. 1508772. https://doi.org/10.3389/fpsyt.2025.1508772
24. Zhou L., Liu Z., Shangguan Z., Yuan X., Li Y., Hu B. JAMFN: Joint attention multi-scale fusion network for depression detection // Proc. of the Annual Conference of the International Speech Communication Association Interspeech. 2023. P. 3417–3421. https://doi.org/10.21437/Interspeech.2023-183
25. Величко А.Н., Карпов А.А. Подход к распознаванию депрессии по речи человека с использованием полуавтоматической разметки данных // Информационно-управляющие системы. 2024. № 4(131). С. 2–11. https://doi.org/10.31799/1684-8853-2024-4-2-11
26. Litvinova T., Ryzhkova E. RusNeuroPsych: open corpus for study relations between author demographic, personality traits, lateral preferences and affect in text // International journal of Open Information Technologies. 2018. V. 6. N 3. P. 32–36.
27. Stankevich M., Ignatiev N., Smirnov I. Predicting depression with social media images // Proc. of the 9th International Conference on Pattern Recognition Applications and Methods ICPRAM. 2020. V. 1. P. 235–240. https://doi.org/10.5220/0009168602350240
28. Stankevich M.A., Smirnov I.V., Kuznetsova Y.M., Kiselnikova N.V., Enikolopov S.N. Predicting depression from essays in Russian // Proc. of the International Conference “Dialogue 2019”. 2019. P. 647–657.
29. Stankevich M., Smirnov I., Kiselnikova N., Ushakova A. Depression detection from social media profiles // Communications in Computer and Information Science. 2020. V. 1223. P. 181–194. https://doi.org/10.1007/978-3-030-51913-1_12
30. Stepanov D., Smirnov A., Ivanov E., Smirnov I., Stankevich M., Danina M. Detection of health-preserving behavior Among VK.com users based on the analysis of graphic, text and numerical data // Lecture Notes in Networks and Systems. 2022. V. 296. P. 574–587. https://doi.org/10.1007/978-3-030-82199-9_39
31. Кисельникова Н., Станкевич М., Данина М., Куминская Е., Лаврова Е. Выявление информативных параметров поведения пользователей социальной сети ВКонтакте как признаков депрессии // Психология. Журнал Высшей школы экономики. 2020. Т. 17. № 1. С. 73–88. https://doi.org/10.17323/1813-8918-2020-1-73-88
32. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 2261–2269. https://doi.org/10.1109/CVPR.2017.243
33. Baltrusaitis T., Zadeh A., Lim Y.C., Morency L.-P. OpenFace 2.0: facial behavior analysis toolkit // Proc. of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018). 2018. P. 59–66. https://doi.org/10.1109/FG.2018.00019
34. Mikolov T., Sutskever I., Chen K., Corrado G.S., Dean J. Distributed representations of words and phrases and their compositionality // Advances in Neural Information Processing Systems. 2013. V. 26. P. 1–9.
35. Boersma P. Praat, a system for doing phonetics by computer // Glot International. 2001. N 5. P. 341–345.
36. Величко А.Н., Карпов А.А. Методика и программная система интегрального анализа деструктивных паралингвистических явлений в разговорной речи // Информационно-управляющие системы. 2023. № 4 (125). С. 2–11. https://doi.org/10.31799/1684-8853-2023-4-2-11
37. Gimeno-Gómez D., Bucur A.M., Cosma A., Martínez-Hinarejos C.D., Rosso P. Reading between the frames: multi-modal depression detection in videos from non-verbal cues // Lecture Notes in Computer Science, 2024. V. 14608. P. 191–209. https://doi.org/10.1007/978-3-031-56027-9_12

20. Dumpala S.H., Dikaios K., Nunes A., Rudzicz F., Uher R., Oore S. Self-supervised embeddings for detecting individual symptoms of depression. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2024, pp. 1450–1454. https://doi.org/10.21437/Interspeech.2024-2344
21. Sadeghi M., Richer R., Egger B., Schindler-Gmelch L., Rupp L.H., Rahimi F., Berking M., Eskofier B.M. Harnessing multimodal approaches for depression detection using large language models and facial expressions. *npj Mental Health Research*, 2024, vol. 3, no. 1, pp. 66. https://doi.org/10.1038/s44184-024-00112-8
22. Wang J., Ravi V., Flint J., Alwan A. Speechformer-CTC: Sequential modeling of depression detection with speech temporal classification. *Speech Communication*, 2024, vol. 163, pp. 103106. https://doi.org/10.1016/j.specom.2024.103106
23. Jin N., Ye R., Li P. Diagnosis of depression based on facial multimodal data. *Frontiers in Psychiatry*, 2025, vol. 16, pp. 1508772. https://doi.org/10.3389/fpsyt.2025.1508772
24. Zhou L., Liu Z., Shangguan Z., Yuan X., Li Y., Hu B. JAMFN: Joint attention multi-scale fusion network for depression detection. *Proc. of the Annual Conference of the International Speech Communication Association Interspeech*, 2023, pp. 3417–3421. https://doi.org/10.21437/Interspeech.2023-183
25. Velichko A.N., Karpov A.A. AN Approach to depression detection in speech using a semi-automatic data annotation. *Information and Control Systems*, 2024, no. 4 (131), pp. 2–11. (in Russian). https://doi.org/10.31799/1684-8853-2024-4-2-11
26. Litvinova T., Ryzhkova E. RusNeuroPsych: open corpus for study relations between author demographic, personality traits, lateral preferences and affect in text. *International journal of Open Information Technologies*, 2018, vol. 6, no. 3, pp. 32–36.
27. Stankevich M., Ignatiev N., Smirnov I. Predicting depression with social media images. *Proc. of the 9th International Conference on Pattern Recognition Applications and Methods ICPRAM*, 2020, vol. 1, pp. 235–240. https://doi.org/10.5220/0009168602350240
28. Stankevich M.A., Smirnov I.V., Kuznetsova Y.M., Kiselnikova N.V., Enikolopov S.N. Predicting depression from essays in Russian. *Proc. of the International Conference “Dialogue 2019”*, 2019, pp. 647–657.
29. Stankevich M., Smirnov I., Kiselnikova N., Ushakova A. Depression detection from social media profiles. *Communications in Computer and Information Science*, 2020, vol. 1223, pp. 181–194. https://doi.org/10.1007/978-3-030-51913-1_12
30. Stepanov D., Smirnov A., Ivanov E., Smirnov I., Stankevich M., Danina M. Detection of health-preserving behavior Among VK.com users based on the analysis of graphic, text and numerical data. *Lecture Notes in Networks and Systems*, 2022, vol. 296, pp. 574–587. https://doi.org/10.1007/978-3-030-82199-9_39
31. Kiselnikova N., Stankevich M., Danina M., Kuminskaya E., Lavrova E. Identification of informative behavior parameters in users of VKontakte social network as markers of depression. Psychology. *Journal of Higher School of Economics*, 2020, vol. 17, no. 1, pp. 73–88. (in Russian). https://doi.org/10.17323/1813-8918-2020-1-73-88
32. Huang G., Liu Z., Van Der Maaten L., Weinberger K.Q. Densely connected convolutional networks. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269, https://doi.org/10.1109/CVPR.2017.243
33. Baltrusaitis T., Zadeh A., Lim Y.C., Morency L.-P. OpenFace 2.0: facial behavior analysis toolkit. *Proc. of the 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 2018, pp. 59–66. https://doi.org/10.1109/FG.2018.00019
34. Mikolov T., Sutskever I., Chen K., Corrado G.S., Dean J. Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 2013, vol. 26, pp. 1–9.
35. Boersma P. Praat, a system for doing phonetics by computer. *Glot International*, 2001, no. 5, pp. 341–345.
36. Velichko A.N., Karpov A.A. Methods and a software system for integrative analysis of destructive paraiinguistic phenomena in colloquial speech. *Information and Control Systems*, 2023, no. 4 (125), pp. 2–11. (in Russian). https://doi.org/10.31799/1684-8853-2023-4-2-11
37. Gimeno-Gómez D., Bucur A.M., Cosma A., Martínez-Hinarejos C.D., Rosso P. Reading between the frames: multi-modal depression detection in videos from non-verbal cues. *Lecture Notes in Computer Science*, 2024, vol. 14608, pp. 191–209. https://doi.org/10.1007/978-3-031-56027-9_12

38. Jaegle A., Gimeno F., Brock A., Zisserman A., Vinyals O., Carreira J. Perceiver: General perception with iterative attention // Proc. of the 38th International Conference on Machine Learning. 2021. V. 139. P. 4651–4664.
39. Li Y., Yang X., Zhao M., Wang Z., Yao Y., Qian W., Qi Sh. FPT-Former: A flexible parallel transformer of recognizing depression by using audiovisual expert-knowledge-based multimodal measures // International Journal of Intelligent Systems. 2024. V. 1564574. P. 1–13. https://doi.org/10.1155/2024/1564574
40. Рюмина Е.В., Карпов А.А. Сравнительный анализ методов устранения дисбаланса классов эмоций в видеоданных выражений лиц // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 5. С. 683–691. https://doi.org/10.17586/2226-1494-2020-20-5-683-691

38. Jaegle A., Gimeno F., Brock A., Zisserman A., Vinyals O., Carreira J. Perceiver: General perception with iterative attention. *Proc. of the 38th International Conference on Machine Learning*, 2021, vol. 139, pp. 4651–4664.
39. Li Y., Yang X., Zhao M., Wang Z., Yao Y., Qian W., Qi Sh. FPT-Former: A flexible parallel transformer of recognizing depression by using audiovisual expert-knowledge-based multimodal measures. *International Journal of Intelligent Systems*, 2024, vol. 1564574, pp. 1–13. https://doi.org/10.1155/2024/1564574
40. Ryumina E.V., Karpov A.A. Comparative analysis of methods for imbalance elimination of emotion classes in video data of facial expressions. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 5, pp. 683–691. (in Russian). https://doi.org/10.17586/2226-1494-2020-20-5-683-691

Авторы

Величко Алёна Николаевна — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, sc 57203962694, https://orcid.org/0000-0002-8503-8512, velichko.a@iias.spb.su

Карпов Алексей Анатольевич — доктор технических наук, профессор, руководитель лаборатории, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, sc 57219469958, https://orcid.org/0000-0003-3424-652X, karpov@iias.spb.su

Authors

Alena N. Velichko — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, sc 57203962694, https://orcid.org/0000-0002-8503-8512, velichko.a@iias.spb.su

Alexey A. Karpov — D.Sc., Professor, Head of Laboratory, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, sc 57219469958, https://orcid.org/0000-0003-3424-652X, karpov@iias.spb.su

Статья поступила в редакцию 26.08.2025
Одобрена после рецензирования 21.09.2025
Принята к печати 26.11.2025

Received 26.08.2025
Approved after reviewing 21.09.2025
Accepted 26.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1117-1124
УДК 004.05

Комбинированная модель качества рекомендательных систем

Алексей Михайлович Цыплов[1]✉, Александр Валерьевич Бухановский[2]

[1,2] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] tsyplov80@mail.ru✉, https://orcid.org/0009-0009-5211-3019
[2] avbukhanovskii@itmo.ru, https://orcid.org/0000-0003-1588-8164

Аннотация
Введение. Рассмотрены подходы к количественной оценке различных эффектов, таких как позиционный сдвиг (Position Bias), сдвиг в сторону популярных объектов (Popularity Bias) и другие, в рекомендательных системах. Предложена новая модель качества рекомендательных систем, которая приводит выбранные метрики к одной единице измерения и определяет для каждого эффекта его влияние на систему. Полученные оценки позволяют проводить более глубокий сравнительный анализ различных систем, а также исследовать поведение системы на разных сегментах пользователей. **Метод.** Для каждой метрики в рамках предложенной модели строится две условные маргинальные плотности распределения: отдельно на релевантных и нерелевантных рекомендациях. На основе сравнения этих плотностей множество возможных значений метрики разделяется на нормальную и критическую. Модель оценивает влияние каждого эффекта на систему на основе частоты попадания значений соответствующей метрики в свою критическую область. **Основные результаты.** Для демонстрации работы модели проведен анализ четырех алгоритмов построения рекомендаций на академическом наборе данных MovieLens-100K. В ходе тестирования оценивались Popularity Bias, отсутствие новизны в рекомендациях и склонность систем рекомендовать объекты исключительно на основе демографических данных пользователей. Для каждого эффекта построена оценка его влияния на систему, приведен пример прогнозирования верхней оценки качества системы в случае устранения соответствующего эффекта. **Обсуждение.** Показано, что метрики таких эффектов, как Popularity Bias или Position Bias, могут менять распределение абсолютных значений в зависимости от рекомендательной системы. Одним из способов более надежно сравнивать разные рекомендательные системы является предложенная модель качества. Модель подходит для оценивания персональных рекомендаций независимо от сферы применения и алгоритма, который был использован для их построения.

Ключевые слова
рекомендательные системы, ранжирование, оценка качества рекомендаций, Popularity Bias, Position Bias, машинное обучение

Ссылка для цитирования: Цыплов А.М., Бухановский А.В. Комбинированная модель качества рекомендательных систем // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1117–1124. doi: 10.17586/2226-1494-2025-25-6-1117-1124

Compound quality model for recommender system evaluation

Aleksei M. Tsyplov[1]✉, Alexander V. Boukhanovsky[2]

[1,2] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] tsyplov80@mail.ru✉, https://orcid.org/0009-0009-5211-3019
[2] avbukhanovskii@itmo.ru, https://orcid.org/0000-0003-1588-8164

Abstract
The study examines approaches to quantifying various effects, such as Position bias, Popularity Bias, and others, in recommender systems. A new quality model of the recommendation algorithms is proposed which reduces the selected metrics to one unit of measurement and determines its impact on the system for each effect. The obtained scores allow for a deeper comparative analysis of various algorithms as well as investigation the behavior of the algorithm in different

© Цыплов А.М., Бухановский А.В., 2025

user segments. For each metric, two conditional marginal distribution densities are built within the framework of the model: separately based on relevant and irrelevant recommendations. Based on the comparison of these densities, the set of possible metric values is divided into normal and critical. The model evaluates the impact of each effect on the system based on the frequency of hitting the values of the corresponding metric in its critical area. To demonstrate how the model works, four recommendation algorithms were analyzed on the MovieLens-100K academic dataset. During the testing, Popularity Bias, the lack of novelty in recommendations, and the tendency of algorithms to recommend objects solely based on user demographic data were evaluated. For each effect, an assessment of its impact on the system is constructed, and an example of predicting an upper estimate of the system quality is given if the corresponding effect is eliminated. The study demonstrated that metrics of effects such as Popularity or Position Bias can change the distribution of absolute values depending on the system. One of the ways to compare different recommendation algorithms more reliably is the proposed quality model. The model is suitable for evaluating personal recommendations, regardless of the scope of application and the algorithm that was used to build them.

Keywords
recommendation systems, ranking, evaluation of the quality of recommendations, popularity bias, position bias, machine learning

For citation: Tsyplov A.M., Boukhanovsky A.V. Compound quality model for recommender system evaluation. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1117–1124 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1117-1124

Введение

На сегодняшний день область применения рекомендательных систем включает в себя множество разнообразных направлений: медиа-системы, социальные сети, онлайн-торговые площадки и т. д. Несмотря на то, что каждое направление обладает своими уникальными особенностями (разные ключевые показатели, формат контента, паттерны поведения пользователей и т. п.), существует ряд общих эффектов, которым могут быть подвержены рекомендательные системы [1, 2]. К таким эффектам можно отнести позиционный сдвиг (Position Bias), склонность к рекомендации популярных объектов (Popularity Bias), усиление цикла обратной связи (Feedback Loop) и другие [3].

Position Bias — эффект, при котором на вероятность взаимодействия пользователя с объектом в большей степени влияет позиция этого объекта в списке рекомендаций, чем его релевантность этому пользователю [1].

Popularity Bias — эффект, при котором на вероятность взаимодействия пользователя с объектом в большей степени влияет популярность этого объекта в списке рекомендаций, чем его релевантность. Под популярностью объекта подразумевается количество взаимодействий с ним за период, предшествующий показу объекта в рекомендациях [4, 5].

Feedback Loop — эффект, который заключается в том, что рекомендательная система со временем «забывает» интересы пользователя, по которым он не дает достаточного количества обратной связи. Иногда этот эффект также называют Selection Bias [6, 7].

Склонность к рекомендации кликбейтов (Clickbait Bias) — эффект, связанный с тем, что пользователи активнее взаимодействуют с объектами с яркими обложками или «кричащими» заголовками, в результате чего история взаимодействий перестает отражать истинные интересы пользователя [1, 6].

Каждый из перечисленных эффектов отдельно рассмотрен в работах [8–10]. Для измерения работы эффекта вводится количественная оценка (обычно такие оценки называют прокси-метриками), а затем предлагается способ ее улучшения. Однако влияние эффектов на рекомендательные системы носит разнонаправленный характер, что затрудняет их совместное устранение [7]. Кроме того, предлагаемые прокси-метрики имеют разные единицы измерения и могут принимать значения из разных диапазонов, что не позволяет использовать их для проведения сравнительного анализа [1]. На данный момент в научных работах отсутствуют универсальные методики измерения, позволяющие не только выявить воздействие различных эффектов, но и сравнить их влияние между собой.

Таким образом, возникает задача построения модели качества рекомендательных систем, которая позволяла бы унифицировать часто используемые метрики и устанавливала бы однозначно интерпретируемую взаимосвязь между этими метриками и качеством рекомендаций. Способ приведения метрик должен работать независимо от диапазона, в котором находятся исходные значения метрик, а также от того, являются они дискретными или непрерывными. Также следует учесть, что значительная часть рассматриваемых метрик вычисляется как среднее значений, рассчитанных для каждого пользователя или объекта рекомендательной системы [1]. В этом случае модель качества должна сохранять возможность рассчитать приведенные метрики для каждого пользователя (или объекта) в отдельности для проведения сравнительного анализа работы рекомендательной системы между различными группами пользователей.

Обзор существующих подходов

Работы, связанные с исследованием характеристик, косвенно влияющих на рекомендательные системы, можно разделить на несколько направлений: исследования, направленные на измерение и устранение конкретных негативных эффектов, обзорные работы с перечислением метрик, а также работы, в которых выполняется сравнительный анализ различных подходов к оценке качества рекомендаций.

В работе [1] приводится 13 различных направлений, по которым можно оценивать рекомендательные системы. Предложено использовать метрики Average Recommendation Popularity (ARP) и Average Coverage of Long Tail items (ACLT) для измерения Popularity

Bias, Discounted Cumulative Gain (DCG) и Click-Through Rate (CTR) — для Position Bias, а для Clickbait Bias строят модель логистической регрессии вовлеченности пользователей по времени взаимодействия с объектом. В исследовании [11] подробна рассмотрена устойчивость системы к Feedback Loop, которая оценивается как способность предсказывать новые предпочтения пользователей.

В работах [8, 12, 13] изучена взаимосвязь между качеством и разнообразием рекомендаций, и предложены подходы, направленные на их увеличение. Низкое разнообразие рекомендаций может свидетельствовать как о сильном воздействии Feedback Loop, так и о высоком Popularity Bias модели [2].

Работы [4, 5, 10, 14, 15] посвящены изучению эффекта Popularity Bias в рекомендательных системах. В [4] для его оценки предлагается использовать частоты попадания в рекомендации среди наиболее популярных объектов, а также применять специальные метрики Popularity Lift (PL) и Miscalibration (MC), основанные на расстоянии Хеллингера между распределениями популярности в рекомендациях и исторических взаимодействиях. В [5] предложено переопределять порядок объектов в списке рекомендаций, смещая популярные объекты в конец списка. В [10] для оценки эффекта применена метрика ARP, а для повышения качества рекомендаций — регуляризация. Работа [15] предлагает подход, позволяющий получить хорошо кластеризуемые эмбеддинги товаров. Это позволяет рекомендовать специализированные непопулярные объекты без снижения качества.

В работе [3] применена коллаборативная информация для предсказания будущих трендов в рекомендациях. Предложенная модель отдает предпочтение объектам, которые будут популярны в будущем, разрывая таким образом Feedback Loop. Для измерения качества прогноза в [3] введены две новых метрики Acceleration (ACC) и Trendiness-Normalized-DCG.

Работы [6, 16] содержат сведения об основных метриках качества рекомендательных систем, которые позволяют получить количественные оценки удовлетворенности пользователей независимо от проявления различных факторов. К ним относятся Hit Rate (HR), Recall, Precision, CTR, Coverage и другие.

В [14, 17] исследовано влияние неравномерного распределения популярности между объектами на качество обучения модели рекомендаций. В [14] сделан акцент на том, как сэмплировать негативные примеры для обучения модели, чтобы нивелировать нежелательные эффекты от неравномерного количества обратной связи, а в [17] для этой цели предсказания модели скорректированы с учетом популярности объектов.

Таким образом, в научных работах представлено множество метрик, позволяющих оценивать рекомендательные системы с различных сторон, однако их сопоставление затруднено из-за разницы в диапазонах значений. Для оценки и интерпретации совокупного влияния различных эффектов можно было бы использовать значения Шепли [18], однако такой подход требует построения объясняющей регрессионной модели, выбор которой будет влиять на результат не меньше, чем сама рекомендательная система [8]. Следовательно, возникает необходимость построения модели качества рекомендательных систем, позволяющей сравнивать прокси-метрики с разными единицами измерения напрямую.

Постановка задачи

Пусть U — множество пользователей и I — множество объектов рекомендательной системы. Рекомендательная система $R = \{R_u\}_{u \in U}$, $R_u \subset I$ предоставляет для каждого пользователя набор персональных рекомендаций. Для каждого элемента $i \in R_u$ вычислен набор прокси-метрик $\boldsymbol{\mu}_{ui} \in \mathrm{M} \subseteq \mathbb{R}^k$, характеризующий данную рекомендацию с точки зрения различных эффектов, и некоторая основная метрика качества $q_{ui} \in Q \subseteq \mathbb{R}$. Единой оценкой качества рекомендаций является значение $\overline{q} \in \overline{Q}$:

$$\overline{q}(R) = \frac{1}{|U|} \sum_{u \in U} \sum_{i \in R_u} q_{ui}.$$

Качество рекомендации q_{ui} обычно вычисляется по некоторой тестовой выборке $S = \{S_u\}_{u \in U}$, где $S_u \subset I$ — объекты, с которыми взаимодействовал пользователь u. На практике часто используют индикаторную функцию множества S_u:

$$q_{ui} = \frac{[i \in S_u]}{|R_u|}, \tag{1}$$

где [] — скобка Айверсона.

Соответствующее таким q_{ui} значение $\overline{q}$ называют уровнем попадания HR [19]. В зависимости от области применения рекомендательной системы формула (1) может быть модифицирована, например, на торговых площадках эту величину умножают на стоимость товара, получая Gross Merchandise Value, а в медиасервисах — на длительность взаимодействия, и называют соответствующее значение Time Spent [16].

Чем больше значение $\overline{q}$, тем лучше рекомендательная система. Множества U и I — конечные, а значит, конечно множество $\mathcal{R} = \{R\}$ всех возможных рекомендательных систем.

$$|\mathcal{R}| = 2^{|U||I|}.$$

Следовательно, для фиксированных U, I, S существует система с наилучшим возможным качеством

$$R^* = \arg\max_{R \in \mathcal{R}} \overline{q}(R) = S.$$

Аналогично основной метрике качества, можно вычислить вектор средних значений $\overline{\boldsymbol{\mu}} \in \overline{\mathrm{M}}$, который характеризует рекомендательную систему с точки зрения исследуемых эффектов.

$$\overline{\boldsymbol{\mu}}(R) = \frac{1}{|U|} \sum_{u \in U} \sum_{i \in R_u} \boldsymbol{\mu}_{ui}.$$

Обычно в работах, связанных с устранением таких эффектов, как Popularity или Position Bias, в экспериментах сравниваются пары $(\overline{\boldsymbol{\mu}}, \overline{q})$ построенных рекомендательных систем, а также эмпирически проверяется, что наилучшему значению прокси-метрики соответствует лучшая рекомендательная система [5, 10, 12].

Прокси-метрики могут иметь разнонаправленный характер, и в общем случае точки экстремума прокси-метрик и основной метрики качества не совпадают [4]. Оптимизация разных прокси-метрик по-разному меняет значение $\overline{q}$, поэтому при работе одновременно с несколькими эффектами выбор прокси-метрики для улучшения играет ключевую роль [2].

Пусть R — некоторая рекомендательная система. Необходимо выявить такой эффект j в системе, устранение которого обеспечит максимальный прирост качества рекомендаций. Пусть $\overline{\mu}^{(j)}$, $j = 1, \ldots, k$ — компоненты вектора $\overline{\boldsymbol{\mu}}$, а $\mathrm{M}_j^* \subseteq \overline{\mathrm{M}}$ — множество всех векторов $\overline{\boldsymbol{\mu}}$, для которых $\overline{\mu}^{(j)}$ принимает наилучшее возможное значение. Тогда:

$$\sup_{R' \in \mathcal{R} | \overline{\boldsymbol{\mu}}(R) \in \mathrm{M}_j^*} (\overline{q}(R') - \overline{q}(R)) \to \max_j. \qquad (2)$$

Рассмотрим этапы решения задачи (2).

Этап 1. Построение проекции P: $\mathrm{M} \to \mathrm{M}'$ пространства метрик M в пространство M′, где все размерности измеряются по одной шкале, у которой наилучшее значение равно нулю, а наихудшее или 1, или $+\infty$, и зависимость между $\overline{\boldsymbol{\mu}}(R)$ и $\sup(\overline{q}(R))$ монотонная. Проекция будет построена на основе некоторой выборки $(\boldsymbol{\mu}_{ui}, q_{ui})$.

Этап 2. Построение аддитивного разложения вида

$$\overline{q}(R) = 1 - \varepsilon - \sum_{j=1}^{k} w_j,$$

где w_j характеризует влияние эффекта j на качество рекомендаций, а ε — случайная ошибка системы. Если j-й эффект не влияет на систему, то $w_j = 0$. Соответственно, для оценки прироста качества после устранения j-го эффекта можно воспользоваться равенством

$$\sup_{R' \in \mathcal{R} | \overline{\boldsymbol{\mu}}(R) \in \mathrm{M}_j^*} (\overline{q}(R') - \overline{q}(R)) = w_j.$$

Таким образом (2) сводится к нахождению наибольшего w_j в аддитивном разложении. Такая задача построения модели качества рекомендаций ранее не рассматривалась в известных работах и представляет собой научную новизну.

Модель

В большинстве метрик качества рекомендательных систем ключевую роль играет индикаторная функция $[i \in S_u]$. Пусть p — вероятность того, что случайная рекомендация (u, i), сделанная исследуемой рекомендательной системой, окажется релевантной. Пусть $\gamma \sim Ber(p)$ — соответствующая ей случайная величина Бернулли. На вероятность p оказывают влияние прокси-метрики $\boldsymbol{\mu}$, которые в свою очередь также являются случайными величинами. В таком случае γ имеет комбинированное распределение

$$p = \mathcal{P}(\gamma = 1) = \int_{\mathrm{M}} \mathcal{P}(\boldsymbol{\gamma = 1 | \mu}) \mathcal{P}(\boldsymbol{\mu}) d\boldsymbol{\mu}$$

и, соответственно,

$$1 - p = \int_{\mathrm{M}} \mathcal{P}((\boldsymbol{\gamma = 0 | \mu}) \mathcal{P}(\boldsymbol{\mu}) d\boldsymbol{\mu}, \qquad (3)$$

где $\mathcal{P}$ — вероятность; M — множество всех возможных значений прокси-метрик. В последующих шагах интеграл в правой части (3) будет разбит на составляющие w_j и ε.

Выполним этап 1 модели — построим проекцию P: $\mathrm{M} \to \mathrm{M}'$. Для каждой компоненты $\mu^{(j)}$ вектора $\boldsymbol{\mu}$ определим две маргинальные условные плотности распределения: $\mathrm{pdf}(\mu^{(j)} | \gamma = 1)$ и $\mathrm{pdf}(\mu^{(j)} | \gamma = 0)$. Сравнивая значения данных плотностей, множество $\mathrm{M}^{(j)}$ всех возможных значений j-ой метрики разделим на множество нормальных значений $\mathrm{M}_1^{(j)}$ и критических $\mathrm{M}_0^{(j)}$.

$$\mathrm{M}_1^{(j)} = \{\mu^{(j)} \in \mathrm{M}^{(j)} | \mathrm{pdf}(\mu^{(j)} | \gamma = 1) > \mathrm{pdf}(\mu^{(j)} | \gamma = 0)\},$$
$$\mathrm{M}_0^{(j)} = \mathrm{M}^{(j)} \backslash \mathrm{M}_1^{(j)}.$$

На рисунке представлена наглядная визуализация такого разбиения.

После того, как разбиение построено для каждого признака, в каждой точке M можно определить, сколько признаков принимают критические значения. Обозначим соответствующий счетчик Count Critical (CC).

$$CC(\boldsymbol{\mu}) = \sum_{j=1}^{k} [\mu^{(j)} \in \mathrm{M}_0^{(j)}].$$

Разобьем M на M_0 и $\mathrm{M}_{>0}$ в зависимости от того, какие значения принимает функция CC.

$$\mathrm{M}_0 = \{\boldsymbol{\mu} \in \mathrm{M} | CC(\boldsymbol{\mu}) = 0\},$$
$$\mathrm{M}_{>0} = \{\boldsymbol{\mu} \in \mathrm{M} | CC(\boldsymbol{\mu}) > 0\}. \qquad (4)$$

Для элементов множества $\mathrm{M}_{>0}$ построим проекцию P, используя функцию CC.

$$P^{(j)}(\boldsymbol{\mu}_{ui}) = \frac{[\mu_{ui}^{(j)} \in \mathrm{M}_0^{(j)}]}{CC(\boldsymbol{\mu}_{ui})}(1 - [i \in S_u]). \qquad (5)$$

Перейдем к этапу 2 модели. Проекция P имеет следующий смысл. Если хотя бы одна из метрик принимает критическое значение, то $\mathcal{P}(\gamma = 0 | \boldsymbol{\mu}) = 1$, причем каждая из метрик равновероятно является причиной

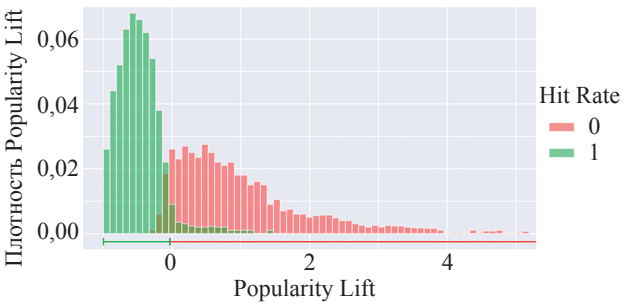

Рисунок. Гистограммы условных распределений Popularity Lift некоторой рекомендательной системы, построенной для набора данных MovieLens100K

Figure. Histograms of the conditional Popularity Lift distribution of a recommender system built for the MovieLens100K dataset

нерелевантности. Если все метрики принимают нормальное значение, то $\mathcal{P}(\gamma = 0|\mathbf{\mu}) = 0$. Тогда $\forall \mathbf{\mu} \neq 0$

$$\mathcal{P}(\gamma = 0|\mathbf{\mu}) = \sum_{i=1}^{k} P^{(j)}(\mathbf{\mu}), \tag{6}$$

и интеграл (3) можно разделить по множествам (4).

$$1 - p = \int_{\mathrm{M}_0} \mathcal{P}(\gamma = 0|\mathbf{\mu}) \times \\ \times \mathcal{P}(\mathbf{\mu}) d\mathbf{\mu} + \int_{\mathrm{M}_{>0}} \mathcal{P}(\gamma = 0|\mathbf{\mu}) \mathcal{P}(\mathbf{\mu}) d\mathbf{\mu}. \tag{7}$$

Первое слагаемое в правой части выражения (7) соответствует ситуации, когда рекомендация оказалось нерелевантна, а все прокси-метрики приняли нормальное значение. Вероятность данного события легко можно оценить по выборке, в рамках модели качества она обозначена ε и отражает долю неудачных рекомендаций, которые нельзя объяснить с точки зрения выбранных прокси-метрик. Второе слагаемое (7) можно разложить в сумму, используя (6).

$$1 - p = \varepsilon + \sum_{i=1}^{k} \int_{\mathrm{M}_{>0}} P^{(j)}(\mathbf{\mu}) \mathcal{P}(\mathbf{\mu}) d\mathbf{\mu}. \tag{8}$$

Равенство (8) определяет модель качества рекомендательной системы. Внедрение метода, направленного на устранение j-го эффекта, позволит устранить критические значения в соответствующей метрике. В результате качество рекомендаций возрастет не более, чем на

$$w_j = \int_{\mathrm{M}_{>0}} P^{(j)}(\mathbf{\mu}) \mathcal{P}(\mathbf{\mu}) d\mathbf{\mu}.$$

Таким образом, решением задачи (2) является пара (j, w_j) с наибольшим w_j.

Объекты экспериментов

Для демонстрации работы модели качества для набора данных MovieLens-100K были построены и проанализированы четыре рекомендательные системы: Top by CTR, Top by CTR within Cohort, Item KNN и MF-BPR.

Набор данных MovieLens-100K содержит информацию 100 тыс. оценок 943 пользователей 1682 фильмам. По каждому пользователю даны демографические данные: пол, возраст, тип занятости, а также не менее 20 его оценок. Для построения выбранных рекомендательных систем данные были разделены на обучающую и тестовую выборки таким образом, чтобы в тестовой выборке оказались последние 10 оценок каждого пользователя. Для всех пользователей подготовлено по 10 рекомендаций.

Top by CTR — рекомендательная система, в которой все фильмы сортируются по убыванию CTR, т. е., отношения количества оценок к количеству пользователей, рассчитанному на обучающей выборке. Каждому пользователю предлагаются фильмы, которые имеют наибольшую оценку кликов (CTR) среди тех, которые он еще не оценивал.

Top by CTR within Cohort использует такую же механику, что и Top by CTR, с небольшим усложнением. Все пользователи разбиваются на группы по полу и возрасту. Чтобы групп было не слишком много, вместо реального возраста используются возрастные группы, составленные на основе децилей возрастов. Внутри каждой группы реализуется система Top by CTR.

Item KNN — рекомендательная система, которая для каждой пары фильмов считает количество пользователей, посмотревших их.

$$\mathrm{sim}(i, i') = \frac{1}{|U|} \sum_{u \in U} [i \in S_u \wedge i' \in S_u].$$

Для оценки релевантности фильма i пользователю u вычисляется суммарное сходство между i и всеми фильмами S_u.

$$\mathrm{rel}(u, i) = \sum_{i' \in S_u} \mathrm{sim}(i, i').$$

Фильмы сортируются по убыванию $\mathrm{rel}(u, i)$, из списка удаляются фильмы, оцененные пользователем ранее.

Matrix Factorization, Bayesian Pairwise Ranking (MF-BPR) — рекомендательная система, в которой каждому пользователю и каждому фильму ставятся в соответствие эмбеддинги e_u, e_i. Релевантность оценивается как скалярное произведение эмбеддингов. Модель обучалась по методологии BPR, при которой для каждого взаимодействия i пользователя u из дискретного равномерного распределения выбирается случайный фильм i', который пользователь не оценивал, а функция потерь вычисляется по формуле

$$\mathrm{loss} = \log_2 \sigma(e_u^{\top}(e_i - e_{i'})).$$

Модель была обучена на 40 эпохах методом Adagrad.

Эксперименты

Каждая рекомендательная система оценивалась по четырем метрикам.

HR для оценки общего качества рекомендаций (1).

PL для оценки влияния Popularity Bias [4].

Novelty Lack (NL) для оценки эффекта недостатка новизны, при котором рекомендательная система предпочитает рекомендовать старые объекты новым. NL рассчитывается так же, как PL, только вместо популярности используется возраст фильма [1].

Коэффициент ранговой корреляции Спирмена между предсказаниями рекомендательной системы и CTR внутри социально-демографической группы (SCbSaCP) [6, 20, 21].

Результаты расчетов по всем пользователям представлены в табл. 1.

С точки зрения абсолютных значений метрик алгоритмы показали предсказуемые результаты. Наилучшее качество продемонстрировала самая сложная система, самый высокий PL показала Top By CTR, основанная на популярности объектов, а самый высокий SCbSaCP — система Top by CTR within Cohort. Применим модель качества к полученным значениям метрик, чтобы оценить влияние соответствующих эффектов на системы (табл. 2).

Таблица 1. Абсолютные значения метрик

Table 1. The absolute scores of the metrics

Рекомендательная система	HR	PL	SCbSaCP	NL
Top By CTR	0,0768	**2,59**	0,55	2,05
Top By CTR within Cohort	0,0721	2,22	**1,00**	**2,47**
Item KNN	0,0965	2,31	0,52	2,10
MF-BPR	**0,1047**	1,85	0,48	1,24

Примечание: жирным шрифтом выделены строки с наибольшими значениями метрик.

Таблица 2. Результат применения комбинированной модели качества

Table 2. The compound quality model projection

Рекомендательная система	p	$P^{(PL)}(\boldsymbol{\mu})$	$P^{(SCbSaCP)}(\boldsymbol{\mu})$	$P^{(NL)}(\boldsymbol{\mu})$	ε
Top By CTR	0,0768	0,0161	0,0454	0,0502	0,8115
Top By CTR within Cohort	0,0721	0,0116	0,0000	0,0469	**0,8694**
Item KNN	0,0965	**0,0212**	**0,0523**	0,0640	0,7660
MF-BPR	**0,1047**	0,0192	0,0380	**0,0684**	0,7697

Примечание: жирным шрифтом выделены строки с наибольшими значениями.

В каждой строке табл. 2 сумма значений равна единице. Результаты проекции значительно отличаются от исходных значений метрик, приведенных в табл. 1. Например, влияние Popularity Bias на систему оказалось наибольшим у системы Item KNN, а не у Top By CTR. Действительно, система Top By CTR часто ошибается из-за того, что предлагает пользователям слишком популярные объекты, но и релевантные рекомендации он подбирает по той же причине. Следовательно, данную систему нельзя улучшить, за счет снижения популярности в рекомендациях. Зато значение $P^{(NL)}(\boldsymbol{\mu}) = 0{,}0502$ указывает на возможность улучшения системы Top By CT за счет удаления старых фильмов из рекомендаций. Верхняя граница p для Top By CTR в случае оптимизации NL составляет 0,0768 + 0,0502 = = 0,1270.

Аналогичные выводы можно сделать и относительно модели MF-BPR. С точки зрения абсолютных значений может показаться, что Popularity Bias имеет более сильное влияние на качество, чем NL. Но на самом деле такое различие объясняется тем, что у PL и NL отличаются выборочные распределения, и в действительности недостаток новизны имеет большее влияние. У моделей Item KNN и MF-BPR меньше ε, что говорит о том, что остальные метрики сильнее разделяются на нормальные и критические значения.

Для более детального анализа все пользователи были разделены на четыре группы (сегмента) на основе квартилей по количеству оценок на обучающей выборке. Первая группа — самые активные пользователи, четвертая группа — наименее активные пользователи. Сравним абсолютные значения метрик и проекцию (5) на примере модели Item KNN (табл. 3).

У активных пользователей высокое абсолютное значение PL и NL. Это объясняется тем, что большая часть оценок новым и популярным фильмам попала в обучающую выборку, а старые и непопулярные фильмы пользователи стали смотреть в последнюю очередь, именно они оказались в тестовой выборке. Для малоактивных пользователей новые популярные фильмы оказались в тестовой выборке, поэтому PL и NL для них имеют более кучный характер. Как следствие, такие пользователи оказываются более чувствительными к исследуемым эффектам и для них можно значительно улучшить рекомендации.

Обсуждение

Рассмотрим подробнее, чем отличается предложенная модель качества от использования прокси-метрик.

Исходные значения метрик необходимы для описания рекомендательной системы с различных сторон.

Таблица 3. Метрики Item KNN для разных сегментов пользователей

Table 3. Item KNN scores for different user segments

Сегмент	HR	PL	SCbSaCP	NL	$P^{(PL)}(\boldsymbol{\mu})$	$P^{(SCbSaCP)}(\boldsymbol{\mu})$	$P^{(NL)}(\boldsymbol{\mu})$	ε
1	0,0615	**3,1137**	0,4930	**3,3596**	0,0141	0,0333	0,0350	**0,8560**
2	0,0853	2,4510	**0,5412**	1,4872	0,0151	0,0440	0,0509	0,8047
3	0,1000	1,9052	0,5280	1,9001	0,0212	0,0511	0,0654	0,7623
4	**0,1370**	1,7684	0,5101	1,6500	**0,0337**	**0,0793**	**0,1028**	0,6472

Примечание: жирным шрифтом выделены строки с наибольшими значениями.

Если одна из метрик близка к экстремуму, это помогает быстро выявить влияние соответствующего эффекта. Прокси-метрики, как правило, разрабатываются таким образом, чтобы их единицы измерения имели четкую интерпретацию. Тем не менее, при значениях, удаленных от экстремальных, оценивание влияния эффекта становится затруднительным. Кроме того, различие величин измерения у разных прокси-метрик делают невозможным их сравнение между собой.

Предложенная модель качества унифицирует прокси-метрики, что позволяет решить проблему их сопоставления. Кроме того, полученные значения вклада каждого эффекта напрямую связаны с количеством ошибочных рекомендаций, что позволяет быстро оценить прирост качества рекомендательной системы в случае его устранения.

Таким образом, совместное использование прокси-метрик и комбинированной модели качества позволяют всесторонне описать работу рекомендательной системы.

Заключение

В работе предложена новая модель, позволяющая численно измерить взаимосвязь между качеством рекомендательной системы и ее произвольными прокси-метриками. Модель названа комбинированной, поскольку в ее основе лежит предположение о том, что факт успешной рекомендации — случайная величина с комбинированным распределением. Разработан алгоритм, позволяющий численно оценивать вклад различных факторов в вероятность ошибки рекомендательной системы. Знание таких вкладов позволяет определить список наиболее важных параметров, которые следует учитывать при дальнейшей разработке системы. Данный подход позволяет сравнивать качество рекомендательной системы на различных сегментах пользователей и сравнивать поведение различных рекомендательных систем между собой.

Литература

1. Anderson A., Maystre L., Anderson I., Mehrotra R., Lalmas M. Algorithmic effects on the diversity of consumption on spotify // Proc. of the Web Conference. 2020. P. 2155–2165. https://doi.org/10.1145/3366423.3380281
2. Avazpour I., Pitakrat T., Grunske L., Grundy J. Dimensions and metrics for evaluating recommendation systems // Recommendation Systems in Software Engineering. 2014. P. 245–273. https://doi.org/10.1007/978-3-642-45135-5_10
3. Ding H., Kveton B., Ma Y., Park Y., Kini V., Gu Y., et al. Trending now: modeling trend recommendations // Proc. of the 17th ACM Conference on Recommender Systems. 2023. P. 294–305. https://doi.org/10.1145/3604915.3608810
4. Cai Y., Guo J., Fan Y., Ai Q., Zhang R., Cheng X. Hard negatives or false negatives: correcting pooling bias in training neural ranking models // Proc. of the 31st ACM International Conference on Information and Knowledge Management. 2022. P. 118–127. https://doi.org/10.1145/3511808.3557343
5. Abdollahpouri H., Mansoury M., Burke R., Mobasher B. The connection between popularity bias, calibration, and fairness in recommendation // Proc. of the 14th ACM Conference on Recommender Systems. 2020. P. 726–731. https://doi.org/10.1145/3383313.3418487
6. Beel J., Langer S., Genzmehr M., Gipp B., Breitinger C., Nürnberger A. Research paper recommender system evaluation: a quantitative literature survey // Proc. of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation. 2013. P. 15–22. https://doi.org/10.1145/2532508.2532512
7. Wasilewski J., Hurley N. Incorporating diversity in a learning to rank recommender system // Proc. of the 29th International Florida Artificial Intelligence Research Society Conference. 2016. P. 1–6.
8. Ricci F., Rokach L., Shapira B. Recommender Systems Handbook. Springer, 2010. 842 p.
9. Said A., Bellogin A. Comparative recommender system evaluation: benchmarking recommendation frameworks // Proc. of the 8th ACM Conference on Recommender Systems. 2014. P. 129–136. https://doi.org/10.1145/2645710.2645746
10. Wilhelm M., Ramanathan A., Bonomo A., Jain S., Chi E.H., Gillenwater J. Practical diversified recommendations on YouTube with determinantal point processes // Proc. of the 27th ACM International Conference on Information and Knowledge Management. 2018. P. 2165–2173. https://doi.org/10.1145/3269206.3272018
11. Chang Bo, Meng C., Ma H., Chang S., Gu Y., Peng Y., et al. Cluster anchor regularization to alleviate popularity bias in recommender systems // Proc. of the Companion Proceedings of the ACM Web Conference. 2024. P. 151–160. https://doi.org/10.1145/3589335.3648312

References

1. Anderson A., Maystre L., Anderson I., Mehrotra R., Lalmas M. Algorithmic effects on the diversity of consumption on spotify. *Proc. of the Web Conference*, 2020, pp. 2155–2165. https://doi.org/10.1145/3366423.3380281
2. Avazpour I., Pitakrat T., Grunske L., Grundy J. Dimensions and metrics for evaluating recommendation systems. *Recommendation Systems in Software Engineering*, 2014, pp. 245–273. https://doi.org/10.1007/978-3-642-45135-5_10
3. Ding H., Kveton B., Ma Y., Park Y., Kini V., Gu Y., et al. Trending now: modeling trend recommendations. *Proc. of the 17th ACM Conference on Recommender Systems*, 2023, pp. 294–305. https://doi.org/10.1145/3604915.3608810
4. Cai Y., Guo J., Fan Y., Ai Q., Zhang R., Cheng X. Hard negatives or false negatives: correcting pooling bias in training neural ranking models. *Proc. of the 31st ACM International Conference on Information and Knowledge Management*, 2022, pp. 118–127. https://doi.org/10.1145/3511808.3557343
5. Abdollahpouri H., Mansoury M., Burke R., Mobasher B. The connection between popularity bias, calibration, and fairness in recommendation. *Proc. of the 14th ACM Conference on Recommender Systems*, 2020, pp. 726–731. https://doi.org/10.1145/3383313.3418487
6. Beel J., Langer S., Genzmehr M., Gipp B., Breitinger C., Nürnberger A. Research paper recommender system evaluation: a quantitative literature survey. *Proc. of the International Workshop on Reproducibility and Replication in Recommender Systems Evaluation*, 2013, pp. 15–22. https://doi.org/10.1145/2532508.2532512
7. Wasilewski J., Hurley N. Incorporating diversity in a learning to rank recommender system. *Proc. of the 29th International Florida Artificial Intelligence Research Society Conference*, 2016, pp. 1–6.
8. Ricci F., Rokach L, Shapira B. *Recommender Systems Handbook*. Springer, 2010, 842 p.
9. Said A., Bellogin A. Comparative recommender system evaluation: benchmarking recommendation frameworks. *Proc. of the 8th ACM Conference on Recommender Systems*, 2014, pp. 129–136. https://doi.org/10.1145/2645710.2645746
10. Wilhelm M., Ramanathan A., Bonomo A., Jain S., Chi E.H., Gillenwater J. Practical diversified recommendations on YouTube with determinantal point processes. *Proc. of the 27th ACM International Conference on Information and Knowledge Management*, 2018, pp. 2165–2173. https://doi.org/10.1145/3269206.3272018
11. Chang Bo, Meng C., Ma H., Chang S., Gu Y., Peng Y., et al. Cluster anchor regularization to alleviate popularity bias in recommender systems. *Proc. of the Companion Proceedings of the ACM Web Conference*, 2024, pp. 151–160. https://doi.org/10.1145/3589335.3648312

12. Bellogin A., Castells P., Cantador I. Precision-oriented evaluation of recommender systems: an algorithmic comparison // Proc. of the 5th ACM Conference on Recommender Systems. 2011. P. 333–336. https://doi.org/10.1145/2043932.2043996
13. Cremonesi P., Koren Y., Turrin R. Performance of recommender algorithms on top-n recommendation tasks // Proc. of the 4th ACM Conference on Recommender Systems. 2010. P. 39–46. https://doi.org/10.1145/1864708.1864721
14. Abdollahpouri H., Burke R., Mobasher B. Managing popularity bias in recommender systems with personalized re-ranking // Proc. of the 32nd International Florida Artificial Intelligence Research Society Conference. 2019. P. 1–6.
15. Yi X., Yang J., Hong L., Cheng D.Z., Heldt L., Kumthekar A., Zhao Z., Wei L., Chi E. Sampling-bias-corrected neural modeling for large corpus item recommendations // Proc. of the 13th ACM Conference on Recommender Systems. 2019. P. 269–277. https://doi.org/10.1145/3298689.3346996
16. Silveira T., Zhang M., Lin X., Liu Y., Ma S. How good your recommender system is? A survey on evaluations in recommendation // International Journal of Machine Learning and Cybernetics. 2019. V. 10. N 5. P. 813–831. https://doi.org/10.1007/s13042-017-0762-9
17. Akiyama T., Obara K., Tanizaki M. Proposal and evaluation of serendipitous recommendation method using general unexpectedness // CEUR Workshop Proceedings. 2010. V. 676. P. 3–10.
18. Scott L.M., Su-In L. A unified approach to interpreting model predictions // Proc. of the 31st Conference on Neural Information Processing Systems. 2017. P. 1–10.
19. Isinkaye F.O., Folajimi Y.O., Ojokoh B.A. Recommendation systems: principles, methods and evaluation // Egyptian Informatics Journal. 2015. V. 16. N 3. P. 261–273. https://doi.org/10.1016/j.eij.2015.06.005
20. Rhee W., Cho S.-M., Suh B. Countering popularity bias by regularizing score differences // Proc. of the 16th ACM Conference on Recommender Systems. 2022. P. 145–155. https://doi.org/10.1145/3523227.3546757
21. Shani G., Gunawardana A. Evaluating recommendation systems // Recommender Systems Handbook. 2010. P. 257–297. https://doi.org/10.1007/978-0-387-85820-3_8

12. Bellogin A., Castells P., Cantador I. Precision-oriented evaluation of recommender systems: an algorithmic comparison. *Proc. of the 5th ACM Conference on Recommender Systems*, 2011, pp. 333–336. https://doi.org/10.1145/2043932.2043996
13. Cremonesi P., Koren Y., Turrin R. Performance of recommender algorithms on top-n recommendation tasks. *Proc. of the 4th ACM Conference on Recommender Systems*, 2010, pp. 39–46. https://doi.org/10.1145/1864708.1864721
14. Abdollahpouri H., Burke R., Mobasher B. Managing popularity bias in recommender systems with personalized re-ranking. *Proc. of the 32nd International Florida Artificial Intelligence Research Society Conference*, 2019, pp. 1–6.
15. Yi X., Yang J., Hong L., Cheng D.Z., Heldt L., Kumthekar A., Zhao Z., Wei L., Chi E. Sampling-bias-corrected neural modeling for large corpus item recommendations. *Proc. of the 13th ACM Conference on Recommender Systems*, 2019, pp. 269–277. https://doi.org/10.1145/3298689.3346996
16. Silveira T., Zhang M., Lin X., Liu Y., Ma S. How good your recommender system is? A survey on evaluations in recommendation. *International Journal of Machine Learning and Cybernetics*, 2019, vol. 10, no. 5, pp. 813–831. https://doi.org/10.1007/s13042-017-0762-9
17. Akiyama T., Obara K., Tanizaki M. Proposal and evaluation of serendipitous recommendation method using general unexpectedness. *CEUR Workshop Proceedings*, 2010, vol. 676, pp. 3–10.
18. Scott L.M., Su-In L. A unified approach to interpreting model predictions. *Proc. of the 31st Conference on Neural Information Processing Systems*, 2017, pp. 1–10.
19. Isinkaye F.O., Folajimi Y.O., Ojokoh B.A. Recommendation systems: principles, methods and evaluation. *Egyptian Informatics Journal*, 2015, vol. 16, no. 3, pp. 261–273. https://doi.org/10.1016/j.eij.2015.06.005
20. Rhee W., Cho S.-M., Suh B. Countering popularity bias by regularizing score differences. *Proc. of the 16th ACM Conference on Recommender Systems*, 2022, pp. 145–155. https://doi.org/10.1145/3523227.3546757
21. Shani G., Gunawardana A. Evaluating recommendation systems. *Recommender Systems Handbook*, 2010, pp. 257–297. https://doi.org/10.1007/978-0-387-85820-3_8

Авторы

Цыплов Алексей Михайлович — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0009-5211-3019, tsyplov80@mail.ru

Бухановский Александр Валерьевич — доктор технических наук, профессор, директор мегафакультета трансляционных информационных технологий, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 6603474810, https://orcid.org/0000-0003-1588-8164, avbukhanovskii@itmo.ru

Authors

Aleksei M. Tsyplov — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0009-5211-3019, tsyplov80@mail.ru

Alexander V. Boukhanovsky — D.Sc., Professor, Head of the School of Translational Information Technologies, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 6603474810, https://orcid.org/0000-0003-1588-8164, avbukhanovskii@itmo.ru

Статья поступила в редакцию 08.08.2025
Одобрена после рецензирования 17.09.2025
Принята к печати 24.11.2025

Received 08.08.2025
Approved after reviewing 17.09.2025
Accepted 24.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1125-1133
УДК 007.52

Метод пространственного противодействия нарушителю в рое беспилотных воздушных судов

Павел Юрьевич Шамрай[1]✉, Данил Анатольевич Заколдаев[2], Андрей Михайлович Бойко[3]

[1,2,3] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
[3] Физико-технический институт им. А.Ф. Иоффе, Санкт-Петербург, 194021, Российская Федерация

[1] pavel.shamray@mail.ru✉, https://orcid.org/0009-0006-0822-5310
[2] d.zakoldaev@mail.ru, https://orcid.org/0000-0002-2520-1998
[3] andray599@gmail.com, https://orcid.org/0000-0002-3120-6170

Аннотация
Введение. Развитие технологий децентрализованного управления роями беспилотных воздушных судов требует создания новых методов обеспечения их устойчивости к внутренним угрозам. Наличие в составе роя агента-нарушителя создает угрозы энергетических или информационных атак. Особенно критична ситуация, когда агент-нарушитель находится в центре роя и его влияние на соседей максимально. Существующие исследования в основном сосредоточены на обнаружении агентов-нарушителей, тогда как методы противодействия, в частности, — пространственного вытеснения нарушителя из группы — мало изучены. В работе представлен и проанализирован разработанный метод пространственного противодействия нарушителю, не требующий его явного обнаружения или обмена информацией между агентами. **Метод.** Предлагаемый метод основан на оригинальной идее проведения аналогии между управлением роем беспилотных воздушных судов (агентов) с процессами, происходящими в полупроводниковом кристалле. Противодействие агенту-нарушителю достигается за счет временного изменения агентами определенных параметров роевого взаимодействия. Вследствие этого меняется пространственная структура роя, а нарушитель, не меняющий свои параметры взаимодействия, начинает движение относительно остальных агентов, оказываясь на краю группы. В работе исследованы следующие варианты реализации разработанного метода противодействия, основанные на сжатии, расширении и последовательном перестроении структуры роя. **Основные результаты.** Выполнено имитационное моделирование поведения роя беспилотных воздушных судов в присутствии нарушителя. Показателем результативности применения метода считалась вероятность наличия у агента-нарушителя менее пяти соседних беспилотных воздушных судов, т. е. нахождение агента-нарушителя на краю группы. Наилучший результат (вероятность, близкая к 1,0) продемонстрировал вариант сжатия роя. Вариант расширения роя показал меньшую результативность (негарантированный вариант), вариант последовательного перестроения оказался неэффективным. Показано, что основной вклад в результативность метода вносит степень изменения расстояния между агентами. Метод, реализованный в варианте сжатия роя, показал эффективность для числа агентов в рое от 19 до 91. **Обсуждение.** Предложенный метод позволяет уменьшить вероятность деструктивного влияния нарушителя в рое беспилотных воздушных судов, основываясь только на локальных навигационных данных агентов, не прибегая к обнаружению нарушителя. Это дает возможность применять метод в системах с ограниченными коммуникационными возможностями. Некоторое повышение энергетических затрат может быть снижено путем оптимизации длительности воздействия и сохранения структуры роя.

Ключевые слова
рой, беспилотные воздушные суда, безопасность, нарушитель, имитационное моделирование, потенциальные поля, децентрализованное управление, кристалл, примесь

Благодарности
Работа выполнена в Университете ИТМО при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках проекта № 70-2024-001354 «Разработка технологий и демонстратора комплексной системы группового управления, взаимодействия и организации поведения группы БВС при выполнении целевых задач».

© Шамрай П.Ю., Заколдаев Д.А., Бойко А.М., 2025

Ссылка для цитирования: Шамрай П.Ю., Заколдаев Д.А., Бойко А.М. Метод пространственного противодействия нарушителю в рое беспилотных воздушных судов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1125–1133. doi: 10.17586/2226-1494-2025-25-6-1125-1133

Method of spatial countering an intruder in a swarm of unmanned aerial vehicles

Pavel Yu. Shamray[1]✉, Danil A. Zakoldaev[2], Andrey M. Boyko[3]

[1,2,3] ITMO University, Saint Petersburg, 197101, Russian Federation
[3] Ioffe Institute, Saint Petersburg, 194021, Russian Federation

[1] pavel.shamray@mail.ru✉, https://orcid.org/0009-0006-0822-5310
[2] d.zakoldaev@mail.ru, https://orcid.org/0000-0002-2520-1998
[3] andray599@gmail.com, https://orcid.org/0000-0002-3120-6170

Abstract
The development of decentralized control technologies for swarms of Unmanned Aerial Vehicles requires the development of new methods for ensuring their resilience to internal threats. The emergence of an intruder agent within a swarm creates threats of energy or information attacks. The situation is especially critical when the intruder agent is located at the center of the swarm, with its influence on its neighbors being greatest. Existing research has focused primarily on detecting intruder agents, while countermeasures, particularly spatial exclusion of the intruder agent from the group, remain poorly understood. This study develops and analyzes a method for spatial countermeasures against the intruder agent that does not require its explicit detection or information exchange between agents. The proposed method is based on the original idea of analogizing the control of a swarm of unmanned aerial vehicles (agents) with the processes occurring in a semiconductor crystal. Countermeasures against the intruder agent are achieved through the temporary modification of certain swarm interaction parameters by the agents. As a result, the spatial structure of the swarm changes, and the intruder, who does not change its interaction parameters, begins to move relative to the other agents, ending up at the edge of the group. Three implementation options for the proposed countermeasure method, based on compression, expansion, and sequential restructuring of the swarm structure, are investigated. Simulation modeling of the behavior of a swarm of unmanned aerial vehicles in the presence of an intruder was performed. The success of the method was measured by the probability of the intruder agent having fewer than five neighboring unmanned aerial vehicles, i.e., being at the edge of the group. The best performance (probability close to 1.0) was demonstrated by the swarm compression option. The swarm expansion option showed lower performance (non-guaranteed option). The sequential restructuring option proved ineffective. It is shown that the degree of change in the distance between agents makes the main contribution to the effectiveness of the method. The proposed method implemented in the swarm compression mode demonstrated effectiveness for swarm numbers ranging from 19 to 91. The proposed method reduces the likelihood of destructive intruder influence in a swarm of unmanned aerial vehicles, relying solely on the agents local navigation data, without resorting to intruder detection. This makes the method applicable to systems with limited communication capabilities. Some increase in energy consumption can be mitigated by optimizing the intrusion duration and maintaining the swarm structure.

Keywords
swarm, unmanned aerial vehicles, security, intruder, simulation, potential fields, decentralized control, crystal, impurity

Acknowledgements
The work was carried out at ITMO University with the financial support of the Ministry of Science and Higher Education of the Russian Federation as part of the project No. 70-2024-001354 “Development of technologies and a demonstrator of a complex system for group control, interaction and organization of behavior of a group of UAVs in performing target tasks”.

For citation: Shamray P.Yu., Zakoldaev D.A., Boyko A.M. Method of spatial countering an intruder in a swarm of unmanned aerial vehicles. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1125–1133 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1125-1133

Введение

С каждым годом растет интерес к группам беспилотных воздушных судов (БВС) [1, 2]. Преимущества групп по сравнению с отдельными БВС заключаются в их большей отказоустойчивости при выполнении задач, повышении вероятности выполнения поставленных задач и большей площади покрытия при мониторинге и поиске объектов. Также группы БВС значительно упрощают реализацию различных приложений, таких как самоорганизующиеся распределенные антенные решетки и сети воздушных ретрансляторов [3]. Однако подобные системы потенциально уязвимы для злоумышленников [4].

Подходы к обеспечению информационной безопасности рассматривают ряд потенциальных атак на канал связи, в том числе спуфинг, отказ в обслуживании, «человек посередине» и другие [4]. Из-за динамической природы сети БВС, характеризующейся постоянным изменением узлов и маршрутов, сложно реализовать эффективные меры противодействия атакам с использованием радиоканала [4]. Тем не менее, в контексте полностью децентрализованного управления роем отпадает необходимость в обмене информацией по радиоканалу между БВС для оперативных целей [5, 6]. В такой ситуации нарушитель может нести другие угрозы, связанные с нанесением физического вреда БВС, увода группы от цели, ведением к ложной цели,

внесением возмущений в структуру за счет подмены локальной навигационной информации. Так, в работе [7] предложен метод спуфинга навигационных датчиков БВС в группе, что ведет к аварии, например, столкновению с препятствием. Более того, злоумышленник может проникнуть в рой не только во время полета, но и в качестве предварительно зараженного агента, внедренного в систему. Он может сохранять радиомолчание в течение всего полета и начать информационную или физическую атаку на наземную станцию или другой объект инфраструктуры, оставаясь до этого незамеченным.

Противодействие нарушителю уже в процессе полета можно разделить на два этапа: обнаружение и устранение, или уклонение. Существующие научные работы затрагивают именно обнаружение нарушителя, в том числе ряд работ, основанных на механизме доверия [8] или технологии распределенного реестра [9]. В работе [10] разработан метод обнаружения вторжений в роях БВС, основанный на применении временных вероятностных автоматов. В предложенном подходе анализируются состояния БВС и переходы между ними с учетом времени и вероятностей, сравнивая их с ожидаемыми значениями и выявляя аномалии в реальном времени. В [11] представлен метод детектирования GPS-спуфинга в группе БВС, сводящийся к проверке существования согласованного набора измеренных координат от всех агентов. В [12] применяются методы машинного обучения для обнаружения вторжения нарушителя в рой БВС.

В [13] рассмотрен метод уклонения роя БВС от нарушителя с помощью введения дополнительного потенциала (escape potential). БВС разделяются на два вида: активные — которые обнаружили нарушителя, и пассивные — соседи которых обнаружили нарушителя. Активные БВС начинают «отталкиваться» от нарушителя, а пассивные БВС следуют за ними.

В настоящей работе исследуется только этап устранения нарушителя. Разработанный метод не требует обнаружения нарушителя, несмотря на то, что он помогает значительно снизить энергетические затраты.

Целью работы является уменьшение деструктивного воздействия нарушителя на децентрализованную группу БВС без необходимости его обнаружения. Для выполнения поставленной цели работы разработан метод пространственного противодействия нарушителю. Идея метода возникла по аналогии с процессами удаления примесей и дефектов в кристаллах [14]. Агенты в рое взаимодействуют между собой посредством виртуальных потенциалов [5]. Так, каждый БВС представляет собой атом, которые вместе образуют кристалл. Рассмотрены следующие варианты реализации метода пространственного противодействия: сжатие, расширение и последовательное перестроение. Во всех вариантах происходит изменение параметров роевого взаимодействия и повышение средней кинетической энергии относительно центра масс роя БВС, что заставляет нарушителя передвигаться по структуре. Отметим, что в настоящей работе рассматривалось только двумерное движение БВС. Проведенное численное моделирование показало, что нарушитель, не обладающий информацией о временах перестроения, вытесняется на край при использовании варианта сжатия. Тем самым минимизируется его влияние на агентов группы.

Описание предлагаемого метода

Модель поведения роя. При проведении имитационного моделирования использовалась следующая модель. Движение БВС осуществляется в двумерном пространстве, а БВС представлено материальной точкой с массой m.

Результирующая сила, действующая на БВС:

$$\mathbf{F}_{res} = \mathbf{F}_{fr} + \mathbf{F}_{ext} + \mathbf{F}_{rnd} + \mathbf{F}_{th},$$

где $\mathbf{F}_{fr}$ — сила трения; $\mathbf{F}_{ext}$ — внешняя сила (например, ветер); $\mathbf{F}_{rnd}$ — некоторая случайная сила, связанная с БВС (например, из-за неточности настроек автопилота, неидеальности пропеллеров); $\mathbf{F}_{th}$ — управляющая сила.

Управляющая сила ограничена максимальной силой $\mathbf{F}_{th_max}$, так как физически ограничена тягой винтомоторной группы.

Управление роем БВС осуществляется с помощью метода, основанного на искусственных потенциалах [5, 15]. Каждый БВС формирует вокруг себя виртуальное потенциальное поле заданной формы, таким образом взаимодействуя с остальными БВС в группе. Зоны действия потенциалов можно разделить на три группы: ближнее отталкивание, ближнее притяжение и дальнее притяжение. Первые два описываются в виде потенциала Леннарда–Джонса [15], а между ними введена нейтральная зона [16]. Тогда сила ближнего притяжения и отталкивания:

$$\mathbf{F}^{(i,j)}_{rep_att}(\mathbf{r}) = \begin{cases} \dfrac{24\varepsilon}{\sigma}\left[2\left(\dfrac{\sigma}{|\mathbf{r}|}\right)^{13} - \left(\dfrac{\sigma}{|\mathbf{r}|}\right)^{7}\right]\dfrac{\mathbf{r}}{|\mathbf{r}|}, \\ |\mathbf{r}| < r_{rep_max} \text{ и } r_{att_min} < |\mathbf{r}| < r_{att_max}, \\ 0,\ r_{rep_max} < |\mathbf{r}| < r_{att_min} \end{cases}$$

где $\mathbf{r}$ — расстояние между i-м и j-м БВС; r_{rep_max} — максимальное расстояние действия силы отталкивания; r_{att_min} — минимальное расстояние действия силы ближнего притяжения; r_{att_max} — максимальное расстояние действия ближней силы отталкивания; σ — постоянная решетки; ε — глубина потенциальной ямы.

Силы ближнего отталкивания и притяжения действуют на коротком расстоянии и не способны собрать далеко расположенные БВС ($|\mathbf{r}| > 1{,}6\sigma$) в единую структуру. Для этого введено дальнее притяжение по аналогии с гравитационным взаимодействием:

$$\mathbf{F}^{(i,j)}_{far_att}(\mathbf{r}) = k_{far_attr} m_i m_j \frac{\mathbf{r}}{|\mathbf{r}|^3},\ |\mathbf{r}| > r_{att_max},$$

где k_{far_attr} — коэффициент дальнего притяжения; $m_{i(j)}$ — масса $i(j)$-го БВС (виртуальная).

Дальние и ближние силы не должны действовать одновременно, чтобы избежать уплотнения структуры [16].

Движение в гравитационных полях происходит по траекториям конического сечения, но для сборки БВС в структуру необходимо их падение на центр тела. Для

решения этой проблемы введено специальное правило: минимизация тангенциальной скорости между БВС с включенным дальним притяжением и точкой притяжения. Кроме того, для ближнего притяжения требуется минимизация относительной скорости между БВС и центром масс ближайших соседей. Для этого введена виртуальная сила трения:

$$\mathbf{F}_{vfr} = k_{vfr}|\mathbf{v}_{rel}|\mathbf{v}_{rel},$$

где k_{vfr} — коэффициент виртуальной силы трения; $\mathbf{v}_{rel}$ — относительная скорость движения БВС.

Результирующая сила, формирующая управляющее воздействие:

$$\mathbf{F}_{th_des} = \mathbf{F}_{rep_att} + \mathbf{F}_{far_att} + \mathbf{F}_{add} + \mathbf{F}_{vfr} + \mathbf{F}_{vrnd},$$

где $\mathbf{F}_{rep_att}$ — сила отталкивания и ближнего притяжения; $\mathbf{F}_{far_att}$ — сила дальнего притяжения; $\mathbf{F}_{add}$ — дополнительная сила (например, созданная потенциалом притяжения цели); $\mathbf{F}_{vfr}$ — виртуальная сила трения; $\mathbf{F}_{vrnd}$ — виртуальная случайная сила.

Случайная сила необходима, чтобы повысить скорость сборки или перестроений.

Итоговое управляющее воздействие ограничивается максимально возможной для БВС тягой $\mathbf{F}_{th_max}$.

Рой БВС определяется как множество агентов, взаимодействующих друг с другом только посредством обмена навигационной информацией, включающую расстояние, скорость и направление до ближайших соседей. Управление БВС в группе осуществляется децентрализованно, поэтому был выбран объектно-ориентированный подход к имитационному моделированию. Более того, это позволяет добавлять БВС-нарушитель со своими собственными параметрами.

Для предложенного метода структура группы имеет ромбическую решетку. Тогда количество ближайших соседей:

$$N_{nb}^{(i)} = \sum_{j=1}^{N} [d_{ij} < 1{,}6r_0],$$

где d_{ij} — расстояние между i-м и j-м БВС; r_0 — шаг решетки; N — количество БВС в группе.

Максимальное количество ближайших соседей при стабилизированной структуре равняется 6.

Для оценки состояния роя используются следующие показатели.

1. Затрачиваемая на движение суммарная энергия группы:

$$E_m = \int_0^T P_m dt = \int_0^T \sum_{i=1}^{N} \mathbf{F}_{fr_i}\mathbf{v}_i dt,$$

где $\mathbf{F}_{fr_i}$ — сила трения воздуха для i-го БВС; P_m — мощность всей группы; $\mathbf{v}_i$ — скорость i-го БВС; T — время полета. При проведении численных экспериментов E_m рассчитывается без учета нарушителя.

2. Средняя кинетическая энергия относительно центра масс БВС в рое:

$$T_s \sim \frac{1}{N}\sum_{i=1}^{N} m_i \mathbf{v}_{rel_i}^2,$$

где $\mathbf{v}_{rel_i}$ — скорость i-го БВС относительно центра масс всей группы.

Модель нарушителя. Нарушитель представляет собой БВС, обладающим таким же набором навигационных и коммуникационных аппаратно-программных средств, что и обычные БВС в группе. В частности, в децентрализованной группе БВС это могут быть датчики измерения расстояний, углов, относительных скоростей, уровня сигнала, камеры. Информации от подобных устройств должно быть достаточно для определения относительного положения и позиционирования в группе БВС. Более того, нарушитель не обладает информацией, определенной до начала миссии. В контексте рассматриваемого метода противодействия, такой информацией может быть период между перестроениями, параметры перестроения. Несмотря на отсутствие информации нарушитель, полагаясь только на датчики, для соседних агентов выглядит как обычный БВС из их группы.

Нарушитель может быть встроен в группу двумя способами. В первом случае злоумышленник может внедрить вредоносный код в обычный БВС между полетами. Во втором случае БВС-нарушитель может внедриться со стороны во время выполнения миссии группой.

Влияние нарушителя зависит от его местоположения относительно группы. Так, в центре он может нанести больший вред, чем находясь с краю. Это следует из того, что создаваемые нарушителем возмущения, распространяясь во все стороны, затрагивают большее количество агентов. Это влияет как на энергетические потери группы, так и возможность увести группу от точки интереса, вести ее к ложной цели, тормозить группу, т. е. вести по неправильному, с точки зрения полетного задания, маршруту.

Метод противодействия нарушителю. Метод противодействия нарушителю основан на аналогии анализа поведения роя с процессами, происходящими в полупроводниковых кристаллах при удалении примесей и дефектов. Так, атом в кристалле может быть замещен атомом примеси, размещаемым в узле решетки или в междоузлии. При этом также получается дефект, в первом случае называемый примесным атомом замещения, а во втором — примесным атомом внедрения. Тип дефекта зависит от размера примеси. В приложении к рою таким «примесным» БВС может являться агент с неисправными датчиками расстояния до соседних бортов, который неправильно оценивает расстояние между собой и соседями, либо нарушителем, не обладающим актуальными значениями алгоритма роевого взаимодействия. В технологиях выращивания кристаллов существует несколько методов избавления от примесей. Одна из них — отжиг, т. е. процесс, при котором воздействие температуры увеличивает скорость движения (диффузии) дефектов в кристалле и в течение времени отжига эти дефекты исчезают. Менее энергозатратный метод — зонная плавка [14], при которой происходит локальное плавление кристалла слой за слоем с последующей перекристаллизацией. В приложении к группе БВС это может быть реализовано как некоторое правило, которое сообщит каждому агенту вблизи дефекта некоторый случайный импульс вместе с уменьшением виртуальной силы трения, что

будет имитировать локальный перегрев. Дальнейшая релаксация пространственной структуры приведет к исчезновению дефектов, по аналогии с процессами, происходящими при росте кристаллов. Так как в данной работе не рассматривается этап обнаружения нарушителя, то осуществляется нагрев всего кристалла без локализации примеси.

Метод пространственного противодействия нарушителю заключается в изменении параметров взаимодействия БВС в рое и, тем самым, повышении средней кинетической энергии относительно центра масс БВС в рое T_s в заранее заданный момент времени τ_k на время $\Delta\tau$.

Были исследованы следующие реализации метода.

1. Сжатие. В заданный заранее момент времени τ_k на время $\Delta\tau$ все агенты роя уменьшают шаг решетки r_0, уменьшают коэффициент виртуальной силы трения k_{vfr}, увеличивают коэффициент виртуальной случайной силы k_{vrnd}, уменьшают глубину потенциальной ямы ε.
2. Расширение. В заданный заранее момент времени τ_k на время $\Delta\tau$ все агенты роя увеличивают шаг решетки r_0, уменьшают коэффициент виртуальной силы трения k_{vfr}, увеличивают коэффициент виртуальной случайной силы k_{vrnd}, уменьшают глубину потенциальной ямы ε.
3. Последовательное перестроение. В заданный заранее момент времени τ_k на время $\Delta\tau_z$ агенты роя, находящиеся в слое на краю, увеличивают шаг решетки r_0, уменьшают коэффициент виртуальной силы трения k_{vfr}, увеличивают k_{vrnd}, уменьшают глубину потенциальной ямы ε. Далее, в момент времени $\tau_k + \Delta\tau_z$ аналогичное изменение параметров происходит на следующем слое. Процедура повторяется для каждого следующего слоя до достижения слоя на противоположной стороне структуры. Далее процедура повторяется заново для всей группы N_z раз.

Стоит отметить, что варианты расширения и сжатия похожи и отличаются только изменением шага решетки r_0 в сторону увеличения или уменьшения, соответственно.

В качестве показателя результативности устранения нарушителя используется вероятность наличия у нарушителя менее пяти соседей в конце симуляции, что означает его нахождение на краю группы:

$$P_{edge}^{(intr)} = \frac{1}{M}\sum_{i=1}^{M}\left[N_{nb}^{(intr)} < 5\right],$$

где $N_{nb}^{(intr)}$ — количество ближайших соседей нарушителя; M — количество экспериментов.

Результаты и обсуждение экспериментальных исследований

Для тестирования предложенного метода в трех вариантах реализации было проведено имитационное моделирование (эксперимент 1 и 2). Во всех численных экспериментах заданы следующие исходные параметры: шаг решетки r_0 = 30 м, масса БВС m = 1 кг, максимальная тяга $\mathbf{F}_{th_max}$ = 19,6 Н. Начальные позиции БВС образуют ромбическую решетку (рис. 1, *a*). Нарушитель помещен в центр группы. Количество симуляций для каждого набора параметров равнялось 100. Время симуляции 90 с.

В эксперименте 1 изменялись параметры из набора $\{r_0; k_{vfr}; k_{vrnd}; \varepsilon; \Delta\tau; \Delta\tau_z; N_z\}$, а количество агентов зафиксировано N = 37. При таком количестве агентов структура симметрична относительно центра, а нарушитель помещен в центр. Это позволяет избежать выделенных направлений движения. В таблице представлены численные значения варьируемых параметров для каждого варианта реализации метода. Целью эксперимента 1 было получение таких наборов параметров, при которых $N_{nb}^{(i)} < 5$.

В эксперименте 2 варьировались следующие параметры: $\{N, \Delta\tau\}$ (таблица). Целью эксперимента 2 было найти зависимость необходимой длительности от количества агентов в рое N. Выбранный набор N обусловлен симметричностью структуры. Эксперимент 2 проводился только для расширения и сжатия, так как

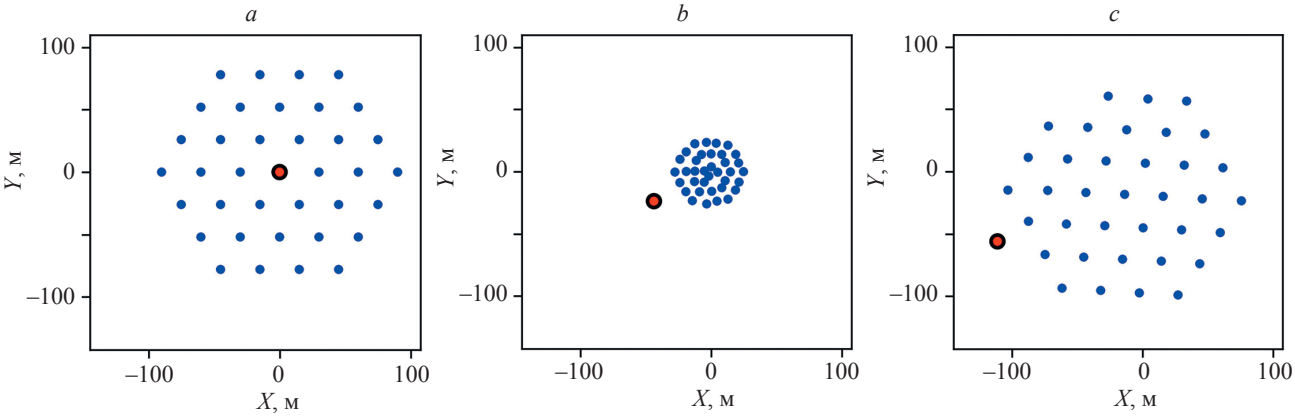

Рис. 1. Процесс изменения структуры роя и выталкивания нарушителя на край группы сжатием в момент времени симуляции: 0 с (*a*); 25 с (*b*); 75 с (*c*). Нарушитель обозначен красной точкой в черном круге. Изначальная структура группы сохранена

Fig. 1. The process of swarm structure changing and pushing the intruder to the edge of the group by compression at: 0 s (*a*); 25 s (*b*); 75 s (*c*). The intruder is marked with a red dot in a black circle. The original spatial structure of the group has been preserved

Таблица. Изменяемые параметры для каждого варианта реализации метода в двух численных экспериментах. Одинаковые наборы изменяемых параметров объединены. В случае неиспользования параметров в данном варианте реализации метода поставлен прочерк (—). Верхние индексы соответствуют параметрам до (I) и после (II) применения метода

Table. Variable parameters for each method realization in two numerical experiments. Identical sets of variable parameters are combined. A dashed (—) parameters are not used in the method. The superscripts correspond to the parameters before (I) and after (II) application of the method

Параметры	Эксперимент 1			Эксперимент 2	
	Сжатие	Расширение	Последовательное перестроение	Расширение	Сжатие
$r_0^{(I)}/r_0^{(II)}$	1; 1,25; ...; 3	(1; 1,5; …; 3)$^{-1}$	(1; 1,5; …; 3)$^{-1}$	2	0,5
$k_{vfr}^{(I)}/k_{vfr}^{(II)}$	1; 10; 100; 1000			1; 10	
$k_{vrnd}^{(I)}/k_{vrnd}^{(II)}$	1; 10; 100			1	
$\varepsilon^{(I)}/\varepsilon^{(II)}$	1; 10; 100			10; 100	
$\Delta\tau$, с	10; 20; 30; 40; 50		—	2; 4; ...; 60	
$\Delta\tau_z$, с	—		2; 5	—	
N_z	—		2; 5; 10	—	
N	37			19; 37; 61; 91	

последовательное перестроение в эксперименте 1 показал отрицательный результат.

По результатам эксперимента 1 для варианта сжатия нашлись наборы параметров, при которых $P_{edge}^{(intr)} = 1$, а наименее затратный с точки зрения E_m оказался набор параметров $\left\{ \frac{r_0^{(I)}}{r_0^{(II)}} = 1{,}7;\ \frac{k_{vfr}^{(I)}}{k_{vfr}^{(II)}} = 10;\ \frac{k_{vrnd}^{(I)}}{k_{vrnd}^{(II)}} = 1;\ \frac{\varepsilon^{(I)}}{\varepsilon^{(II)}} = 100;\ \Delta\tau = 50\text{ с} \right\}$ (рис. 1). Причем исходная структура группы сохраняется. Такой результат свидетельствует о том, что при сжатии происходит замораживание агентов в исходной структуре при одновременном уменьшении шага решетки r_0. Уменьшение глубины потенциальной ямы ε помогает снизить степень взаимодействия, тем самым сохраняя исходную структуру. При этом уменьшение виртуальной силы трения $\mathbf{F}_{vfr}$ помогает лишь незначительно уменьшить расходуемую энергию. При возвращении к исходной структуре в центре может оставаться незанятое пространство. Проведенный анализ также показал, что есть минимальная степень сжатия, при которой метод начинает работать, а расходуемая энергия E_m растет с увеличением степени сжатия (рис. 2).

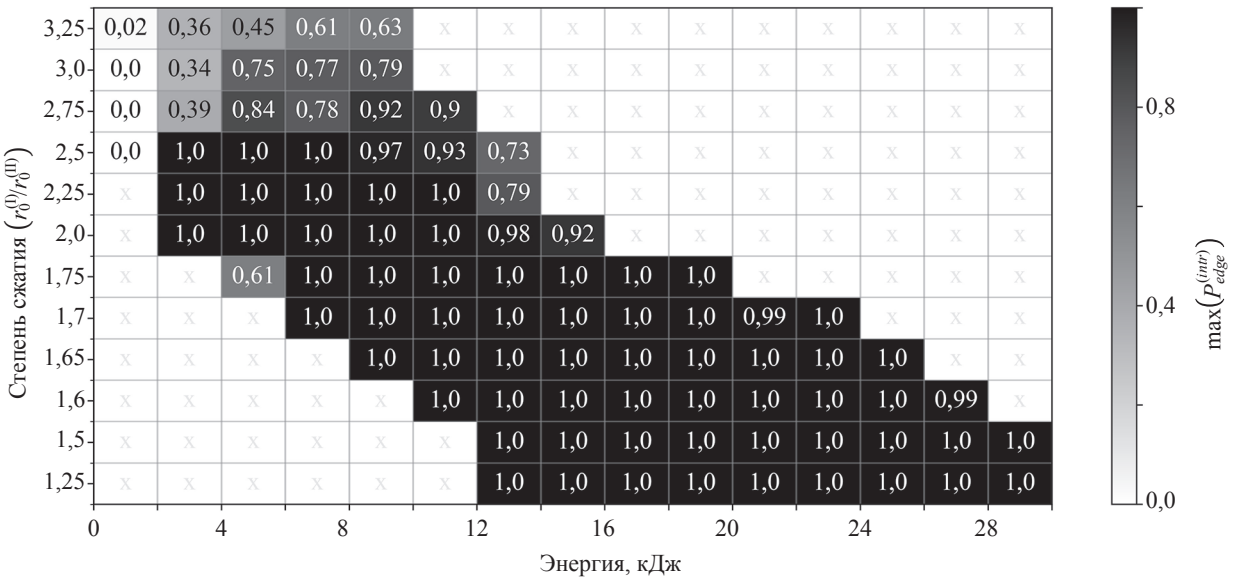

Рис. 2. Зависимость вероятности успеха метода в варианте сжатия $P_{edge}^{(intr)}$ от степени изменения шага решетки r_0 и средней затраченной энергии на движение каждым агентом в группе. По оси абсцисс отложены промежутки энергии в 2 кДж. Вероятность успеха вычисляется как максимальная вероятность при заданном $r_0^{(I)}/r_0^{(II)}$ на промежутке энергии $\left(E_m^i/E_m^{i+1}\right)$

Fig. 2. Dependence of the success probability of the compression method $P_{edge}^{(intr)}$ on the degree of change in the lattice pitch r_0 and the average energy expended on the movement of each UAV in the group. The *x*-axis shows energy intervals of 2 kJ. The success probability is calculated as the maximum probability for a given $r_0^{(I)}/r_0^{(II)}$ on the energy interval $\left(E_m^i/E_m^{i+1}\right)$

Для метода в варианте расширения не нашлось ни одного набора параметров, гарантирующих выталкивание нарушителя с $P_{edge}^{(intr)} = 1$, поэтому дальнейший анализ производился для экспериментов с $P_{edge}^{(intr)} > 0{,}5$. Наибольшая вероятность $P_{edge}^{(intr)} = 0{,}71$ была получена для набора $\left\{ \frac{r_0^{(\mathrm{I})}}{r_0^{(\mathrm{II})}} = 0{,}4;\ \frac{k_{vfr}^{(\mathrm{I})}}{k_{vfr}^{(\mathrm{II})}} = 1000;\ \frac{k_{vrnd}^{(\mathrm{I})}}{k_{vrnd}^{(\mathrm{II})}} = 1;\ \frac{\varepsilon^{(\mathrm{I})}}{\varepsilon^{(\mathrm{II})}} = 10;\ \Delta\tau = 50\ \text{с} \right\}$. С увеличением времени $\Delta\tau$ вероятность $P_{edge}^{(intr)}$ также увеличивалась. Нарушитель дрейфует от одного агента к другому, так как у него остался прежний шаг решетки. Этот процесс является вероятностным и поэтому трудно предугадать время окончания действия метода. Более того, для сжатия затрачивается меньше энергии на движение БВС, чем для расширения (рис. 3).

Метод в варианте реализации последовательного перестроения показал наихудший результат. В заданном диапазоне изменения параметров нарушитель не выталкивался на край структуры. Стоит сделать поправку, что при длительной работе метода (большом количестве

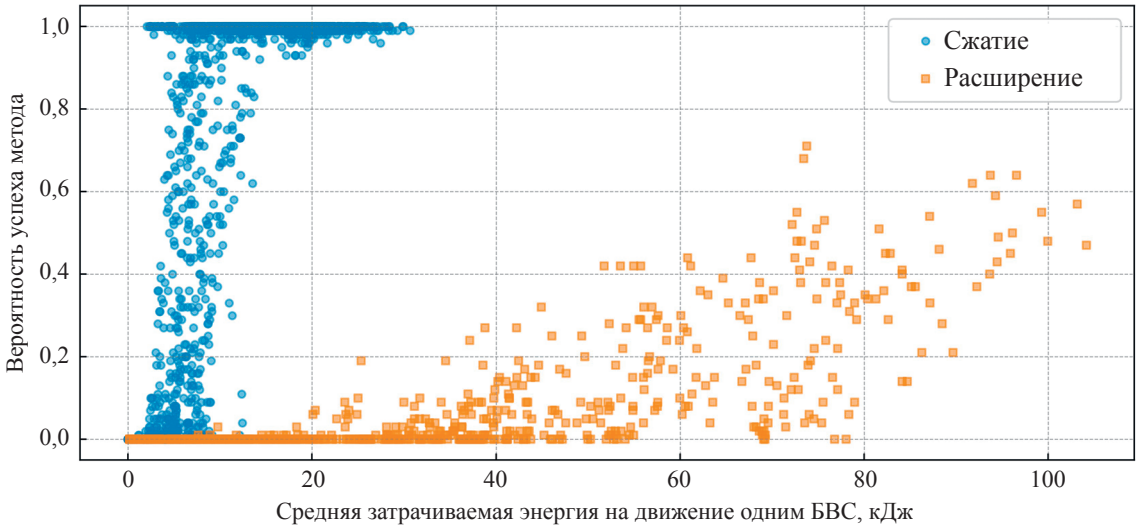

Рис. 3. Зависимость вероятности вытолкнуть нарушителя на край группы от средней затрачиваемой на движение энергии агента (за все время симуляции). Реализации сжатием и расширением разделены цветом. Разные точки обозначают разные наборы параметров

Fig. 3. Dependence of the probability of pushing an intruder to the edge of the group on the average energy spent on the UAV movement (for the entire simulation time). The compression and expansion methods are separated by color. Different dots indicate different parameter sets

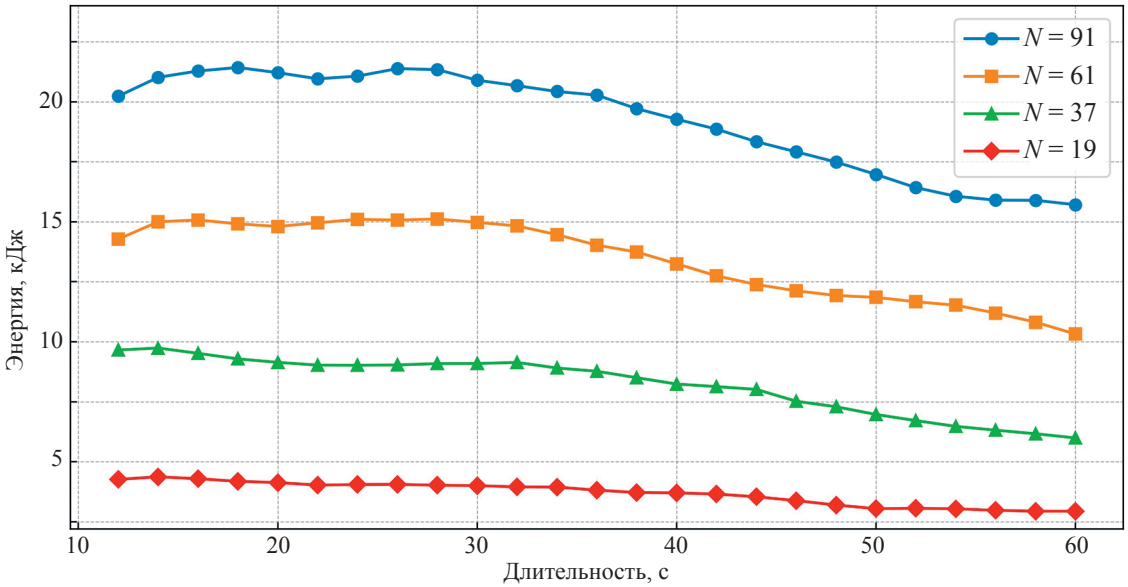

Рис. 4. Зависимость суммарной затрачиваемой энергии E_m от продолжительности действия метода $\Delta\tau$ для варианта реализации метода. Данные отфильтрованы для $P_{edge}^{(intr)} = 1$

Fig. 4. Dependence of the total expended energy E_m vs. the duration of the method $\Delta\tau$. Data is filtered by $P_{edge}^{(intr)} = 1$

N_z) и сильном изменении виртуальной силы трения k_{vfr} в более чем 100 раз, поведение становится похоже на расширение, а изначальная структура не сохраняется.

В эксперименте 2 работоспособность метода через сжатие была подтверждена для всех заданных $N = \{19; 37; 61; 91\}$. Для набора параметров $\left\{ \frac{r_0^{(\mathrm{I})}}{r_0^{(\mathrm{II})}} = 2; \frac{k_{vfr}^{(\mathrm{I})}}{k_{vfr}^{(\mathrm{II})}} = 1; \frac{k_{vrnd}^{(\mathrm{I})}}{k_{vrnd}^{(\mathrm{II})}} = 1; \frac{\varepsilon^{(\mathrm{I})}}{\varepsilon^{(\mathrm{II})}} = 100 \right\}$, при которых сохраняется первоначальная структура, суммарная затрачиваемая энергия E_m уменьшается при увеличении времени действия метода $\Delta\tau$ для всех N (рис. 4). Это может быть связано с тем, что структура группы успевает перестроиться в исходную структуру и вытесненный на край нарушитель уже вносит меньший вклад в колебания после возвращения к исходным параметрам. Для варианта расширения при различных N вероятность $P_{edge}^{(intr)}$ так же не достигает 1 для всех N.

Заключение

В работе представлен новый метод пространственного противодействия нарушителю в рое беспилотных воздушных судов. Противодействие агенту-нарушителю не требует его обнаружения и достигается за счет временного изменения агентами определенных параметров роевого взаимодействия.

Выполнено численное моделирование трех вариантов реализации метода: через сжатие, расширение и последовательное перестроение структуры роя (в количестве от 19 до 91 агентов). Работоспособность метода была подтверждена для варианта сжатия. Противодействие через расширение оказалось не гарантированно, а последовательное перестроение не дало положительных результатов. Для уверенной работы метода степень уменьшения шага решетки r_0 должна быть больше 1,65. С увеличением этого значения уменьшается средняя затрачиваемая на движение энергия. Более того, в случае сохранения изначальной структуры группы, средняя энергия, затрачиваемая агентом группы на движение, уменьшается с увеличением длительности работы метода.

Противодействие через расширение оказалось не гарантированно, а последовательное перестроение не дало положительных результатов.

Дальнейшие модификация и улучшение метода могут включать в себя рассмотрение нарушителя, который адаптируется к изменяющимся параметрам. Так, можно рассмотреть вариант поочередного сжатия и расширения, а также провести оптимизацию параметров метода относительно затрачиваемой на движение энергии. С другой стороны, обнаружение нарушителя поможет применять метод противодействия точечно, минимизируя затрачиваемую энергию во время полета. Также вызывает интерес противодействие нарушителю в трехмерной структуре роя беспилотных воздушных судов.

Для реализации предложенного метода необходимы исключительно локальные навигационные данные, что позволяет применять его в системах с ограниченными коммуникационными возможностями.

Литература

1. Phadke A., Medrano F.A., Sekharan C.N., Chu T. An analysis of trends in UAV swarm implementations in current research: simulation versus hardware // Drone Systems and Applications. 2024. V. 12. P. 1–10. https://doi.org/10.1139/dsa-2023-0099
2. Cetinsaya B., Reiners D., Cruz-Neira C. From PID to swarms: A decade of advancements in drone control and path planning — A systematic review (2013–2023) // Swarm and Evolutionary Computation. 2024. V. 89. P. 101626. https://doi.org/10.1016/j.swevo.2024.101626
3. Dao N.-N., Pham Q., Tu N., Thanh T.T., Bao V.N.Q., Lakew D.S., Cho S. Survey on aerial radio access networks: toward a comprehensive 6G access infrastructure // IEEE Communications Surveys & Tutorials. 2021. V. 23. N 2. P. 1193–1225. https://doi.org/10.1109/comst.2021.3059644
4. Wang X., Zhao Z., Yi L., Ning Z., Guo L., Richard Yu F., Guo S. A Survey on security of UAV swarm networks: attacks and countermeasures // ACM Computing Surveys. 2024. V. 57. N 3. P. 1–37. https://doi.org/10.1145/3703625
5. Boyko A., Girgidov R. Maintaining the spatial stability of a swarm of autonomous unmanned aerial vehicles // Robotics and Technical Cybernetics. 2021. V. 9. N 2. P. 85–90. https://doi.org/10.31776/rtcj.9201
6. Nie Z., Zhang Q., Wang X., Wang F., Hu T. Triangular lattice formation in robot swarms with minimal local sensing // IET Cyber-Systems and Robotics. 2023. V. 5. N 2. P. e12087. https://doi.org/10.1049/csy2.12087
7. Yao Y. (Elaine), Dash P., Pattabiraman K. Poster: may the swarm be with you: sensor spoofing attacks against drone swarms // Proc. of the ACM SIGSAC Conference on Computer and Communications Security. 2022. P. 3511–3513. https://doi.org/10.1145/3548606.3563535

References

1. Phadke A., Medrano F.A., Sekharan C.N., Chu T. An analysis of trends in UAV swarm implementations in current research: simulation versus hardware. *Drone Systems and Applications*, 2024, vol. 12, pp. 1–10. https://doi.org/10.1139/dsa-2023-0099
2. Cetinsaya B., Reiners D., Cruz-Neira C. From PID to swarms: A decade of advancements in drone control and path planning — A systematic review (2013–2023). *Swarm and Evolutionary Computation*, 2024, vol. 89, pp. 101626. https://doi.org/10.1016/j.swevo.2024.101626
3. Dao N.-N., Pham Q., Tu N., Thanh T.T., Bao V.N.Q., Lakew D.S., Cho S. Survey on aerial radio access networks: toward a comprehensive 6G access infrastructure. *IEEE Communications Surveys & Tutorials*, 2021, vol. 23, no. 2, pp. 1193–1225. https://doi.org/10.1109/comst.2021.3059644
4. Wang X., Zhao Z., Yi L., Ning Z., Guo L., Richard Yu F., Guo S. A Survey on security of UAV swarm networks: attacks and countermeasures. *ACM Computing Surveys*, 2024, vol. 57, no. 3, pp. 1–37. https://doi.org/10.1145/3703625
5. Boyko A., Girgidov R. Maintaining the spatial stability of a swarm of autonomous unmanned aerial vehicles. *Robotics and Technical Cybernetics*, 2021, vol. 9, no. 2, pp. 85–90. https://doi.org/10.31776/rtcj.9201
6. Nie Z., Zhang Q., Wang X., Wang F., Hu T. Triangular lattice formation in robot swarms with minimal local sensing. *IET Cyber-Systems and Robotics*, 2023, vol. 5, no. 2, pp. e12087. https://doi.org/10.1049/csy2.12087
7. Yao Y. (Elaine), Dash P., Pattabiraman K. Poster: may the swarm be with you: sensor spoofing attacks against drone swarms. *Proc. of the ACM SIGSAC Conference on Computer and Communications Security*, 2022, pp. 3511–3513. https://doi.org/10.1145/3548606.3563535

8. Зикратов И.А., Зикратова Т.В., Лебедев И.С. Доверительная модель информационной безопасности мультиагентных робототехнических систем с децентрализованным управлением // Научно-технический вестник информационных технологий, механики и оптики. 2014. Т. 90. № 2. С. 47–52.
9. Петренко В.И., Тебуева Ф.Б., Стручков И.В., Рябцев С.С. Модель доверенного взаимодействия агентов в децентрализованной киберфизической среде // Вестник Дагестанского государственного технического университета. Технические науки. 2023. Т. 50. № 2. С. 134–141. https://doi.org/10.21822/2073-6185-2023-50-2-134-141
10. Subbarayalu V., Vensuslaus M.A. An intrusion detection system for drone swarming utilizing timed probabilistic automata // Drones. 2023. V. 7. N 4. P. 248. https://doi.org/10.3390/drones7040248
11. Bi S., Li K., Hu S., Ni W., Wang C., Wang X. Detection and mitigation of position spoofing attacks on cooperative UAV swarm formations // IEEE Transactions on Information Forensics and Security. 2024. V. 19. P. 1883–1895. https://doi.org/10.1109/tifs.2023.3341398
12. Mughal U.A., Hassler S.C., Ismail M. Machine learning-based intrusion detection for swarm of unmanned aerial vehicles // Proc. of the IEEE Conference on Communications and Network Security (CNS). 2023. P. 1–9. https://doi.org/10.1109/cns59707.2023.10288962
13. Novák F. Walter V., Petráček P., Báča T., Saska M. Fast collective evasion in self-localized swarms of unmanned aerial vehicles // Bioinspiration and Biomimetics. 2021. V. 16. N 6. P. 066025. https://doi.org/10.1088/1748-3190/ac3060
14. Pfann W.G. Principles of zone-melting // Journal of Metals. 1952. V. 4. N 7. P. 747–753. https://doi.org/10.1007/bf03398137
15. Son J.-H., Ahn H.-S., Cha J. Lennard-jones potential field-based swarm systems for aggregation and obstacle avoidance // Proc. of the 17th International Conference on Control, Automation and Systems (ICCAS). 2017. P. 1068–1072. https://doi.org/10.23919/iccas.2017.8204374
16. Boyko A., Girgidov R. Key features of a swarm assemlby algorythm for autonomous unmanned aerial vehicles (UAVs) in absence of GNSS and stable radio communication // Robotics and Technical Cybernetics. 2022. V. 10. N 1. P. 25–31. https://doi.org/10.31776/rtcj.10103

8. Zikratov I.A., Zikratova T.V., Lebedev I.S. Trust model for information security of multi-agent robotic systems with a decentralized management. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2014, vol. 2, no. 90, pp. 47–52. (in Russian)
9. Petrenko V.I., Tebueva F.B., Struchkov I.V., Ryabtsev S.S. Model of trusted interaction of agents in decentralized cyber-physical environmentv. *Herald of Dagestan State Technical University. Technical Sciences*, 2023, vol. 50, no. 2, pp. 134–141. (in Russian). https://doi.org/10.21822/2073-6185-2023-50-2-134-141
10. Subbarayalu V., Vensuslaus M.A. An intrusion detection system for drone swarming utilizing timed probabilistic automata. *Drones*, 2023, vol. 7, no. 4, pp. 248. https://doi.org/10.3390/drones7040248
11. Bi S., Li K., Hu S., Ni W., Wang C., Wang X. Detection and mitigation of position spoofing attacks on cooperative UAV swarm formations. *IEEE Transactions on Information Forensics and Security*, 2024, vol. 19, pp. 1883–1895. https://doi.org/10.1109/tifs.2023.3341398
12. Mughal U.A., Hassler S.C., Ismail M. Machine learning-based intrusion detection for swarm of unmanned aerial vehicles. *Proc. of the IEEE Conference on Communications and Network Security (CNS)*, 2023, pp. 1–9. https://doi.org/10.1109/cns59707.2023.10288962
13. Novák F. Walter V., Petráček P., Báča T., Saska M. Fast collective evasion in self-localized swarms of unmanned aerial vehicles. *Bioinspiration and Biomimetics*, 2021, vol. 16, no. 6, pp. 066025. https://doi.org/10.1088/1748-3190/ac3060
14. Pfann W.G. Principles of zone-melting. *Journal of Metals*, 1952, vol. 4, no. 7, pp. 747–753. https://doi.org/10.1007/bf03398137
15. Son J.-H., Ahn H.-S., Cha J. Lennard-jones potential field-based swarm systems for aggregation and obstacle avoidance. *Proc. of the 17th International Conference on Control, Automation and Systems (ICCAS)*, 2017, pp. 1068–1072. https://doi.org/10.23919/iccas.2017.8204374
16. Boyko A., Girgidov R. Key features of a swarm assemlby algorythm for autonomous unmanned aerial vehicles (UAVs) in absence of GNSS and stable radio communication. *Robotics and Technical Cybernetics*, 2022, vol. 10, no. 1, pp. 25–31. https://doi.org/10.31776/rtcj.10103

Авторы

Шамрай Павел Юрьевич — ведущий инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 56493192000, https://orcid.org/0009-0006-0822-5310, pavel.shamray@mail.ru

Заколдаев Данил Анатольевич — кандидат технических наук, доцент, директор международного научно-образовательный центра «Безопасность и надежность критических цифровых технологий», Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57021875400, https://orcid.org/0000-0002-2520-1998, d.zakoldaev@mail.ru

Бойко Андрей Михайлович — кандидат физико-математических наук, научный сотрудник, Физико-технический институт им. А.Ф. Иоффе, Санкт-Петербург, 194021, Российская Федерация; старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0000-0002-3120-6170, andray599@gmail.com

Authors

Pavel Yu. Shamray — Leading Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 56493192000, https://orcid.org/0009-0006-0822-5310, pavel.shamray@mail.ru

Danil A. Zakoldaev — PhD, Associate Professor, Head of the International Scientific and Academic Center “Security and Safety for Critical Digital Technologies”, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57021875400, https://orcid.org/0000-0002-2520-1998, d.zakoldaev@mail.ru

Andrey M. Boyko — PhD (Physics & Mathematics), Scientific Researcher, Ioffe Institute, Saint Petersburg, 194021, Russian Federation; Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0000-0002-3120-6170, andray599@gmail.com

Статья поступила в редакцию 17.06.2025
Одобрена после рецензирования 22.09.2025
Принята к печати 12.11.2025

Received 17.06.2025
Approved after reviewing 22.09.2025
Accepted 12.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1134-1141

ReflectivePrompt: Reflective evolution in autoprompting algorithms

Viktor N. Zhuravlev[1], Artur R. Khairullin[2], Ernest A. Dyagin[4], Alena N. Sitkina[4], Nikita I. Kulin[5]✉

[1,2,3,4,5] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] viktor.zhuravlev2003@mail.ru, https://orcid.org/0009-0009-0788-9790
[2] arkhairullin@itmo.ru, https://orcid.org/0009-0007-0442-7764
[3] tsenred@yandex.com, https://orcid.org/0009-0000-3865-4446
[4] sitkina.alena2017@yandex.ru, https://orcid.org/0009-0002-6046-8943
[5] kylin98@list.ru✉, https://orcid.org/0000-0002-3952-6080

Abstract

Autoprompting is the process of automatically selecting optimized prompts for language models, which has been gaining popularity with the rapid advancement of prompt engineering driven by extensive research in the field of Large Language Models. This paper presents ReflectivePrompt — a novel autoprompting method based on evolutionary algorithms that employs a reflective evolution approach for more precise and comprehensive search of optimal prompts. ReflectivePrompt utilizes short-term and long-term reflection operations before crossover and elitist mutation to enhance the quality of the modifications they introduce. This method allows for the accumulation of knowledge obtained throughout the evolution process and updates it at each epoch based on the current population. ReflectivePrompt was tested on 33 datasets for classification and text generation tasks using open-access large language models: T-lite-instruct-0.1 and Gemma3-27b-it. The method demonstrates, on average, a significant improvement (e.g., 28 % on BBH compared to EvoPrompt) in metrics relative to current state-of-the-art approaches, thereby establishing itself as one of the most effective solutions in evolutionary algorithm-based autoprompting.

Keywords

LLM, autoprompting, evolutionary algorithms, reflective evolution, prompt engineering

For citation: Zhuravlev V.N., Khairullin A.R., Dyagin E.A., Sitkina A.N., Kulin N.I. ReflectivePrompt: Reflective evolution in autoprompting algorithms. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1134–1141. doi: 10.17586/2226-1494-2025-25-6-1134-1141

УДК 004.89

ReflectivePrompt: использование рефлексивной эволюции в алгоритмах автопромптинга

Виктор Николаевич Журавлев[1], Артур Рустамович Хайруллин[2], Эрнест Александрович Дягин[3], Алена Николаевна Ситкина[4], Никита Игоревич Кулин[5]✉

[1,2,3,4,5] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] viktor.zhuravlev2003@mail.ru, https://orcid.org/0009-0009-0788-9790
[2] arkhairullin@itmo.ru, https://orcid.org/0009-0007-0442-7764
[3] tsenred@yandex.com, https://orcid.org/0009-0000-3865-4446
[4] sitkina.alena2017@yandex.ru, https://orcid.org/0009-0002-6046-8943
[5] kylin98@list.ru✉, https://orcid.org/0000-0002-3952-6080

Аннотация

Автопромптинг — процесс автоматического подбора оптимизированных промптов к языковым моделям, который набирает свою популярность с быстрым развитием промпт-инжиниринга, обусловленного многочисленными исследованиями в области больших языковых моделей. В работе представлен ReflectivePrompt — новый метод автопромптинга на основе эволюционных алгоритмов, использующий подход рефлексивной эволюции для более точного и расширенного поиска оптимальных промптов. ReflectivePrompt использует операции

© Zhuravlev V.N., Khairullin A.R., Dyagin E.A., Sitkina A.N., Kulin N.I., 2025

краткосрочной и долгосрочной рефлексии перед операциями скрещивания и элитарной мутации для повышения качества проводимых ими изменений. Предложенный метод позволяет накапливать знания, полученные на протяжении всей эволюции, и обновлять их на каждой эпохе исходя из текущей популяции. ReflectivePrompt был протестирован на 33 наборах данных по задачам классификации и генерации текста с использованием больших языковых моделей с открытым доступом: T-lite-instruct-0.1, Gemma3-27b-it. Представленный метод продемонстрировал значительное увеличение (например, 28 % в среднем на бенчмарке Big-Bench-Hard относительно EvoPrompt) по метрикам в сравнении с известными методами в данной области, тем самым показал себя одним из самых эффективных вариантов в рамках автопромптинга на основе эволюционных алгоритмов.

Ключевые слова
большие языковые модели, автоматический промптинг, эволюционные алгоритмы, рефлексивная эволюция, промпт-инжиниринг

Ссылка для цитирования: Журавлев В.Н., Хайруллин А.Р., Дягин Э.А., Ситкина А.Н., Кулин Н.И. ReflectivePrompt: использование рефлексивной эволюции в алгоритмах автопромптинга // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1134–1141 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-6-1134-1141

Introduction

Large Language Models (LLMs) have demonstrated significant results in solving Natural Language Processing (NLP) tasks [1, 2]. Prompting and prompt engineering are universal methods for improving the performance of LLMs that do not require access to model weights and gradients during training. Instead, they enhance the efficiency of LLM inference by providing carefully crafted and well-structured instructions (prompts) as input to the model [3]. Currently, there are many different prompting techniques, such as Few-Shot [4], Role-Based [5], Chain-of-Thought [6], Plan-and-Solve [7], and others. What all these techniques have in common is that they can be time-consuming to manually create, iterate, and optimize, often requiring expert knowledge and experience. The reason for this is that models are highly sensitive to input data, necessitating careful and precise application of these techniques [8].

Autoprompting addresses this issue by automating the generation and selection of prompts [9]. It is based on various optimization methods and principles, including reinforcement learning, evolutionary, gradient-based, and gradient-free approaches, among others [9–12]. In particular, prompt optimization can be either discrete or continuous [13]. Continuous optimization involves representing the prompt as a numerical tensor, while discrete optimization treats the prompt as a sequence of tokens. The latter approach offers several advantages: it does not require derivative computations, meaning there is no need to access the model internal parameters and gradients. This allows working with black-box models and avoids additional computational overhead [14]. Additionally, it preserves prompt interpretability, enabling humans to analyze and edit them [14], and allows optimization for any metric (including non-differentiable ones) [11, 15–17].

However, this approach also has challenges: the optimization space for prompts is vast, and prompts generated through search methods may lack diversity [11]. Nevertheless, there are numerous heuristic optimization algorithms that employ stochastic strategies, making the optimization process less sensitive to local optima. Evolutionary algorithms are one such example [18].

In this work, we analyzed the Reflective Evolution algorithm [19] and integrated it to address the problem of automatic prompt generation. The resulting solution, called ReflectivePrompt, was tested on 33 datasets to demonstrate its effectiveness compared to existing methods.

Evolutionary algorithms

Evolutionary algorithms are a family of optimization methods based on the principles of biological evolution: natural selection, mutation, crossover, and inheritance. These algorithms operate with populations of solutions, gradually improving them according to a given fitness function [18]. Among such algorithms, the genetic algorithm [20] can be distinguished which works with gene sequences (in our case, sequences of phrases in prompts). Within this algorithm, starting with an initial population of individuals, selection, crossover of selected individuals (creating offspring based on a combination of parental information), mutation of the offspring (random modification of certain parts), and population update based on offspring evaluations are performed iteratively.

This approach is highly flexible when applied to problems from various domains. The usage of evolutionary operators (crossover, mutation) and population-based search reduces the risk of getting stuck in local optima, maintaining a balance between exploring new solutions and exploiting existing ones, thereby leading to a high diversity of individuals in the final population while ensuring their quality according to the objective function remains high [11, 15–17].

Related works

One solution employing genetic algorithms is EvoPrompt [11]. The improvement of the candidate prompt population occurs iteratively through selection, evolution (generation of new candidates using evolutionary operators), and population updates based on the evaluation of new candidates. The implementation of evolutionary operators (mutation and crossover) is achieved through queries to an LLM, enabling the utilization of its expertise in solving NLP tasks while maintaining prompt readability. During new candidate generation, two parents are first selected from the previous population using roulette-wheel selection (selection phase) [21] followed by the application of crossover and subsequent mutation of the resulting offspring. The study also presents a differential evolution algorithm [22], which involves mutating different segments

of two donor prompts from the same population, combining them with a mutating candidate, and performing crossover with the current best prompt. The authors' position EvoPrompt as a general framework for integrating LLMs into evolutionary algorithms, with experimental results demonstrating that differential evolution exhibits superior performance on more complex tasks.

Semantic Prompt Evolution based on a LLM (SPELL) [15] employs a genetic algorithm operating iteratively through repeated reproduction and selection steps where selection is performed via roulette-wheel while reproduction involves generating offspring based on a list of parent prompts and their corresponding scores. Notably, although reproduction is also conducted through LLM queries, this solution lacks explicit separation between crossover and mutation. Instead, it utilizes predefined prompt instructing modifications to the parent prompt set (replacing, adding, or deleting words, altering tone) to generate offspring.

An alternative approach implemented in Plum [16] is based on metaheuristics. Unlike EvoPrompt prompt mutation methodology, Plum explicitly defines a set of prompt modification operations: adding, deleting, rephrasing words/phrases, or swapping their positions, thereby generating multiple neighboring prompts. The solution architecture comprises: a well-defined set of neighboring prompts for each prompt, a metaheuristic algorithm with its inherent hyperparameters, and auxiliary functions (including crossover). The study examines six algorithms: hill climbing [23], simulated annealing [24], genetic algorithm (two variants – with mutation and crossover, and mutation-only) [20], tabu search [25], and harmony search [26]. Each algorithm performs candidate mutation through the application of predefined modification operations, enabling exploration within the discrete prompt space. It should be noted that only the rephrasing operation is executed via LLM queries, while other operations are performed manually. As described by the authors, experimental results demonstrate this approach capability to identify novel structural prompt modifications that enhance performance.

In Promptbreeder [17] paper, the authors propose an extended genetic algorithm mutation approach incorporating predefined mutation prompts and "thinking styles" (concise descriptions of cognitive strategies, e.g., "Let's think step by step"), in addition to utilizing Chain-of-Thought [6] and Plan-and-Solve [7] techniques. At each iteration, candidates are improved through the application of a randomly selected mutation from a uniform distribution. The authors identify five mutation classes: direct mutation, hypermutation, estimation of distribution mutation [27], Lamarckian mutation [28], and prompt crossover/context shuffling. The first two classes further include zero-order and first-order mutations, totaling ten distinct mutations, each implemented through LLM queries. First-order direct mutation modifies candidates using specific mutation prompts, while zero-order mutation utilizes the initial problem statement to address method divergence. When problem specifications lack precision, Lamarckian mutation facilitates prompt reconstruction based on the last output yielding correct results. The algorithm key innovation involves hypermutation which modifies the mutation prompts themselves (via hypermutation prompts), thereby enhancing not only prompt solutions but also the improvement mechanisms. According to the authors, this diversity of operators enables continuous reformulation and representation of problems by LLMs, leading to more effective solutions [17]. This approach demonstrates adaptability across various domains while optimizing prompts and preserving their interpretability.

Proposed method: ReflectivePrompt

Reflective Evolution

Reflective Evolution is an approach described in the article Ye H. et al. [19]. Its essence lies in using a language model to generate prompts aimed at enhancing the efficiency of mutation and crossover operations. The processes of prompt creation are referred to as short-term and long-term reflection. According to the authors, such reflective actions can be interpreted as obtaining a "verbal gradient" within the prompt space. Short-term reflection involves generating crossover prompts based solely on the current parent population, while long-term reflection, as the name suggests, entails accumulating knowledge, dependencies, and methods for improving efficiency throughout the entire evolutionary operation.

The application of reflection helps guide the direction of mutation and crossover operations while also expanding the search space, potentially moving beyond the initial population's predefined prompt space. In the original article, this approach was successfully applied to solving problems, such as Guided Local Search [29], Ant Colony Optimization [30], Electronic Design Automation [31], the Decap Placement Problem [32], the Traveling Salesman Problem [33], and other combinatorial optimization tasks.

Proposed solution

In this study we developed a novel approach that combines methods of Reflective Evolution with large language models for the automatic generation of higher-quality prompts — ReflectivePrompt. ReflectivePrompt employs short-term and long-term reflection operations for subsequent use in crossover and elitist mutation. All performed operations and their corresponding queries to the language model were modified and refined to directly optimize prompts.

Specifically, the beginning of each instruction was changed to: "You are an expert in the domain of optimization prompts. Your task is to give hints to design better prompts". This adjustment is motivated by the specifics of the autoprompting task for which reflective evolution was applied. Techniques describing possible modification operations performed during crossover and mutation, previously used in the SPELL algorithm, were incorporated. Thus, the following was added to the model queries defining short-term and long-term reflection: "For example, you can try to recommend word replacements, active/positive voice conversions, adding words, or deleting words". This enables the LLM to generate more precise, well-described hints that affect not only the semantic content of prompts but also their structural aspects.

ReflectivePrompt simplifies user interaction by generating an initial population of prompts based on just a

single input prompt. In this approach, the prompt is rephrased using the LLM and structured output techniques [34].

A key feature of ReflectivePrompt is delegating the decision on the specifics of mutation to the model itself. In previously described solutions, the mutation type was either predefined and fixed or randomly selected from a uniform distribution. In this approach, however, the model generates hints autonomously and tends to decide whether to apply structural transformations to the prompt or only modify its semantic meaning and phrasing.

When performing crossover and mutation operations, the LLM is provided with a brief task description which helps generate more problem-targeted prompts while preserving the logical structure of the instruction. Empirical observations have shown that even large models can achieve decent metric values using prompts that are partially or entirely irrelevant to the task. As a result, the final prompts may deviate significantly from the intended meaning. ReflectivePrompt avoids this issue and, in the vast majority of cases, generates semantically correct prompts that are more comprehensible to human perception and logic.

The general scheme of Reflective Evolution within ReflectivePrompt is illustrated in Fig. 1.

Particular attention should be paid to the two selection operations. The parent population selection of prompts chooses pairs of parent prompts from the current population. In this process, each prompt can be included in multiple parent pairs. The main constraint, which is related to the original Reflective Evolution algorithm, is that prompts in a parent pair must have different fitness function values. Parent selection is performed using the roulette-wheel method [21]. The probability vector for being selected for each individual is represented by the normalized vector of their fitness scores. The second selection operation, which mimics the survival of the fittest, also employs the roulette-wheel method, but in this case, the probabilities are obtained by applying a softmax operation with a temperature of 0.1 to the fitness function value vector. This temperature value yields a less uniform distribution in cases where all prompts have approximately similar scores, thereby increasing the probability of selection for individuals with higher fitness values.

Another crucial aspect is the preservation of elite individuals in the population. Before the start of each epoch, the individual that has demonstrated the best performance throughout the entire evolutionary process is reintroduced into the population, even if it was not selected at the end of the previous iteration. This approach enhances the algorithm convergence speed, as the best individuals are not lost over time due to unfavorable selection outcomes.

The examples of ReflectivePrompt optimization are shown below.

Example 1. The optimized prompt for SST-2 dataset:

Initial prompt:

"Please provide a Sentiment Classification task."

Optimized prompt:

"Could you please conduct a Sentiment Classification exercise on movie reviews?"

Example 2. The optimized prompt for BBH/logical_deduction_three_objects dataset:

Initial prompt:

"Identify the position of each object in a fixed order based on the given statements in each input-output pair."

Optimized prompt:

"Determine the order of three objects based on the given statements. Each paragraph will contain three objects and a set of statements describing theirproperties, sizes, or positions. You will be provided with a list of statementsabout the objects, and your task is to analyze these statements to deduce thecorrect order of the objects.

For each paragraph, you will find three objects and a set of statements. These statements will describe the objects dimensions, weights, or positions relative to each other. Use this information to deduce the correct order of the objects. For example, consider the paragraph below:

1. The object on the left is smaller than the object in the middle.

2. The object in the middle is larger than the object on the right.

3. The object on the right is transparent.

Based on these statements, determine the correct order of the objects. The possible orderings are:

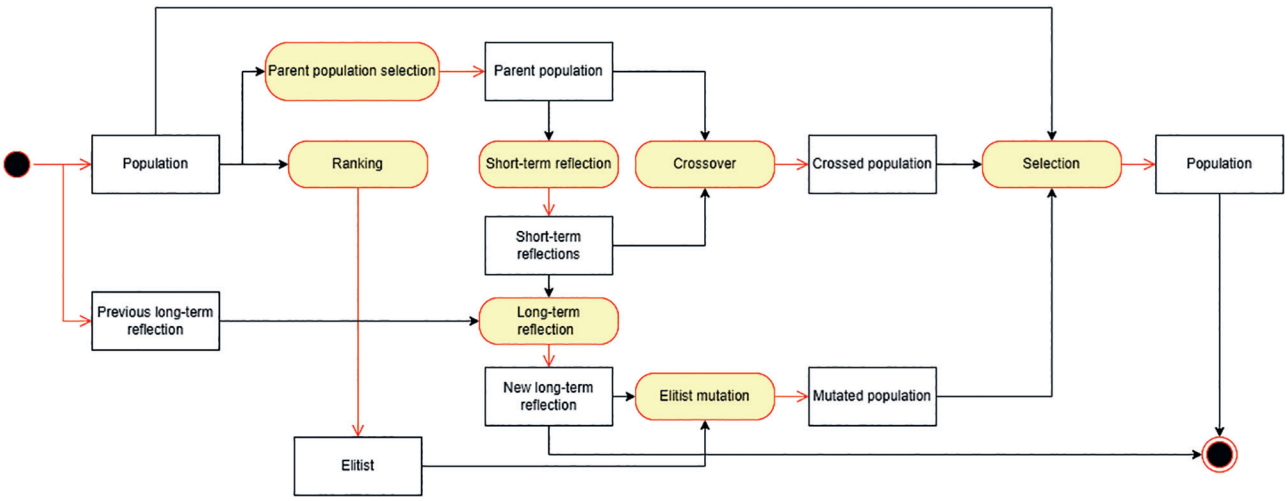

Fig. 1. The Reflective Evolution pipeline in ReflectivePrompt

A. Left, Middle, Right
B. Middle, Left, Right
C. Right, Middle, Left
D. Right, Left, Middle

Please, select the correct order from the options provided."

Experimental Results

ReflectivePrompt was evaluated on 33 datasets for text classification and generation tasks. As baselines and reference points for comparison, we used results from EvoPrompt, SPELL, PromptBreeder, and Plum. The autoprompting algorithms were executed using large language models from different families and sizes (T-lite-instruct-0.1[1], Gemma3-27b-it [35]). This choice of LLMs was made due to the use of open source white-box models which are more user-friendly and can be utilized by everyone. Also the significant difference in the number of model parameters leads to better testing coverage and makes our results more unbiased.

[1] T-lite-instruct-0.1 — Hugging Face, 2025. URL: https://huggingface.co/AnatoliiPotapov/T-lite-instruct-0.1 (accessed: 19.06.2025).

Classification tasks

For classification tasks, the following datasets and benchmarks were used: MNLI, MR, SST-2, YAHOO, and BBH (a subset of datasets with strictly formatted answers that can be treated as classification tasks). The metric selected for evaluation and optimization during evolution was the F1-score. The results of each method are presented in Fig. 2.

Generation tasks

ReflectivePrompt and its counterparts were evaluated on the following datasets: BBH (dyck_languages, multistep_arithmetic_two, object_counting, word_sorting), GSM8K, and SamSUM. The metric used for evaluation and optimization was METEOR. The main results are shown in Fig. 3.

Discussion

The conducted experiments demonstrate that ReflectivePrompt effectively handles both classification and text generation tasks. Across all evaluated datasets, ReflectivePrompt either outperformed or matched the performance of existing evolutionary algorithm-based autoprompting methods. The method showed particularly strong results on the BBH benchmark, comprising 23 classification tasks and 4 text generation tasks. For

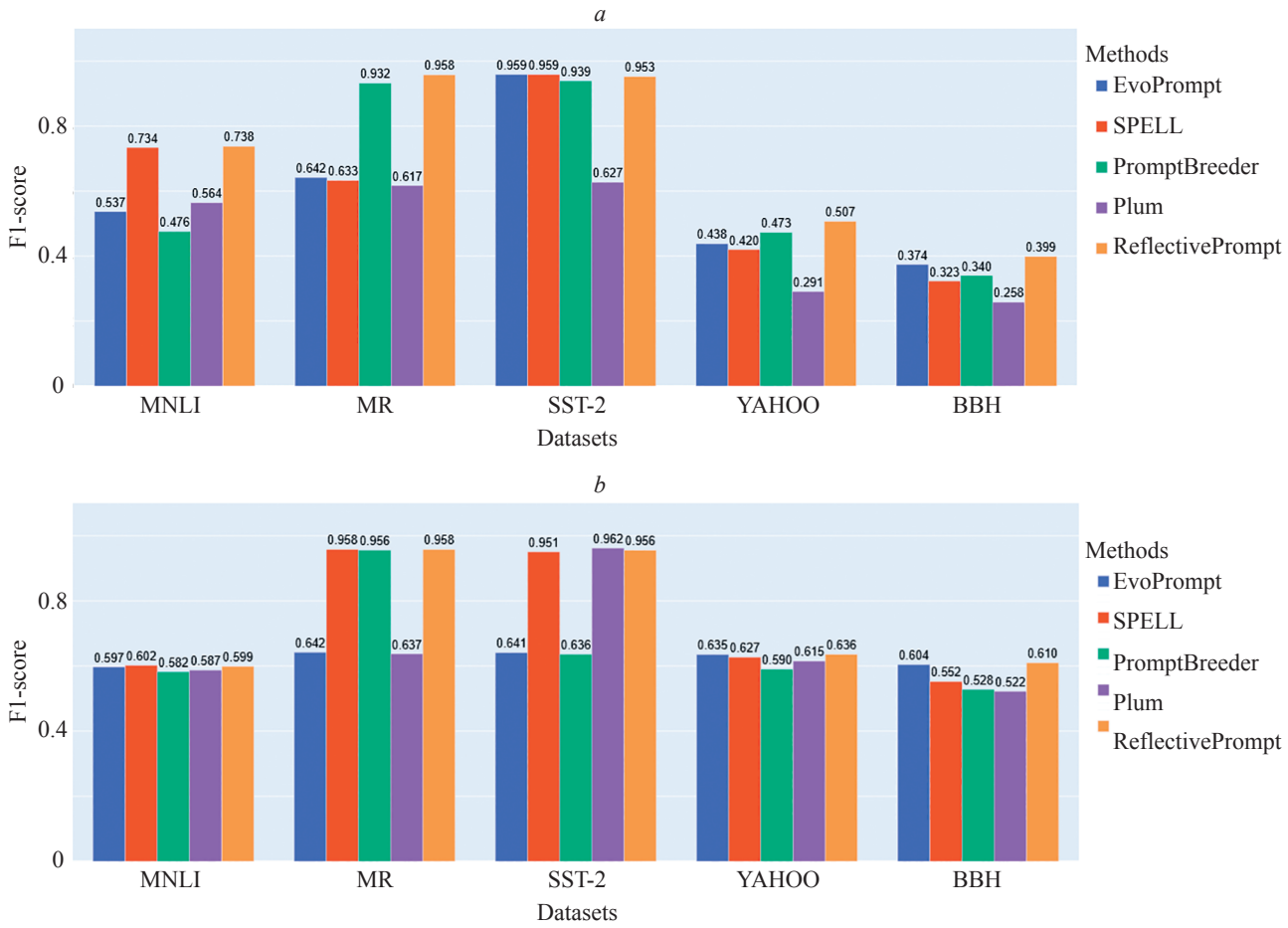

Fig. 2. Histogram of F1-score values.
Modes T-lite-instruct-0.1 (*a*) and Gemma3-27b-it (*b*)

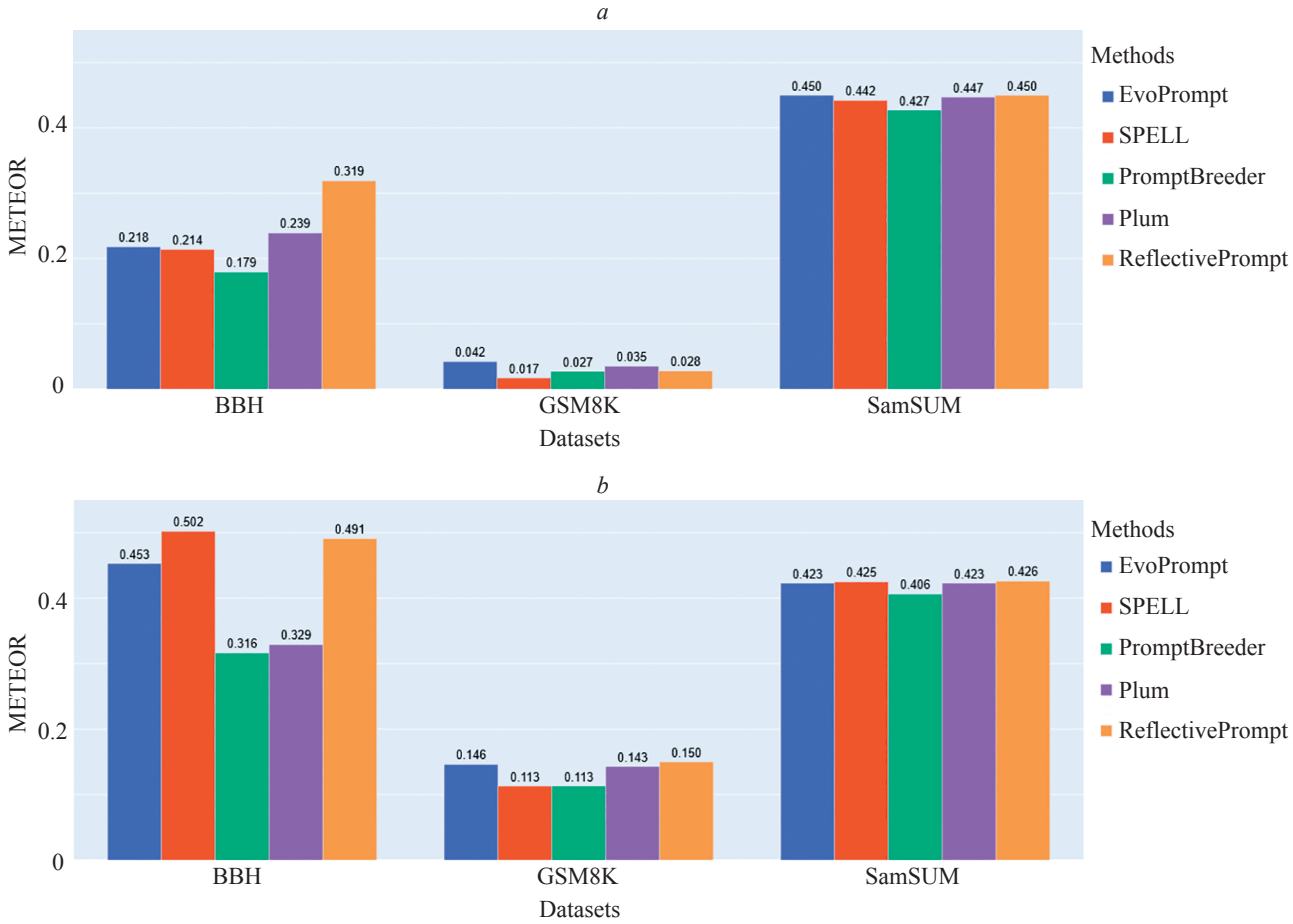

Fig. 3. Histogram of METEOR scores for text generation datasets.
Modes T-lite-instruct-0.1 (*a*) and Gemma3-27b-it (*b*)

classification tasks, the average F1-score improved by 6.59 % on the T-lite-instruct-0.1 model and by 0.96 % on the Gemma3-27b-it model. In text generation tasks, the average METEOR score increased by 33.34 % on the T-lite-instruct-0.1 model (comparisons and improvements were calculated relative to the maximum average metrics achieved by existing solutions).

It should be noted that ReflectivePrompt performance significantly depends on the underlying LLM. The effectiveness of Reflective Evolution relies on the quality of generated hints, and weaker language models may produce suggestions that are not fully relevant to the optimization task.

This work creates a scope for future research into Reflective Evolution for autoprompting applications. The current ReflectivePrompt implementation could potentially be further refined for more targeted prompt optimization. Moreover, the concept of Reflective Evolution could be generalized and adapted to other metaheuristic optimization algorithms, representing a promising direction for future studies. For example, there was a recent research where ReflectivePrompt evolution outperforms reinforcement learning on a group of benchmarks [36].

Conclusion

The proposed ReflectivePrompt algorithm, which employs Reflective Evolution for prompt optimization, was evaluated on 33 datasets covering various natural language processing domains. It demonstrated consistent improvements over existing evolutionary algorithm-based autoprompting methods. ReflectivePrompt proves to be a competitive solution, showing that exploring Reflective Evolution for autoprompting can yield significant benefits and advance current methods to new levels of performance.

References

1. Kadavath S., Conerly T., Askell A., Henighan T., Drain D., Perez E., et al. Language models (mostly) know what they know. *arXiv*, 2022, arXiv:2207.05221. https://doi.org/10.48550/arXiv.2207.05221
2. Wei J., Bosma M., Zhao V.Y., Guu K., Yu A.W., Lester B., et al. Finetuned language models are zero-shot learners. *arXiv*, 2021, arXiv:2109.01652. https://doi.org/10.48550/arXiv.2109.01652

Литература

1. Kadavath S., Conerly T., Askell A., Henighan T., Drain D., Perez E., et al. Language models (mostly) know what they know // arXiv. 2022. arXiv:2207.05221. https://doi.org/10.48550/arXiv.2207.05221
2. Wei J., Bosma M., Zhao V.Y., Guu K., Yu A.W., Lester B., et al. Finetuned language models are zero-shot learners // arXiv. 2021. arXiv:2109.01652. https://doi.org/10.48550/arXiv.2109.01652

3. Liu P., Yuan W., Fu J., Jiang Z., Hayashi H., Neubig G. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 2023, vol. 55, no. 9, pp. 1–35. https://doi.org/10.1145/3560815
4. Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., et al. Language models are few-shot learners. *arXiv*, 2020, arXiv:2005.14165. https://doi.org/10.48550/arXiv.2005.14165
5. Wang N., Peng Z., Que H., Liu J., Zhou W., Wu Y., et al. RoleLLM: benchmarking, eliciting, and enhancing role-playing abilities of large language models. *Proc. of the Annual Meeting of the Association for Computational Linguistics*, 2024, pp. 14743–14777. https://doi.org/10.18653/v1/2024.findings-acl.878
6. Wei J., Wang X., Schuurmans D., Bosma M., Ichter B., Xia F., et al. Chain-of-thought prompting elicits reasoning in large language models. *arXiv*, 2022, arXiv:2201.11903. https://doi.org/10.48550/arXiv.2201.11903
7. Wang L., Xu W., Lan Y., Hu Z., Lan Y., Lee R.K.-W., Lim E.-P. Plan-and-solve prompting: improving zero-shot chain-of-thought reasoning by large language models. *Proc. of the 61st Annual Meeting of the Association for Computational Linguistics*, 2023, vol. 1, pp. 2609–2634. https://doi.org/10.18653/v1/2023.acl-long.147
8. Leidinger A., van Rooij R., Shutova E. The language of prompting: What linguistic properties make a prompt successful?. *Proc. of the Findings of the Association for Computational Linguistics: EMNLP*, 2023, pp. 9210–9232. https://doi.org/10.18653/v1/2023.findings-emnlp.618
9. Shin T., Razeghi Y., Logan R.L., Wallace E., Singh S. AutoPrompt: Eliciting knowledge from language models with automatically generated prompts. *Proc. of the Conference on Empirical Methods in Natural Language*, 2020, pp. 4222–4235. https://doi.org/10.18653/v1/2020.emnlp-main.346
10. Kwon M., Kim G., Kim J., Lee H., Kim J. StablePrompt: automatic prompt tuning using reinforcement learning for large language models. *arXiv*, 2024, arXiv:2410.07652. https://doi.org/10.48550/arXiv.2410.07652
11. Guo Q., Wang R., Guo J., Li B., Song K., Tan X., et al. EvoPrompt: Connecting large language models with evolutionary algorithms yields powerful prompt optimizers. *arXiv*, 2023, arXiv:2309.08532. https://doi.org/10.48550/arXiv.2309.08532
12. Prasad A., Hase P., Zhou X., Bansal M. GrIPS: gradient-free, edit-based instruction search for prompting large language models. *Proc. of the 17th Conference of the European Chapter of the Association for Computational Linguistics*, 2023, pp. 3845–3864. https://doi.org/10.18653/v1/2023.eacl-main.277
13. Schulhoff S., Ilie M., Balepur N., Kahadze K., Liu A., Si C., et al. The prompt report: a systematic survey of prompt engineering techniques. *arXiv*, 2024, arXiv:2406.06608. https://doi.org/10.48550/arXiv.2406.06608
14. Liu P., Yuan W., Fu J., Jiang Z., Hayashi H., Neubig G. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 2023, vol. 55, no. 9, pp. 1–35. https://doi.org/10.1145/3560815
15. Li Y.B., Wu K. Spell: semantic prompt evolution based on a LLM. *arXiv*, 2023, arXiv:2310.01260. https://doi.org/10.48550/arXiv.2310.01260
16. Pan R., Xing S., Diao S., Sun W., Liu X., Shum K., et al. Plum: prompt learning using metaheuristic. *arXiv*, 2023, arXiv:2311.08364. https://doi.org/10.48550/arXiv.2311.08364
17. Fernando C., Banarse D., Michalewski H., Osindero S., Rocktäschel T. Promptbreeder: self-referential self-improvement via prompt evolution. *arXiv*, 2023, arXiv:2309.16797. https://doi.org/10.48550/arXiv.2309.16797
18. Eiben A.E., Smith J.E. *Introduction to Evolutionary Computing*. Springer, 2015, 287 p.
19. Ye H., Wang J., Cao Z., Berto F., Hua C., Kim H., et al. ReEvo: large language models as hyper-heuristics with reflective evolution. *arXiv*, 2024, arXiv:2402.01145. https://doi.org/10.48550/arXiv.2402.01145
20. Holland J.H. Genetic algorithms. *Scientific American*, 1992, vol. 267, no. 1, pp. 66–72. https://doi.org/10.1038/scientificamerican0792-66
21. Lipowski A., Lipowska D. Roulette-wheel selection via stochastic acceptance. *Physica A: Statistical Mechanics and its Applications*, 2012, vol. 391, no. 6, pp. 2193–2196. https://doi.org/10.1016/j.physa.2011.12.004
22. Storn R., Price K. Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *Journal of*

3. Liu P., Yuan W., Fu J., Jiang Z., Hayashi H., Neubig G. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing // ACM Computing Surveys. 2023. V. 55. N 9. P. 1–35. https://doi.org/10.1145/3560815
4. Brown T.B., Mann B., Ryder N., Subbiah M., Kaplan J., Dhariwal P., et al. Language models are few-shot learners // arXiv. 2020. arXiv:2005.14165. https://doi.org/10.48550/arXiv.2005.14165
5. Wang N., Peng Z., Que H., Liu J., Zhou W., Wu Y., et al. RoleLLM: benchmarking, eliciting, and enhancing role-playing abilities of large language models // Proc. of the Annual Meeting of the Association for Computational Linguistics. 2024. P. 14743–14777. https://doi.org/10.18653/v1/2024.findings-acl.878
6. Wei J., Wang X., Schuurmans D., Bosma M., Ichter B., Xia F., et al. Chain-of-thought prompting elicits reasoning in large language models // arXiv. 2022. arXiv:2201.11903. https://doi.org/10.48550/arXiv.2201.11903
7. Wang L., Xu W., Lan Y., Hu Z., Lan Y., Lee R.K.-W., Lim E.-P. Plan-and-solve prompting: improving zero-shot chain-of-thought reasoning by large language models // Proc. of the 61st Annual Meeting of the Association for Computational Linguistics. 2023. V. 1. P. 2609–2634. https://doi.org/10.18653/v1/2023.acl-long.147
8. Leidinger A., van Rooij R., Shutova E. The language of prompting: What linguistic properties make a prompt successful? // Proc. of the Findings of the Association for Computational Linguistics: EMNLP. 2023. P. 9210–9232. https://doi.org/10.18653/v1/2023.findings-emnlp.618
9. Shin T., Razeghi Y., Logan R.L., Wallace E., Singh S. AutoPrompt: Eliciting knowledge from language models with automatically generated prompts // Proc. of the Conference on Empirical Methods in Natural Language. 2020. P. 4222–4235. https://doi.org/10.18653/v1/2020.emnlp-main.346
10. Kwon M., Kim G., Kim J., Lee H., Kim J. StablePrompt: automatic prompt tuning using reinforcement learning for large language models // arXiv. 2024. arXiv:2410.07652. https://doi.org/10.48550/arXiv.2410.07652
11. Guo Q., Wang R., Guo J., Li B., Song K., Tan X., et al. EvoPrompt: Connecting large language models with evolutionary algorithms yields powerful prompt optimizers // arXiv. 2023. arXiv:2309.08532. https://doi.org/10.48550/arXiv.2309.08532
12. Prasad A., Hase P., Zhou X., Bansal M. GrIPS: gradient-free, edit-based instruction search for prompting large language models // Proc. of the 17th Conference of the European Chapter of the Association for Computational Linguistics. 2023. P. 3845–3864. https://doi.org/10.18653/v1/2023.eacl-main.277
13. Schulhoff S., Ilie M., Balepur N., Kahadze K., Liu A., Si C., et al. The prompt report: a systematic survey of prompt engineering techniques // arXiv. 2024. arXiv:2406.06608. https://doi.org/10.48550/arXiv.2406.06608
14. Liu P., Yuan W., Fu J., Jiang Z., Hayashi H., Neubig G. Pre-train, prompt, and predict: a systematic survey of prompting methods in natural language processing // ACM Computing Surveys. 2023. V. 55. N 9. P. 1–35. https://doi.org/10.1145/3560815
15. Li Y.B., Wu K. Spell: semantic prompt evolution based on a LLM // arXiv. 2023. arXiv:2310.01260. https://doi.org/10.48550/arXiv.2310.01260
16. Pan R., Xing S., Diao S., Sun W., Liu X., Shum K., et al. Plum: prompt learning using metaheuristic // arXiv. 2023. arXiv:2311.08364. https://doi.org/10.48550/arXiv.2311.08364
17. Fernando C., Banarse D., Michalewski H., Osindero S., Rocktäschel T. Promptbreeder: self-referential self-improvement via prompt evolution // arXiv. 2023. arXiv:2309.16797. https://doi.org/10.48550/arXiv.2309.16797
18. Eiben A.E., Smith J.E. Introduction to Evolutionary Computing. Springer, 2015. 287 p.
19. Ye H., Wang J., Cao Z., Berto F., Hua C., Kim H., et al. ReEvo: large language models as hyper-heuristics with reflective evolution // arXiv. 2024. arXiv:2402.01145. https://doi.org/10.48550/arXiv.2402.01145
20. Holland J.H. Genetic algorithms // Scientific American. 1992. V. 267. N 1. P. 66–72. https://doi.org/10.1038/scientificamerican0792-66
21. Lipowski A., Lipowska D. Roulette-wheel selection via stochastic acceptance // Physica A: Statistical Mechanics and its Applications. 2012. V. 391. N 6. P. 2193–2196. https://doi.org/10.1016/j.physa.2011.12.004
22. Storn R., Price K. Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces // Journal of

Global Optimization, 1997, vol. 11, no. 4, pp. 341–359. https://doi.org/10.1023/a:1008202821328
23. Russell S., Norvig P. *Artificial Intelligence: a Modern Approach*. Pearson, 2009, 1152 p.
24. Kirkpatrick S., Gelatt C.D., Vecchi M.P. Optimization by simulated annealing. *Science*, 1983, vol. 220, no. 4598, pp. 671–680. https://doi.org/10.1126/science.220.4598.671
25. Glover F. Future paths for integer programming and links to artificial intelligence. *Computers and Operations Research*, 1986, vol. 13, no. 5, pp. 533–549. https://doi.org/10.1016/0305-0548(86)90048-1
26. Geem Z.W., Kim J.H., Loganathan G.V. A new heuristic optimization algorithm: harmony search. *Simulation*, 2001, vol. 76, no. 2, pp. 60–68. https://doi.org/10.1177/003754970107600201
27. Larranaga P. A review on estimation of distribution algorithms. *Genetic Algorithms and Evolutionary Computation*, 2002, vol. 2, pp. 57–100. https://doi.org/10.1007/978-1-4615-1539-5_3
28. Ross B.J. A Lamarckian evolution strategy for genetic algorithms. *Practical Handbook of Genetic Algorithms*, 2019, pp. 1–16. https://doi.org/10.1201/9780429128356-1
29. Voudouris C., Tsang E.P., Alsheddy A. Guided local search. *International Series in Operations Research & Management Science*, 2010, vol. 146, pp. 321–361. https://doi.org/10.1007/978-1-4419-1665-5_11
30. Dorigo M., Maniezzo V., Colorni A. Ant system: Optimization by a colony of cooperating agents. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 1996, vol. 26, no. 1, pp. 29–41. https://doi.org/10.1109/3477.484436
31. Shibasaka K., Kanazawa K., Yasunaga M. Decoupling-capacitor allocation problem solved by genetic algorithm. *Proc. of the IEEE Electrical Design of Advanced Packaging Systems Symposium (EDAPS)*, 2013, pp. 225–228. https://doi.org/10.1109/edaps.2013.6724430
32. Kim H., Kim M., Berto F., Kim J., Park J. DevFormer: a symmetric transformer for context-aware device placement. *arXiv*, 2022, arXiv:2205.13225. https://doi.org/10.48550/arXiv.2205.13225
33. Gohil A., Tayal M., Sahu T., Sawalpurkar V. Travelling salesman problem: parallel implementations & analysis. *arXiv*, 2022, arXiv:2205.14352. https://doi.org/10.48550/arXiv.2205.14352
34. Liu M.X., Liu F., Fiannaca A.J., Koo T., Dixon L., Terry M., Cai C.J. "We Need Structured Output": towards user-centered constraints on large language model output. *Proc. of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–9. https://doi.org/10.1145/3613905.3650756
35. Kamath A., Ferret J., Pathak S., Vieillard N., Merhej R., Perrin S., et al. Gemma 3 technical report. *arXiv*, 2025, arXiv:2503.19786. https://doi.org/10.48550/arXiv.2503.19786
36. Agrawal L.A., Tan S., Soylu D., Ziems N., Khare R., Opsahl-Ong K., et al. GEPA: reflective prompt evolution can outperform reinforcement learning. *arXiv*, 2025, arXiv:2507.19457. https://doi.org/10.48550/arXiv.2507.19457

Global Optimization. 1997. V. 11. N 4. P. 341–359. https://doi.org/10.1023/a:1008202821328
23. Russell S., Norvig P. Artificial Intelligence: a Modern Approach. Pearson, 2009. 1152 p.
24. Kirkpatrick S., Gelatt C.D., Vecchi M.P. Optimization by simulated annealing // Science. 1983. V. 220. N 4598. P. 671–680. https://doi.org/10.1126/science.220.4598.671
25. Glover F. Future paths for integer programming and links to artificial intelligence // Computers and Operations Research. 1986. V. 13. N 5. P. 533–549. https://doi.org/10.1016/0305-0548(86)90048-1
26. Geem Z.W., Kim J.H., Loganathan G.V. A new heuristic optimization algorithm: harmony search // Simulation. 2001. V. 76. N 2. P. 60–68. https://doi.org/10.1177/003754970107600201
27. Larranaga P. A review on estimation of distribution algorithms // Genetic Algorithms and Evolutionary Computation. 2002. V. 2. P. 57–100. https://doi.org/10.1007/978-1-4615-1539-5_3
28. Ross B.J. A Lamarckian evolution strategy for genetic algorithms // Practical Handbook of Genetic Algorithms. 2019. P. 1–16. https://doi.org/10.1201/9780429128356-1
29. Voudouris C., Tsang E.P., Alsheddy A. Guided local search // International Series in Operations Research & Management Science. 2010. V. 146. P. 321–361. https://doi.org/10.1007/978-1-4419-1665-5_11
30. Dorigo M., Maniezzo V., Colorni A. Ant system: Optimization by a colony of cooperating agents // IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 1996. V. 26. N 1. P. 29–41. https://doi.org/10.1109/3477.484436
31. Shibasaka K., Kanazawa K., Yasunaga M. Decoupling-capacitor allocation problem solved by genetic algorithm // Proc. of the IEEE Electrical Design of Advanced Packaging Systems Symposium (EDAPS). 2013. P. 225–228. https://doi.org/10.1109/edaps.2013.6724430
32. Kim H., Kim M., Berto F., Kim J., Park J. DevFormer: a symmetric transformer for context-aware device placement // arXiv. 2022. arXiv:2205.13225. https://doi.org/10.48550/arXiv.2205.13225
33. Gohil A., Tayal M., Sahu T., Sawalpurkar V. Travelling salesman problem: parallel implementations & analysis // arXiv. 2022. arXiv:2205.14352. https://doi.org/10.48550/arXiv.2205.14352
34. Liu M.X., Liu F., Fiannaca A.J., Koo T., Dixon L., Terry M., Cai C.J. "We Need Structured Output": towards user-centered constraints on large language model output // Proc. of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems. 2024. P. 1–9. https://doi.org/10.1145/3613905.3650756
35. Kamath A., Ferret J., Pathak S., Vieillard N., Merhej R., Perrin S., et al. Gemma 3 technical report // arXiv. 2025. arXiv:2503.19786. https://doi.org/10.48550/arXiv.2503.19786
36. Agrawal L.A., Tan S., Soylu D., Ziems N., Khare R., Opsahl-Ong K., et al. GEPA: reflective prompt evolution can outperform reinforcement learning // arXiv. 2025. arXiv:2507.19457. https://doi.org/10.48550/arXiv.2507.19457

Authors

Viktor N. Zhuravlev — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0009-0788-9790, viktor.zhuravlev2003@mail.ru
Artur R. Khairullin — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0007-0442-7764, arkhairullin@itmo.ru
Ernest A. Dyagin — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0000-3865-4446, tsenred@yandex.com
Alena N. Sitkina – Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0002-6046-8943, sitkina.alena2017@yandex.ru
Nikita I. Kulin — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57222386134, https://orcid.org/0000-0002-3952-6080, kylin98@list.ru

Авторы

Журавлев Виктор Николаевич — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0009-0788-9790, viktor.zhuravlev2003@mail.ru
Хайруллин Артур Рустамович — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0007-0442-7764, arkhairullin@itmo.ru
Дягин Эрнест Александрович — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0000-3865-4446, tsenred@yandex.com
Ситкина Алена Николаевна — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0002-6046-8943, sitkina.alena2017@yandex.ru
Кулин Никита Игоревич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57222386134, https://orcid.org/0000-0002-3952-6080, kylin98@list.ru

Received 11.06.2025
Approved after reviewing 01.09.2025
Accepted 21.11.2025

Статья поступила в редакцию 11.06.2025
Одобрена после рецензирования 01.09.2025
Принята к печати 21.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1142-1149

Transformer-based automated coronary artery segmentation with domain-specific pretraining

Nouar Ismail[1], Alexandra S. Vatian[2], Artem D. Beresnev[3], Alexey A. Zubanenko[4], Natalia F. Gusarova[5]✉, Igor A. Men'kov[6]

[1,2,3,4,5] ITMO University, Saint Petersburg, 197101, Russian Federation
[4] Imaging Medical Vision LLC, Saint Petersburg, 191119, Russian Federation
[6] The S.M. Kirov Military Medical Academy, Saint Petersburg, 194044, Russian Federation

[1] noauresmail@gmail.com, https://orcid.org/0009-0003-3545-9719
[2] alexvatyan@gmail.com, https://orcid.org/0000-0002-5483-716X
[3] artem.beresnev@itmo.ru, https://orcid.org/0000-0002-4646-6856
[4] zubdocmri@gmail.com, https://orcid.org/0000-0001-6953-5239
[5] natfed@list.ru✉, https://orcid.org/0000-0002-1361-6037
[6] i.menkov@gmail.com, https://orcid.org/0000-0002-1569-2180

Abstract
Automated segmentation of coronary arteries in coronary computed tomography angiography plays an important role in the diagnosis and treatment of coronary artery disease. Manual segmentation of coronary arteries requires significant labor costs and is accompanied by subjective errors, which necessitates the development of accurate and reliable automated methods for coronary artery segmentation. The paper presents an approach based on a deep neural network with the Swin-UNETR architecture which combines the advantages of visual transformers and the U-Net structure. To improve the accuracy, a domain-specific transfer learning strategy was used: the model was pre-trained on the ImageCAS dataset, and then further trained on a specialized dataset created for Automated Segmentation of Coronary Arteries (ASOCA) Challenge with expert labeling of coronary arteries. The accuracy of the model was assessed on 10 test Computed Tomography Coronary Angiography cases from the ASOCA dataset. The average Dice coefficient was 0.8778, and the average 95th percentile Hausdorff distance (HD_{95}) was 11.66 mm. The obtained results demonstrate that the accuracy of the proposed method is at the level of the leading models presented in the official ASOCA Challenge rating and exceeds the average inter-rater labeling. The proposed method provides high accuracy of coronary artery segmentation. In the future, the introduction of post-processing methods such as connected component filtering or vessel tracking, and spatial attention mechanisms can improve the accuracy of arterial contour localization and the adaptability of the model to various types of computed tomography data.

Keywords
coronary artery, computed tomography coronary angiography (CTCA), deep learning, Swin-UNETR, vision transformers, medical image segmentation, transfer learning, ImageCAS, ASOCA

Acknowledgements
This work was supported by the Russian Science Foundation, Grant No. 23-11-00346.

For citation: Ismail N., Vatian A.S., Beresnev A.D., Zubanenko A.A., Gusarova N.F., Men'kov I.A. Transformer-based automated coronary artery segmentation with domain-specific pretraining. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1142–1149. doi: 10.17586/2226-1494-2025-25-6-1142-1149

© Ismail N., Vatian A.S., Beresnev A.D., Zubanenko A.A., Gusarova N.F., Men'kov I.A., 2025

УДК 004.932

Автоматизированная сегментация коронарных артерий на основе трансформера с доменно-специфической предварительной подготовкой

Нуар Исмаил[1], Александра Сергеевна Ватьян[2], Артем Дмитриевич Береснев[3], Алексей Александрович Зубаненко[4], Наталия Федоровна Гусарова[5]✉, Игорь Анатольевич Меньков[6]

[1,2,3,4,5] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
[4] ООО ИМВИЖН, Санкт-Петербург, 191119, Российская Федерация
[6] Военно-медицинская академия имени С.М. Кирова, Санкт-Петербург, 194044, Российская Федерация

[1] noauresmail@gmail.com, https://orcid.org/0009-0003-3545-9719
[2] alexvatyan@gmail.com, https://orcid.org/0000-0002-5483-716X
[3] artem.beresnev@itmo.ru, https://orcid.org/0000-0002-4646-6856
[4] zubdocmri@gmail.com, https://orcid.org/0000-0001-6953-5239
[5] natfed@list.ru✉, https://orcid.org/0000-0002-1361-6037
[6] i.menkov@gmail.com, https://orcid.org/0000-0002-1569-2180

Аннотация

Введение. Автоматизированная сегментация коронарных артерий при компьютерной томографической коронарографии играет важную роль в диагностике и лечении ишемической болезни сердца. Ручная сегментация коронарных артерий требует значительных трудозатрат и сопровождается субъективными погрешностями, что обусловливает необходимость разработки точных и надежных автоматизированных методов сегментации. **Метод.** В работе представлен подход на основе глубокой нейронной сети с архитектурой Swin-UNETR, сочетающей преимущества визуальных трансформеров и структуры U-Net. Для повышения точности применена доменно-специфическая стратегия трансферного обучения: модель предварительно обучена на наборе данных ImageCAS, после чего дообучена на специализированном — Automated Segmentation of Normal and Diseased Coronary Arteries (ASOCA) с экспертной разметкой коронарных артерий. **Основные результаты.** Оценка точности модели выполнена на 10 тестовых случаях из набора ASOCA. Средний коэффициент Дайса составил 0,8778, а среднее значение 95 % — расстояния Хаусдорфа (HD_{95}) — 11,66 мм. Полученные результаты демонстрируют, что точность предложенного метода находится на уровне ведущих моделей, представленных в официальном рейтинге ASOCA Challenge, и превосходит средний показатель межэкспертной разметки. **Обсуждение.** Предложенный метод обеспечивает высокую точность сегментации коронарных артерий. Вместе с тем отмечена вариабельность результатов по HD_{95}, обусловленная отсутствием специализированной постобработки, чувствительностью к мелким сосудам и различиями в характеристиках тестовых данных. В перспективе внедрение методов постобработки, таких как фильтрация связанных компонентов или отслеживание сосудов, а также механизмов пространственного внимания, может повысить точность локализации артериального контура и адаптируемость модели к различным типам данных компьютерной томографии.

Ключевые слова

коронарные артерии, компьютерная томографическая коронарография с контрастным усилением, глубокое обучение, Swin-UNETR, визуальный трансформер, сегментация медицинских изображений, трансферное обучение, ImageCAS, ASOCA

Благодарности

Исследование поддержано грантом Российского научного фонда, № 23-11-00346.

Ссылка для цитирования: Исмаил Н., Ватьян А.С., Береснев А.Д., Зубаненко А.А., Гусарова Н.Ф., Меньков И.А. Автоматизированная сегментация коронарных артерий на основе трансформера с доменно-специфической предварительной подготовкой // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1142–1149 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-6-1142-1149

Introduction and motivation

Coronary Artery Disease (CAD) is one of the leading causes of death worldwide [1]. It occurs when atherosclerotic plaque builds up in the coronary arteries, causing narrowing or blockage that reduces blood flow to the heart muscle. This condition, known as stenosis, can lead to chest pain (angina), ischemia, or even myocardial infarction if left untreated [2]. Accurate and early detection of coronary stenosis is essential for effective diagnosis and treatment planning.

Coronary Computed Tomography Angiography (CCTA) is increasingly used as a non-invasive method for initial diagnosis. CCTA provides 3D images of coronary arteries and can often help avoid unnecessary invasive procedures. The analysis of the resulting images is increasingly performed in specialized software packages which can be divided into three groups:

1. Software packages certified for use in clinical practice — Syngo.via (Siemens Healthineers)[1], Vitrea Advanced Visualization (Canon Medical Systems) [3], to name a few;
2. AI-based and cloud-based solutions used as auxiliary tools — HeartFlow FFRct [4], Cleerly [5];
3. Research-oriented or semi-clinical software — 3D Slicer [6] containing vascular modeling toolkit, MITK [7, 8], to name a few.

Resources of groups 1 and 2 are proprietary, i.e. paid, which prevents their use when official localization

[1] Syngo.CT Coronary Analysis (syngo.CT Applications) | Instructions for Use. C2-via-CA.621.07.02.02.

is impossible. In addition, the development of clinical practice creates new needs of clinicians in CCTA analysis, the response to which is an incessant flow of research in this area.

Automated segmentation methods initially relied on traditional approaches, such as Active Shape Models (ASMs), Active Contour Models (ACMs), and atlas-based techniques [9]. The introduction of deep learning, particularly Convolutional Neural Networks (CNNs), revolutionized coronary artery segmentation by enabling automated feature extraction directly from imaging data, significantly improving accuracy and robustness [10]. Models like U-Net, 3D U-Net, and nnU-Net have achieved strong results [11, 12]. Recently, transformer-based models have gained popularity for image segmentation. Several adaptations, including the 3D U-Net, were developed to handle volumetric data, enhancing performance in complex segmentation tasks [13]. Chen et al. [13] proposed a multi-channel 3D U-Net incorporating vesselness maps, improving accuracy by emphasizing tubular structures in Computed Tomography Angiography (CTA) images.

Despite significant advancements, coronary artery segmentation remains challenging due to issues like low contrast, small vessel diameter, and severe class imbalance in imaging data. To address these challenges, specialized loss functions such as focal loss have been employed, as demonstrated by Chen et al. [14] with their 3D Dense-U-Net model which achieved improved segmentation by effectively mitigating class imbalance. Additionally, attention mechanisms and residual connections have been explored to further enhance segmentation accuracy. Wang et al. [15] introduced the DR-LCT-UNet, integrating dense residual modules with local contextual transformers, effectively capturing rich context and achieving superior results compared to traditional 3D U-Net models. Zhang et al. [16] successfully utilized Swin-UNET for medical images segmentation, highlighting its efficiency and accuracy compared to conventional CNN-based models.

Training transformer-based models from scratch requires large datasets. However, models pretrained on large-scale medical imaging datasets capture robust low- and mid-level features — such as edges, textures, and anatomical patterns — which can be effectively adapted to niche tasks like coronary artery segmentation where labeled data are scarce [17, 18]. These pretrained weights not only accelerate convergence by reducing the number of required training epochs but also enhance generalization on the target dataset, mitigating overfitting [19]. In transformer-based architectures like Swin UNETR, pretrained on volumetric CT datasets (e.g., Beyond the Cranial Vault (BTCV), Medical Segmentation Decathlon [20]), transfer learning has shown superior performance in downstream segmentation tasks, outperforming fully convolutional models [21].

The best methods of computed tomography angiography of coronary vessels are accumulated and compared within the Automated Segmentation of Normal and Diseased Coronary Arteries (ASOCA) challenge [22].

For an objective comparison of the proposed technical solutions, the ASOCA Challenge Dataset [23] was developed. It is a large-scale dataset consisting of over 1000 cases from realistic clinical scenarios at Guangdong Provincial People's Hospital from April 2012 to December 2018. This publicly available dataset is significantly larger than previous datasets and is intended as a benchmark for comparing different network architectures. Since 2020, it has become the de facto standard for assessing the effectiveness of emerging developments in the field of CCTA analysis automation. We emphasize that such assessments do not replace clinical trials but are a necessary step in assessing the potential of the proposed technical solution on the way to its clinical certification.

In our study, we aim to create a lightweight, open-source solution of group 3 in the field of research-oriented or semi-clinical software for CCTA segmentation that meets modern requirements for the level of software implementation and accuracy. In our solution, pretrained Swin UNET weights from ImageCAS [24] and BTCV serve as a solid foundation, enabling more efficient and effective fine-tuning on the smaller ASOCA dataset. We utilize the ASOCA dataset to evaluate and benchmark the performance of the proposed segmentation method in comparison with SOTA models presented at ASOCA Challenge.

Research methodology

Scope and Technologies

In our study, we leveraged a modern deep learning pipeline composed of PyTorch, PyTorch Lightning, and MONAI to implement and manage coronary artery segmentation models. PyTorch is a widely adopted, open-source deep learning framework that provides flexibility, high performance through its C++/CUDA backend, and ease of use via Python integration. To streamline training and reduce boilerplate code, we utilized PyTorch Lightning — a lightweight wrapper that structures the code into modular components, such as LightningModule, while handling tasks like GPU management, logging, and checkpointing. For medical imaging-specific needs, we employed Medical Open Network for AI (MONAI), an open-source framework built on PyTorch. MONAI provides domain-tailored tools for medical image processing, including specialized data loaders, transforms, architectures, and evaluation metrics. The framework consists of three core components: MONAI Core, which supports model development and training; MONAI Deploy, which facilitates integration into clinical workflows; and MONAI Label, an AI-assisted annotation tool that integrates with platforms like 3D Slicer to accelerate dataset labeling. This integrated toolchain enabled efficient experimentation and deployment within a medical imaging context.

This study proposes a deep learning-based segmentation framework for extracting coronary arteries from contrast-enhanced CT Coronary Angiography (CTCA) images (Fig. 1). The methodology is structured into four main stages: dataset analysis, transfer learning, fine-tuning, and evaluation. All experiments were conducted on a high-performance workstation equipped with an NVIDIA TITAN RTX GPU featuring 24 GB of VRAM, allowing efficient processing of high-resolution 3D volumes and large batch sizes without memory bottlenecks.

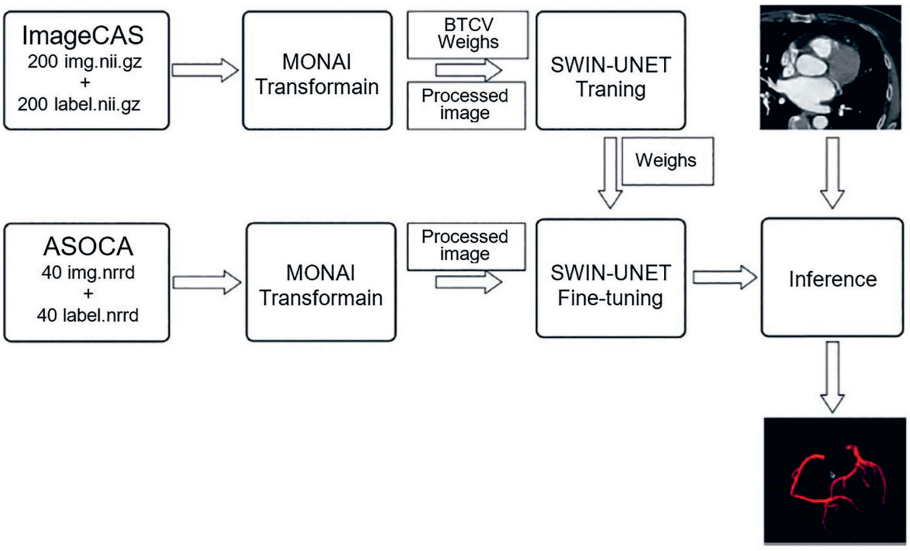

Fig. 1. Overview of the Proposed Coronary Artery Segmentation Pipeline

Data Analysis and Preprocessing

To ensure consistency across datasets and guide transform selection, we conducted an exploratory analysis using the DataAnalyzer class from MONAI Auto3DSeg framework. This tool was used independently on both the ImageCAS and ASOCA datasets to extract voxel spacing, intensity range, and average shape. The analysis gave us insight into our preprocessing pipeline, which includes resampling all volumes to a unified spacing, intensity normalization to the [0, 1] range, channel formatting, and region-of-interest cropping. These transformations were essential to harmonize differences in scanner settings and resolution between datasets.

Model Initialization and Pretraining

For the segmentation backbone, we adopted Swin-UNETR, a transformer-based encoder-decoder model tailored for volumetric medical image segmentation. The encoder consists of hierarchical Swin Transformer blocks which model both short- and long-range dependencies through shifted window attention. The model was first initialized with pretrained weights from the BTCV dataset which provides anatomical diversity through multi-organ labels in abdominal CTs. We then pretrained the model on ImageCAS, a dataset containing over 100 CTCA scans, using a curated subset of 111 cases. This pretraining phase helped the model adapt to coronary artery-specific textures and geometry before exposure to the target dataset. The training process employed the AdamW optimizer with a base learning rate of $2 \cdot 10^{-4}$, a warmup scheduler, and a combined Dice + Focal Loss function to handle class imbalance.

Fine-Tuning on ASOCA

Once the model converged on ImageCAS, it was fine-tuned using the ASOCA Challenge dataset which provides high-quality, expert-annotated coronary artery segmentations. This phase aimed to adapt the pretrained model to the exact imaging protocol and labeling schema used in the ASOCA benchmark. Training was conducted for up to 500 epochs, with early stopping based on validation Dice score. Evaluation was performed every 10 epochs using a sliding window inference approach and was monitored via TensorBoard logs and quantitative metrics.

Evaluation Metrics. We evaluated the segmentation quality using the Dice Similarity Coefficient (DSC) and the 95th percentile Hausdorff Distance (HD_{95}):

The DSC measures the overlap between predicted and ground truth segmentations. It is defined as:

$$DSC = \frac{2|P \cap G|}{|P| + |G|}.$$

Here P is the predicted segmentation; G is the ground truth segmentation; $|P \cap G|$ is the number of overlapping voxels. A Dice score of 1 indicates perfect overlap.

The HD measures the maximum distance from a point on one surface to the nearest point on the other. We use the HD_{95} version to reduce sensitivity to outliers:

$$HD_{95}(P, G) = \max\left\{\sup_{p \in P} \inf_{g \in G} \|p - g\|, \sup_{g \in G} \inf_{p \in P} \|g - p\|\right\}_{95\,\%}.$$

Here $\|p - g\|$ is the Euclidean distance between boundary points. The subscript 95 % denotes the 95th percentile, which makes the metric more robust to outliers compared to the standard Hausdorff Distance. This metric reflects boundary agreement and penalizes extreme deviations, which is critical in clinical applications like coronary segmentation.

Inference and Post-processing. After training, the best model was used to infer segmentation on the ASOCA test set. Inference predictions were post-processed using thresholding, connected component filtering, and re-orientation transforms to match the original affine metadata. Outputs were saved in NIfTI format and converted to the required structure for ASOCA submission. Additionally, qualitative comparison with expert annotations was visualized in both 2D cross-sectional slices and 3D surface renderings. The pipeline begins with two datasets: ImageCAS, containing 200 CTCA scans in NIfTI format, and ASOCA, comprising 40 scans in NRRD format. Each dataset includes corresponding segmentation labels.

Both datasets undergo preprocessing and augmentation using tailored MONAI transform pipelines. The Swin-UNET model is initially pretrained on ImageCAS, using BTCV weights for initialization. After convergence, the best model weights are fine-tuned on the ASOCA dataset to adapt to its specific annotation style and imaging characteristics.

The fine-tuned model is then used to perform inference on unseen ASOCA images. Finally, 2D slices and 3D reconstructions are generated to qualitatively evaluate segmentation performance and visualize anatomical accuracy.

Results and Comparative Analysis

An example of images segmented by our method, in comparison with manual segmentation and ground truth, is shown in Fig. 2.

Our results are presented in Table 1. For comparison, Table 2 shows the best results from the ASOCA Challenge.

To evaluate the performance of our proposed Swin-UNET-based segmentation pipeline, we tested the model on 10 representative CTCA cases from the ASOCA dataset. As shown in Table 1, the model achieved a mean DSC of 0.8778 and an average Hausdorff Distance (HD_{95}) of 11.656 mm across all cases. Notably, the Dice scores remained consistently high, with several cases, such as Case 6 (0.9340), Case 8 (0.9253), and Case 9 (0.8959), demonstrating very close agreement with expert annotations. Meanwhile, low HD_{95} values in Case 6 through Case 9 (all ≤ 1.00 mm) suggest precise boundary localization in many test cases.

Table 1. Dice Score and 95th percentile Hausdorf Distance (HD_{95}) for proposed model

Case No.	Dice Score	HD_{95}, mm
1	0.8573	34.76
2	0.8700	1.41
3	0.8637	4.12
4	0.8442	22.47
5	0.9340	21.02
6	0.8540	1.00
7	0.9253	1.00
8	0.8573	1.00
9	0.8959	1.00
10	0.8342	24.78
Mean	0.8778	11.66

When comparing these results to the official ASOCA Challenge leaderboard (Table 2), our model average Dice score of 0.8778 is competitive with top-performing methods such as “Submission 1” (0.87 ± 0.04) and the “Top 7 Ensemble” (0.87 ± 0.04). While ensemble methods show slightly lower HD values (e.g., 1.56 (SD:2.55) mm), our model Dice scores fall well within the top five single-model submissions, as shown in the leaderboard. In particular, our results exceed those of Submission 5 to Submission 8, and are on par with or slightly above Submission 2 and Submission 4 in segmentation accuracy.

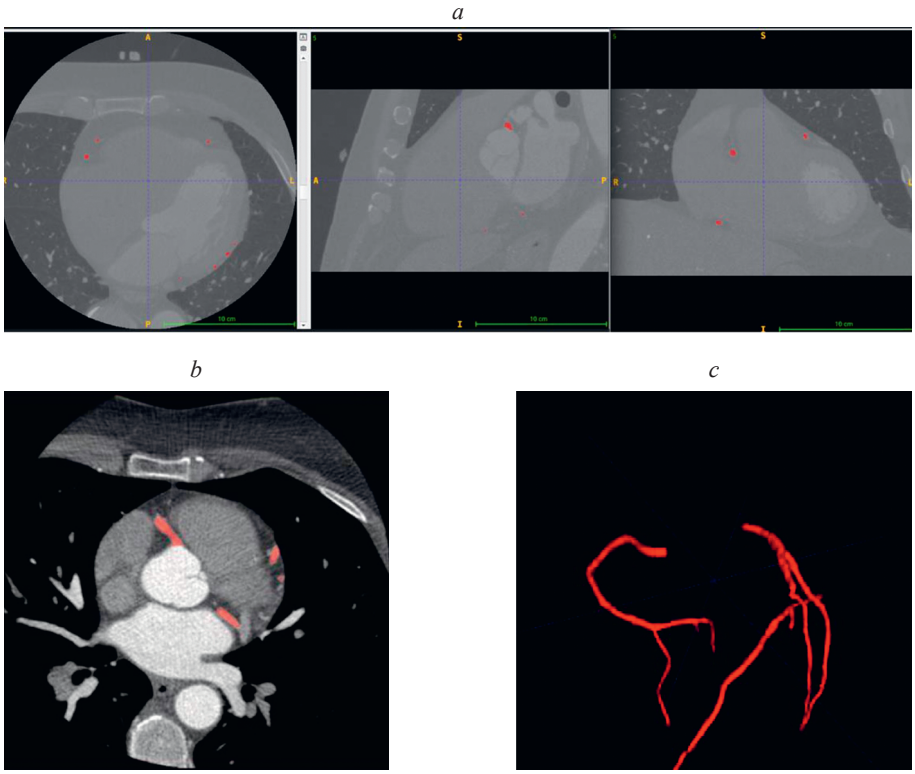

Fig. 2. Example of Coronary Artery Segmentation: manual segmentation in three planes using specialized software (*a*); ground truth method (*b*); predicted segmentation (3D reconstruction) (*c*)

Table 2. Summary of Dice Score and Hausdorf Distance for each method (official results from ASOCA Challenge)

Submission	Dice Score	HD_{95}, mm (mean ±SD)
Top 3 Ensemble	0.88 ± 0.04	1.56 (SD:2.55)
Top 5 Ensemble	0.88 ± 0.04	1.73 (SD:2.08)
Top 6 Ensemble	0.87 ± 0.04	1.61 (SD:2.17)
Submission 1	0.87 ± 0.04	4.16 (SD:7.13)
Submission 2	0.84 ± 0.05	2.34 (SD:2.92)
Submission 3	0.86 ± 0.07	6.22 (SD:15.52)
Submission 4	0.87 ± 0.05	6.57 (SD:14.27)
Submission 5	0.80 ± 0.04	3.31 (SD:3.42)
Submission 6	0.84 ± 0.06	7.73 (SD:11.83)
Submission 7	0.78 ± 0.10	4.77 (SD:3.51)
Submission 8	0.73 ± 0.05	6.55 (SD:7.40)

Note: Data from from [23, Table 1]. (SD: ...) represents the standard deviation across 20 test cases. High SD values for HD_{95} reflect variability due to anatomical features and artifacts; negative values when subtracting SD are not realistic, as $HD_{95} \geq 0$ mm.

However, our Hausdorff Distances show greater variance, with high values in certain cases (e.g., Case 1: 34.76 mm, Case 10: 24.78 mm), likely due to minor discontinuities or spurious predictions in elongated artery branches. These outliers suggest that post-processing techniques, such as connected component filtering or vessel tracking, could further enhance the robustness of boundary accuracy without sacrificing overall Dice performance.

To summarize, despite not being an ensemble approach, our model achieves performance comparable to several of the top single-system entries on the ASOCA leaderboard. This affirms that the Swin-UNET with transfer learning from BTCV and ImageCAS is a strong standalone model and may serve as a reliable reference architecture for coronary artery segmentation in future research.

Conclusion

In this study, we presented a deep learning-based pipeline for coronary artery segmentation in Computed Tomography Coronary Angiography images, leveraging the Swin-UNETR architecture and domain-specific transfer learning. By pretraining the model on the ImageCAS dataset and fine-tuning it on the Automated Segmentation of Coronary Arteries Challenge (ASOCA) Challenge dataset, we achieved a mean Dice score of 0.8778 across 10 test cases — results that are competitive with top-performing entries in the ASOCA benchmark. The effectiveness of our approach highlights the importance of using anatomically relevant pretraining data and transformer-based architectures for vascular segmentation tasks. While our model performed strongly overall, some variability in boundary localization suggests that incorporating post-processing techniques or spatial attention mechanisms could further enhance segmentation robustness. Overall, our pipeline provides a reliable and efficient solution for automated coronary artery segmentation and can serve as a reference framework for future research applications.

References

1. Roth G.A., Abate D., Abate K.H., Abay S.M., Abbafati C., Abbasi N., et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, 2018, vol. 392, no. 10159, pp. 1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7
2. Montalescot G., Sechtem U., Achenbach S., Andreotti F., Arden C., Budaj A., et al. 2013 ESC guidelines on the management of stable coronary artery disease. *European Heart Journal*, 2013, vol. 34, no. 38, pp. 2949–3003. https://doi.org/10.1093/eurheartj/eht296
3. Cury R.C., Leipsic J., Abbara S., Achenbach S., Berman D., Bittencourt M., et al. CAD-RADS™ 2.0 - 2022 Coronary Artery Disease-Reporting and Data System. *Journal of Cardiovascular Computed Tomography*, 2022, vol. 16, no. 6, pp. 536–557. https://doi.org/10.1016/j.jcct.2022.07.002
4. Yang S., Chung J., Lesina K., Doh J., Jegere S., Erglis A., et al. Long-term prognostic implications of CT angiography-derived fractional flow reserve: Results from the DISCOVER-FLOW study. *Journal of Cardiovascular Computed Tomography*, 2024, vol. 18, no. 3, pp. 251–258. https://doi.org/10.1016/j.jcct.2024.01.016
5. Hurreh M., Knickelbine T., Earls J., Aquino M., Idris A., Cheng V. Enhanced cardiac CTA segmental atherosclerosis detection with artificial intelligence quantitative computed tomography (AIQCT) analysis vs conventional reader interpretation. *Journal of Cardiovascular Computed Tomography*, 2023, vol. 17, no. 4 suppl., pp. S37-S38. https://doi.org/10.1016/j.jcct.2023.05.093
6. Fedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J-C., Pujol S., et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic Resonance Imaging*, 2012, vol. 30, no. 9, pp. 1323–1341. https://doi.org/10.1016/j.mri.2012.05.001

Литература

1. Roth G.A., Abate D., Abate K.H., Abay S.M., Abbafati C., Abbasi N., et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017 // The Lancet. 2018. V. 392. N 10159. P. 1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7
2. Montalescot G., Sechtem U., Achenbach S., Andreotti F., Arden C., Budaj A., et al. 2013 ESC guidelines on the management of stable coronary artery disease // European Heart Journal. 2013. V. 34. N 38. P. 2949–3003. https://doi.org/10.1093/eurheartj/eht296
3. Cury R.C., Leipsic J., Abbara S., Achenbach S., Berman D., Bittencourt M., et al. CAD-RADS™ 2.0 - 2022 Coronary Artery Disease-Reporting and Data System // Journal of Cardiovascular Computed Tomography. 2022. V. 16. N 6. P. 536–557. https://doi.org/10.1016/j.jcct.2022.07.002
4. Yang S., Chung J., Lesina K., Doh J., Jegere S., Erglis A., et al. Long-term prognostic implications of CT angiography-derived fractional flow reserve: Results from the DISCOVER-FLOW study // Journal of Cardiovascular Computed Tomography. 2024. V. 18. N 3. P. 251–258. https://doi.org/10.1016/j.jcct.2024.01.016
5. Hurreh M., Knickelbine T., Earls J., Aquino M., Idris A., Cheng V. Enhanced cardiac CTA segmental atherosclerosis detection with artificial intelligence quantitative computed tomography (AIQCT) analysis vs conventional reader interpretation // Journal of Cardiovascular Computed Tomography. 2023. V. 17. N 4 Suppl. P. S37-S38. https://doi.org/10.1016/j.jcct.2023.05.093
6. Fedorov A., Beichel R., Kalpathy-Cramer J., Finet J., Fillion-Robin J-C., Pujol S., et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network // Magnetic Resonance Imaging. 2012. V. 30. N 9. P. 1323–1341. https://doi.org/10.1016/j.mri.2012.05.001

7. Yushkevich P.A., Piven J., Hazlett H.C., Smith R.G., Ho S., Gee J.C., Gerig G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *NeuroImage*, 2006, vol. 31, no. 3, pp. 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015
8. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., et al. An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv*, 2020, arXiv:2010.11929. https://doi.org/10.48550/arXiv.2010.11929
9. Fossan F.E., Müller L.O., Sturdy J., Bråten A.T., Jorgensen A., Wiseth R., Hellevik L.R. Machine learning augmented reduced-order models for FFR-prediction. *Computer Methods in Applied Mechanics and Engineering*, 2021, vol. 384, pp. 113892. https://doi.org/10.1016/j.cma.2021.113892
10. Gharleghi, R., Adikari D., Ellenberger K., Ooi S.Y., Ellis C., Chen C.M., et al. Automated segmentation of normal and diseased coronary arteries — The ASOCA challenge. *Computerized Medical Imaging and Graphics*, 2022, vol. 97, pp. 102049. https://doi.org/10.1016/j.compmedimag.2022.102049
11. Zeng A., Wu C., Lin G., Xie W., Hong J., Huang M.P., et al. ImageCAS: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images. *Computerized Medical Imaging and Graphics*, 2023, vol. 109, pp. 102287. https://doi.org/10.1016/j.compmedimag.2023.102287
12. Çiçek Ö., Abdulkadir A., Lienkamp S.S., Brox T., Ronneberger O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation. *Lecture Notes in Computer Science*, 2016, vol. 9901, pp. 424–432. https://doi.org/10.1007/978-3-319-46723-8_49
13. Chen Y.-C., Lin Y.-C., Wang C.-P., Lee C.-Y., Lee W.-J., Wang T.-D., Chen C.-M. Coronary artery segmentation in cardiac CT Angiography using 3D Multi-channel U-net. *arXiv*, 2019, arXiv:1907.12246. https://doi.org/10.48550/arXiv.1907.12246
14. Liang R., Huang X., Ma J., Ma G., Wang K. 3D U-Net with attention and focal loss for coronary tree segmentation. *Research Square. Preprint CC BY 4.0 License*, 2021, pp. 1–17. https://doi.org/10.21203/rs.3.rs-717054/v1
15. Wang Q., Xu L., Wang L., Yang X., Sun Y., Yang B., Greenwald S.E. Automatic coronary artery segmentation of CCTA images using UNet with a local contextual transformer. *Frontiers in Physiology*, 2023, vol. 14, pp. 1138257. https://doi.org/10.3389/fphys.2023.1138257
16. Zhang L., Yin X., Liu X., Liu Z. Medical image segmentation by combining feature enhancement Swin Transformer and UperNet. *Scientific Reports*, 2025, vol. 15, no. 1, pp. 14565. https://doi.org/10.1038/s41598-025-97779-6
17. Salehi A.W., Khan S., Gupta G., Alabduallah B.I., Almjally A., Alsolai H., et al. A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. *Sustainability*, 2023, vol. 15, no. 7, pp. 5930. https://doi.org/10.3390/su15075930
18. Matsoukas C., Haslum J.F., Sorkhei M., Soderberg M., Smith K. What makes transfer learning work for medical images: Feature reuse & other factors. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 9215–9224. https://doi.org/10.1109/CVPR52688.2022.00901
19. Yosinski J., Clune O., Bengio Y., Lipson H. How transferable are features in deep neural networks? *Proc. of the 28th International Conference on Neural Information Processing Systems*, 2014, vol. 2, pp. 3320–3328.
20. Wang X., Zhang X., Wang G., He J., Li Z., ZhuW., et al. OpenMEDLab: An open-source platform for multi-modality foundation models in medicine. *arXiv*, 2024, arXiv:2402.18028v2. https://doi.org/10.48550/arXiv.2402.18028
21. Tang Y., Yang D., Li W., Roth H.R., Landman B., Xu D., Nath V., Hatamizadeh A. Self-supervised pre-training of Swin Transformers for 3D medical image analysis. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 20698–20708. https://doi.org/10.1109/cvpr52688.2022.02007
22. Yong K.R., Lee S., Uppaluri A., Tan M.O., Ng J.J., Venkatesh S.H., et al. How to report a coronary CT angiogram: Step-by-Step guide for the novice reader. *RadioGraphics*, 2025, vol. 45, no. 7, pp. e250015. https://doi.org/10.1148/rg.250015
23. Gharleghi R., Adikari D., Ellenberger K., Webster M., Ellis C., Sowmya A., Ooi S., Beier S. Annotated computed tomography coronary angiogram images and associated data of normal and diseased arteries. *Scientific Data*, 2023, vol. 10, no. 1, pp. 128. https://doi.org/10.1038/s41597-023-02016-2

7. Yushkevich P.A., Piven J., Hazlett H.C., Smith R.G., Ho S., Gee J.C., Gerig G. User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability // NeuroImage. 2006. V. 31. N 3. P. 1116–1128. https://doi.org/10.1016/j.neuroimage.2006.01.015
8. Dosovitskiy A., Beyer L., Kolesnikov A., Weissenborn D., Zhai X., Unterthiner T., et al. An image is worth 16×16 words: Transformers for image recognition at scale // arXiv. 2020. arXiv:2010.11929. https://doi.org/10.48550/arXiv.2010.11929
9. Fossan F.E., Müller L.O., Sturdy J., Bråten A.T., Jorgensen A., Wiseth R., Hellevik L.R. Machine learning augmented reduced-order models for FFR-prediction // Computer Methods in Applied Mechanics and Engineering. 2021. V. 384. P. 113892. https://doi.org/10.1016/j.cma.2021.113892
10. Gharleghi, R., Adikari D., Ellenberger K., Ooi S.Y., Ellis C., Chen C.M., et al. Automated segmentation of normal and diseased coronary arteries — The ASOCA challenge // Computerized Medical Imaging and Graphics. 2022. V. 97. P. 102049. https://doi.org/10.1016/j.compmedimag.2022.102049
11. Zeng A., Wu C., Lin G., Xie W., Hong J., Huang M.P., et al. ImageCAS: A large-scale dataset and benchmark for coronary artery segmentation based on computed tomography angiography images // Computerized Medical Imaging and Graphics. 2023. V. 109. P. 102287. https://doi.org/10.1016/j.compmedimag.2023.102287
12. Çiçek Ö., Abdulkadir A., Lienkamp S.S., Brox T., Ronneberger O. 3D U-Net: Learning dense volumetric segmentation from sparse annotation // Lecture Notes in Computer Science. 2016. V. 9901. P. 424–432. https://doi.org/10.1007/978-3-319-46723-8_49
13. Chen Y.-C., Lin Y.-C., Wang C.-P., Lee C.-Y., Lee W.-J., Wang T.-D., Chen C.-M. Coronary artery segmentation in cardiac CT Angiography using 3D Multi-channel U-net // arXiv. 2019. arXiv:1907.12246. https://doi.org/10.48550/arXiv.1907.12246
14. Liang R., Huang X., Ma J., Ma G., Wang K. 3D U-Net with attention and focal loss for coronary tree segmentation // Research Square. Preprint CC BY 4.0 License. 2021. P. 1–17. https://doi.org/10.21203/rs.3.rs-717054/v1
15. Wang Q., Xu L., Wang L., Yang X., Sun Y., Yang B., Greenwald S.E. Automatic coronary artery segmentation of CCTA images using UNet with a local contextual transformer // Frontiers in Physiology. 2023. V. 14. P. 1138257. https://doi.org/10.3389/fphys.2023.1138257
16. Zhang L., Yin X., Liu X., Liu Z. Medical image segmentation by combining feature enhancement Swin Transformer and UperNet // Scientific Reports. 2025. V. 15. N 1. P. 14565. https://doi.org/10.1038/s41598-025-97779-6
17. Salehi A.W., Khan S., Gupta G., Alabduallah B.I., Almjally A., Alsolai H., et al. A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope // Sustainability. 2023. V. 15. N 7. P. 5930. https://doi.org/10.3390/su15075930
18. Matsoukas C., Haslum J.F., Sorkhei M., Soderberg M., Smith K. What makes transfer learning work for medical images: Feature reuse & other factors // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 9215–9224. https://doi.org/10.1109/CVPR52688.2022.00901
19. Yosinski J., Clune O., Bengio Y., Lipson H. How transferable are features in deep neural networks? // Proc. of the 28th International Conference on Neural Information Processing Systems. 2014. V. 2. P. 3320–3328.
20. Wang X., Zhang X., Wang G., He J., Li Z., ZhuW., et al. OpenMEDLab: An open-source platform for multi-modality foundation models in medicine // arXiv. 2024. arXiv:2402.18028v2. https://doi.org/10.48550/arXiv.2402.18028
21. Tang Y., Yang D., Li W., Roth H.R., Landman B., Xu D., Nath V., Hatamizadeh A. Self-supervised pre-training of Swin Transformers for 3D medical image analysis // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022. P. 20698–20708. https://doi.org/10.1109/cvpr52688.2022.02007
22. Yong K.R., Lee S., Uppaluri A., Tan M.O., Ng J.J., Venkatesh S.H., et al. How to report a coronary CT angiogram: Step-by-Step guide for the novice reader // RadioGraphics. 2025. V. 45. N 7. P. e250015. https://doi.org/10.1148/rg.250015
23. Gharleghi R., Adikari D., Ellenberger K., Webster M., Ellis C., Sowmya A., Ooi S., Beier S. Annotated computed tomography coronary angiogram images and associated data of normal and diseased arteries // Scientific Data. 2023. V. 10. N 1. P. 128. https://doi.org/10.1038/s41597-023-02016-2

24. Kirişli H.A., Schaap M., Metz C.T., Dharampal A.S., Meijboom W.B., Papadopoulou S.L., et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography. *Medical Image Analysis*, 2013, vol. 17, no. 8, pp. 859–876. https://doi.org/10.1016/j.media.2013.05.007

24. Kirişli H.A., Schaap M., Metz C.T., Dharampal A.S., Meijboom W.B., Papadopoulou S.L., et al. Standardized evaluation framework for evaluating coronary artery stenosis detection, stenosis quantification and lumen segmentation algorithms in computed tomography angiography // Medical Image Analysis. 2013. V. 17. N 8. P. 859–876. https://doi.org/10.1016/j.media.2013.05.007

Authors

Nouar Ismail — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, https://orcid.org/0009-0003-3545-9719, noauresmail@gmail.com

Alexandra S. Vatian — PhD, Dean, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57191870868, https://orcid.org/0000-0002-5483-716X, alexvatyan@gmail.com

Artem D. Beresnev — PhD, Vice Dean, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57202210221, https://orcid.org/0000-0002-4646-6856, artem.beresnev@itmo.ru

Alexey A. Zubanenko — CEO, Imaging Medical Vision LLC, Saint Petersburg, 191119, Russian Federation; PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57215436184, https://orcid.org/0000-0001-6953-5239, zubdocmri@gmail.com

Natalia F. Gusarova — PhD, Senior Researcher, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57162764200, https://orcid.org/0000-0002-1361-6037, natfed@list.ru

Igor A. Men'kov — PhD (Medicine), Head of Department, The S.M. Kirov Military Medical Academy, Saint Petersburg, 194044, Russian Federation, https://orcid.org/0000-0002-1569-2180, i.menkov@gmail.com

Авторы

Исмаил Нуар — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, https://orcid.org/0009-0003-3545-9719, noauresmail@gmail.com

Ватьян Александра Сергеевна — кандидат технических наук, декан, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57191870868, https://orcid.org/0000-0002-5483-716X, alexvatyan@gmail.com

Береснев Артем Дмитриевич — кандидат технических наук, заместитель декана, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57202210221, https://orcid.org/0000-0002-4646-6856, artem.beresnev@itmo.ru

Зубаненко Алексей Александрович — генеральный директор, ООО ИМВИЖН, Санкт-Петербург, 191119, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57215436184, https://orcid.org/0000-0001-6953-5239, zubdocmri@gmail.com

Гусарова Наталия Федоровна — кандидат технических наук, старший научный сотрудник, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57162764200, https://orcid.org/0000-0002-1361-6037, natfed@list.ru

Меньков Игорь Анатольевич — кандидат медицинских наук, заведующий отделением, Военно-медицинская академия имени С.М. Кирова, Санкт-Петербург, 194044, Российская Федерация, https://orcid.org/0000-0002-1569-2180, i.menkov@gmail.com

Received 04.06.2025
Approved after reviewing 10.10.2025
Accepted 24.11.2025

Статья поступила в редакцию 04.06.2025
Одобрена после рецензирования 10.10.2025
Принята к печати 24.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1150-1159

A weighted ensemble model combining ARIMA, LSTM, and GBM for robust time series prediction

Arumugam Vignesh[1]✉, Natarajan Vijayalakshmi[2]

[1,2] SRM Institute of Science and Technology — Ramapuram, Chennai, 600087, India

[1] vignesha2@srmist.edu.in✉, https://orcid.org/0000-0002-4211-0131
[2] vijayaln@srmist.edu.in, https://orcid.org/0000-0002-9100-9937

Abstract
Time series forecasting has been used in research and applications in a number of domains such as environmental forecasting, healthcare, finance, supply chain management, and energy consumption. Accurate prediction of future values is necessary for strategic planning operational efficiency and well-informed decision-making regarding time-dependent variables. A hybrid time series forecasting architecture is proposed that combines the strengths of machine learning and statistical models, in particular Gradient Boosting Machines (GBM), Auto-Regressive Integrated Moving Average (ARIMA), and Long Short-Term Memory (LSTM) networks. While LSTM networks and GBM are able to capture complex dependencies and nonlinear patterns, the ARIMA model captures the linear components within the time series. The hybrid model exploits ARIMA interpretability, LSTM temporal memory ability, and GBM ensemble learning efficiency by integrating these three models. Comprehensive experiments conducted on benchmark data sets have shown that the accuracy and reliability of predictions of the proposed hybridization significantly exceeds both individual models and traditional baseline models. The results show that for a variety of real-world applications, hybrid architectures can deliver reliable and accurate time series predictions.

Keywords
Auto-Regressive Integrated Moving Average (ARIMA), Long Short-Term Memory (LSTM), Gradient Boosting Machines (GBM), time series forecasting

For citation: Vignesh A., Vijayalakshmi N. A weighted ensemble model combining ARIMA, LSTM, and GBM for robust time series prediction. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1150–1159. doi: 10.17586/2226-1494-2025-25-6-1150-1159

УДК 004.891

Взвешенная ансамблевая модель, сочетающая ARIMA, LSTM и GBM для надежного прогнозирования временных рядов

Арумугам Вигнеш[1]✉, Натараян Виджаялакшми[2]

[1,2] Институт науки и технологий SRM — Рамапурам, Ченнаи, 600087, Индия

[1] vignesha2@srmist.edu.in✉, https://orcid.org/0000-0002-4211-0131
[2] vijayaln@srmist.edu.in, https://orcid.org/0000-0002-9100-9937

Аннотация
Прогнозирование временных рядов используется в исследованиях и приложениях в ряде областей, таких как прогнозирование состояния окружающей среды, здравоохранение, финансы, управление цепочками поставок и энергопотребление. Точное прогнозирование будущих значений необходимо для стратегического планирования, эффективности работы и принятия обоснованных решений относительно переменных, зависящих от времени. Предлагается гибридная архитектура прогнозирования временных рядов, сочетающая в себе сильные стороны машинного обучения и статистических моделей, в частности сетей градиентного бустинга (Gradient Boosting Machines, GBM), авторегрессионных интегрированных скользящих средних (Auto-Regressive Integrated Moving Average, ARIMA) и сетей с долговременной краткосрочной памятью (Long Short-Term Memory, LSTM). В то время как сети LSTM и GBM способны улавливать сложные и нелинейные закономерности, модель ARIMA

© Vignesh A., Vijayalakshmi N., 2025

фиксирует линейные компоненты временных рядов. Гибридная модель использует интерпретируемость ARIMA, временную память LSTM и эффективность ансамблевого обучения GBM, интегрируя эти три модели. Комплексные эксперименты, проведенные на эталонных наборах данных, показали, что точность и надежность прогнозов предлагаемой гибридизации значительно превосходят как отдельные модели, так и традиционные базовые. Результаты показывают, что для различных реальных приложений гибридные архитектуры могут обеспечивать надежные и точные прогнозы временных рядов.

Ключевые слова
авторегрессионные интегрированные скользящие сети (ARIMA), долговременная краткосрочная память (LSTM), градиентный бустинг (GBM), прогнозирование временных рядов

Ссылка для цитирования: Вигнеш А., Виджаялакшми Н. Взвешенная ансамблевая модель, сочетающая ARIMA, LSTM и GBM для надежного прогнозирования временных рядов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1150–1159. (на англ. яз.). doi: 10.17586/2226-1494-2025-25-6-1150-1159

Introduction

Time series forecasting has long been used in research and applications in a number of domains, such as environmental forecasting, healthcare, finance, supply chain management, and energy consumption. Accurate prediction of future values is necessary for strategic planning operational efficiency and well-informed decision-making regarding time-dependent variables. Several statistical and computational techniques have been developed over the decades to address this challenge but for a significant portion of the 20th century traditional models like the Auto-Regressive Integrated Moving Average (ARIMA) dominated the forecasting landscape. ARIMA and its variants are widely used due to their effectiveness in capturing linear temporal dependencies interpretability and mathematical elegance [1–5]. However real-world time series data often exhibit complex nonlinear behaviors that are impossible for linear frameworks to effectively model. The advancement of machine learning and more recently deep learning has opened up new possibilities for modeling such intricate patterns especially in noisy and large-scale datasets. In sequential data models networks like Long Short-Term Memory (LSTM) and Gradient Boosting Machines (GBM) have shown superior performance in detecting nonlinearities and long-range dependencies. These models have their own limitations despite their high predictive abilities. For instance, LSTMs although adept at sequence learning may encounter instability during training and incur significant computational costs and GBMs may be unable to detect temporal order due to their inherent structure [6–11]. Furthermore, they can be challenging to understand due to their opaqueness particularly in high-stakes sectors like healthcare and finance. Hybrid approaches that seek to integrate multiple modeling paradigms into a single forecasting framework have become more and more popular due to the complementary benefits and drawbacks of statistical models and machine learning [12–17]. Hybrid models aim to combine the interpretability and linear modeling capabilities of statistical techniques such as ARIMA with the nonlinear pattern recognition power of machine learning algorithms. This partnership could improve the generalizability robustness and accuracy of forecasting. However many hybrid approaches currently in use are either domain-specific have limited scalability or cannot generalize across datasets with different properties [18–29]. To address these shortcomings the current study presents a brand-new hybrid time series forecasting framework that successfully blends LSTM GBM and ARIMA models.

Literature Review

Time series forecasting has garnered significant attention across various disciplines, particularly in energy management, finance, and communication systems. Traditional statistical models, hybrid approaches, and machine learning-based methods have all been explored to improve accuracy and robustness. In the domain of energy consumption, Lee and Tong [11] proposed an enhanced grey prediction model optimized using genetic programming which demonstrated improved performance in short-term energy consumption forecasting by dynamically tuning the model structure. Furthermore, Pérez-Lombard et al. [15] examined building energy usage trends and underlined the necessity of better forecasting methods to aid in energy management and efficiency. Zhou et al. [25] complemented this by suggesting a real-time smart home control system that allows for responsive optimization of energy use. According to Liboschik et al. [12], the *tscount* R package has made the Generalized Linear Models for count time series available in the field of statistical forecasting. Discrete-valued time series modeling has benefited greatly from this method especially in cases where data distributions deviate from normalcy. By adding an adaptive Moving Average (MA) to the Markov-switching regression model, Pomorski and Gorse [16] further investigated the use of adaptive techniques and improved its ability to capture regime changes in financial and economic data. Additionally deep learning models and machine learning have become effective tools for time series forecasting. Xu et al. [22] revealed an Autoregressive (AR) model based on Deep Belief Networks that outperforms traditional linear models in capturing nonlinear patterns in time series data. Similar to Sen et al. [17], it was successfully integrated spatial and temporal features by combining Convolutional Neural Networks (CNNs) with LSTM networks to forecast stock prices. These models demonstrate how deep learning architectures can represent intricate dependencies that are frequently missed by conventional statistical techniques, because they can combine the best features of several methods hybrid models that combine machine learning and

statistical methods have become popular. Saleti et al. [18] suggested a MA method combined with a hybrid ARIMA-LSTM model to improve forecasting accuracy. The efficacy of combining LSTMs capacity to handle nonlinearities with ARIMAs strength in modeling linear patterns was validated by their findings. Similar to this, Xu et al. [22] demonstrated the synergy between memory-based deep learning models and conventional time series methods for anomaly detection by creating a combined LSTM-ARIMA model to identify anomalies in communication networks. Thangarajan et al. [20] provided additional support for the use of hybrid modeling techniques which put forth a new model that combines ARIMA and Artificial Neural Networks. The non-stationary and nonlinear features of real-world datasets were successfully captured by their model which enhanced prediction performance. This hybrid approach is in line with Zhao et al. [13] description of the new requirements in smart energy systems where more flexible and precise forecasting solutions are required due to complex data from numerous sensors and systems. New computational approaches are also very important. Peleg et al. [14] presented a new way to improve fast-adaptive moment estimation using the Triple Exponential Moving Average (TEMA). This could be incorporated into machine learning models to provide more responsive forecasting in dynamic environments. In addition, the author also highlighted the economic aspects impacting forecasting models and energy planning while reviewing electric energy storage technologies. The allocation of energy resources and storage efficiency are directly impacted by precise time series forecasts making these insights essential. Yip et al. [23] highlighted the significance of model transparency and interpretability in critical infrastructure systems by using linear regression techniques to identify energy theft and malfunctioning meters in smart grid applications. Young et al. [24] demonstrated the value of time series techniques in the fields of public policy and healthcare by using an interrupted time-series analysis to assess the effect of tobacco plain packaging on Quitline calls. They provide a framework for real-time energy control to maximize energy use in smart home settings. Based on real-time data inputs user preferences and grid signals the authors in [25–29] created a control-oriented algorithmic system that dynamically schedules monitors and modifies the power consumption of household appliances.

Methodology

The proposed study aims to improve the accuracy and robustness of time series forecasting by developing a hybrid model that combines three distinct approaches: ARIMA, LSTM networks, and GBM. Each model contributes uniquely: ARIMA captures linear patterns, LSTM models nonlinear dependencies with long-term memory, and GBM enhances predictive performance through boosting weak learners. This section explains the overall research design, dataset preprocessing, model architecture, mathematical formulation, algorithm, and evaluation metrics.

Dataset Description

This study employed three real-world multivariate time series datasets — Electricity Consumption (UCI Machine Learning Repository)[1], Stock Prices (Yahoo Finance)[2], and Temperature (National Oceanic and Atmospheric Administration, NOAA)[3] — to evaluate the performance of the ARIMA-LSTM-GBM ensemble model. After preprocessing the datasets for stationarity and applying normalization, each dataset was split into 70 % training and 30 % testing sets. Model performance was assessed using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), ensuring a robust and generalized evaluation of forecasting accuracy.

Data Preprocessing

Prior to model training, the raw time series data undergoes multiple preprocessing steps to enhance model effectiveness:

— Handling missing values. Linear interpolation or forward-fill is applied to impute missing data;

— Stationarity check. Augmented Dickey-Fuller test is used to evaluate stationarity. If the p-value is above 0.05, the series is differenced;

— Normalization. For LSTM and GBM models, Min-Max normalization is applied:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}},$$

where X represents the original data value; $X_{\min}$ and $X_{\max}$ denote the minimum and maximum values of the feature, respectively, and X' is the normalized value;

— Train-test split. The dataset is divided into 80 % training and 20 % testing sets.

ARIMA Model (Linear Component)

The ARIMA model is denoted as *ARIMA* (p, d, q), where p indicates the order of the AR terms; d represents the degree of differencing applied to make the series stationary, and, q specifies the number of lagged forecast errors included in the MA component.

An ARIMA model is represented as:

$$y_t = c + \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t,$$

where y_t is the observed value at time t; c is a constant term; φ_i is AR coefficients; θ_j are MA coefficients; ϵ_t is denotes the error term (assumed to be white noise).

The model is fitted using Akaike Information Criterion and Bayesian Information Criterion for order selection.

Once the ARIMA model was fitted on the time series data, its residuals — representing the nonlinear patterns

[1] UCI Machine Learning Repository. (2012). Individual household electric power consumption data set. University of California, Irvine. Available at: https://archive.ics.uci.edu/ml/datasets/individual%2Bhousehold%2Belectric%2Bpower%2Bconsumption (accessed: 06.09.2025).

[2] Yahoo Finance. (n.d.). Historical stock price data. Available at: https://finance.yahoo.com/quote/%5EGSPC/history/ (accessed: 06.09.2025).

[3] National Centers for Environmental Information. (2023). NOAA Global Surface Temperature dataset (NOAAGlobalTemp). National Oceanic and Atmospheric Administration. Available at: https://www.ncei.noaa.gov/products/land-based-station/noaa-global-temp (accessed: 06.09.2025).

not captured by ARIMA — were extracted and used as input for the LSTM and GBM models. To ensure consistent feature ranges and prevent scale-related learning issues, the residuals were scaled using Min-Max normalization to the range [0, 1]. This normalization was applied only on the training set, and the same scaling parameters were used to transform the test data. This step preserved the temporal structure while enhancing convergence speed and learning stability for the LSTM and GBM models.

LSTM Network (Nonlinear Component)

LSTM networks are a type of Recurrent Neural Network capable of learning long-term dependencies. They are particularly effective for modeling temporal sequences with nonlinear trends.

LSTM Unit Structure

Each LSTM unit comprises:

1. Forget gate state $\mathbf{f}_t$

$$\mathbf{f}_t = \sigma(\mathbf{W}_f[h_{t-1}, \mathbf{x}_t] + \mathbf{b}_f),$$

where $\mathbf{x}_t$ is input vector at time t; $[h_{t-1}, \mathbf{x}_t]$ is previous hidden state; $\sigma(\cdot)$ refers to the sigmoid activation function, h_{t-1} is the hidden state at time $t-1$; $\mathbf{W}_f$ is weight matrix; $\mathbf{b}_f$ is bias vector; $\mathbf{f}_t$ is forget gate vector;

2. Input gate state $\mathbf{i}_t$ and $\ddot{\mathbf{C}}_t$

$$\mathbf{i}_t = \sigma(\mathbf{W}_i[h_{t-1}, \mathbf{x}_t] + \mathbf{b}_i),$$
$$\ddot{\mathbf{C}}_t = \tanh(\mathbf{W}_c[h_{t-1}, \mathbf{x}_1] + \mathbf{b}_c),$$

where $\mathbf{i}_t$ is input gate vector; $\ddot{\mathbf{C}}_t$ is candidate cell state vector; $\mathbf{W}_i$, $\mathbf{W}_c$ are weight matrices; $\mathbf{b}_c$ is bias vectors;

3. Cell state update

$$\mathbf{C}_t = \mathbf{f}_t \odot \mathbf{C}_{t-1} + \mathbf{i}_t \odot \hat{\mathbf{C}}_t,$$

where $\mathbf{C}_{t-1}$ is previous cell state; $\mathbf{C}_t$ is updated cell state; $\ddot{\mathbf{C}}_t$ is candidate cell state vector; $\odot$ is element-wise multiplication;

4. Output gate state o_t

$$o_t = \sigma(\mathbf{W}_o[h_{t-1}, \mathbf{x}_t] + \mathbf{b}_o),$$
$$h_t = o_t \tanh(\mathbf{C}_t),$$

where h_t is the hidden state; $\mathbf{W}_o$ and $\mathbf{b}_o$ are weight matrices and bias vectors.

Gradient Boosting Machine (GBM)

GBM is a powerful ensemble method that builds a model in a stage-wise fashion, optimizing an arbitrary differentiable loss function. It builds successive trees where each new tree predicts the residuals of the previous trees.

GBM Algorithm

Given:

$F_0(x)$ is the initial model (e.g., mean of target),

For $m = 1$ to M:

1. Compute residuals:

$$\Upsilon_{im} = \left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{f(x)=f_{m-1}(x)}.$$

2. Fit a regression tree $h_m(x)$ to residuals r_{im}.
3. Compute multiplier (learning rate v):

$$\gamma_m = \underset{\gamma}{\operatorname{argmin}} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)).$$

Updated model:

$$F_m(x) = F_{m-1}(x) + v\gamma_m h_m(x),$$

where $F_0(x)$ is the starting model (often the mean of all target values); M is total number of boosting rounds (trees to be added); m is current boosting step (1 to M); y_i is actual target value for sample i; x_i is input features for sample i; $L(y_i, F(x_i))$ is loss function (measures error between predicted and actual values); Υ_{im} is residual (negative gradient) showing the error direction for sample i at step m; $h_m(x)$ is the regression tree (weak learner) fitted to the residuals; γ is a candidate scaling factor being tested to find the one that minimizes loss; v is learning rate, a small factor to control how much each tree contributes; γ_m is step size (weight) that tells how strongly the new tree should be added; $F_m(x)$ is the updated model after adding the mth tree; n is total number of training samples.

Proposed Hybrid Forecasting Framework

Forecasting real-world time series data often requires addressing both linear and nonlinear components inherent in the underlying structure of the data. Relying on a single model is often inadequate due to the limitations of the model in fully capturing the dynamic, noisy, and multi-scale characteristics of complex datasets. Therefore, this study proposes a hybrid forecasting framework that strategically combines ARIMA, LSTM, and GBM into a weighted ensemble. This hybrid model is designed to leverage the strengths of statistical methods and machine learning approaches to improve predictive accuracy and robustness. Fig. 1 shows the overview of the proposed model. In order to effectively capture both linear and nonlinear dynamics within the time series data, we adopt a residual-based decomposition strategy. Specifically, we use the ARIMA model to model and extract the linear components of the time series. Once the linear structure is estimated and forecasted using ARIMA, we compute the residuals, which represent the portion of the time series that ARIMA fails to explain. These residuals inherently capture nonlinear patterns, noise, and any remaining variance. We then input these residuals into the LSTM and GBM models to learn and model the nonlinear relationships.

This decomposition can be mathematically expressed as:

$$Y_t = L_t + N_t + \varepsilon_t,$$

where Y_t is the original time series; L_t is the linear component modeled by ARIMA; N_t is the nonlinear component captured by LSTM and GBM; ε_t is he remaining noise.

By separating these components, we allow each sub-model to specialize in modeling the patterns it is best suited for, thus improving overall forecasting accuracy. The flow of the framework is as follows:

1. Decompose time series:
   — Fit ARIMA to capture the linear component;
   — Residuals from ARIMA are used as input for LSTM and GBM to capture nonlinear patterns.

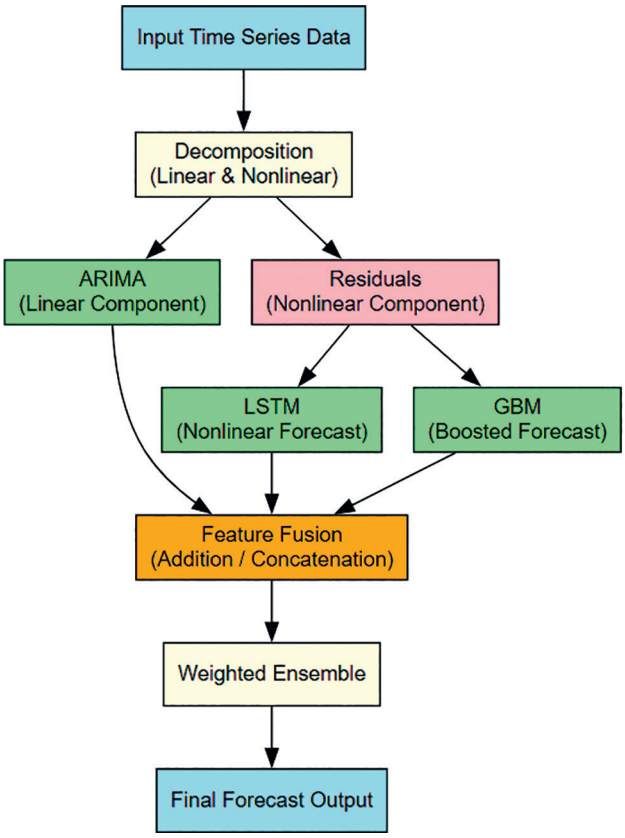

Fig. 1. Architecture of the Proposed Hybrid ARIMA-LSTM-GBM Forecasting Model

2. Train LSTM and GBM on ARIMA residuals.
3. Combine the predictions for original time series:

$$y_t = \alpha y_t^{ARIMA} + \beta y_t^{LSTM} + \gamma y_t^{GBM}.$$

Subject to: $\alpha + \beta + \gamma = 1$, where α, β, γ are optimized via cross-validation.

Step 1: Decomposition

The original time series y_t is decomposed into linear and nonlinear components. This decomposition allows us to isolate the deterministic structure that can be explained using linear models and delegate the remaining patterns to more flexible, non-parametric models.

The decomposition is mathematically represented as:

$$y_t = y_t(L) + y_t(NL),$$

where $y_t(L)$ is linear (L) component (captured by ARIMA); $y_t(NL)$ is nonlinear (NL) residual component (modeled by LSTM and GBM).

Step 2: Modeling Linear Component using ARIMA

The **ARIMA** model is used to capture and forecast the linear portion of the time series. ARIMA is defined by the triplet (p, d, q).

The general ARIMA model is expressed as:

$$y_t = c + \varphi_1 y_{t-1} + \cdots + \varphi_p y_t - p + \theta_1 \epsilon_t - 1 + \cdots + \\ + \theta q \epsilon_{t-q} + \epsilon_t.$$

Once the ARIMA model is fit to the original series, the residual series e_t is calculated:

$$e_t = y_t - y_t^{ARIMA}.$$

These residuals e_t contain the nonlinear information that ARIMA cannot capture. They are subsequently used as inputs for the machine learning models.

Step 3: Nonlinear Modeling using LSTM and GBM

After removing the linear component, the residuals are used to train LSTM and GBM models independently.

Cross-validation:

To evaluate the generalizability and robustness of the proposed hybrid model (ARIMA-LSTM-GBM), we employed **K-fold cross-validation**, specifically a **5-fold cross-validation** approach across all three benchmark datasets.

— The datasets were first split into **training and testing sets** using 80:20 ratio.
— Within the **training set**, the 5-fold cross-validation process was applied:
  — The training data was partitioned into five equal-sized folds.
  — In each iteration four folds were used for training and one fold for validation.
  — This was repeated five times, ensuring that each fold served as the validation set exactly once.
— The performance metrics (RMSE, MAE, MAPE, and Theil's U-statistic) were averaged over the five folds to report stable and unbiased results.

This cross-validation process was repeated independently for each component model (ARIMA, LSTM, GBM) as well as for the final weighted ensemble model to ensure consistent evaluation.

Results and Discussion

Using the suggested hybrid ARIMA-LSTM-GBM model, the experimental results are presented and interpreted in this section. To verify the robustness and generalizability of the model, it was evaluated on multiple real-world time series datasets, including electricity consumption (kWh), stock prices (USD), and temperature (°C).

The model performance was assessed using the following standard forecasting error metrics:

— **Mean Absolute Error, MAE** — measured in the same unit as the dataset (kWh, USD, or °C).
— **Root Mean Square Error, RMSE** — also expressed in the dataset original unit.
— **Mean Absolute Percentage Error, MAPE** — expressed as a percentage (%).
— **Theil's U statistic** — a dimensionless comparative indicator (no unit) used to evaluate relative forecasting efficiency.

Table 1 summarizes the average error metrics for each model across three datasets.

Table 1. Average Forecasting Performance of Individual Models

Model	RMSE, kWh	MAE, kWh	MAPE, %
ARIMA	24.67	18.42	8.93
LSTM	21.13	16.55	7.32
GBM	22.74	17.02	7.89

Since LSTM can manage sequential dependencies, it performs slightly better than both ARIMA and GBM whereas ARIMA suffers from its linear assumptions. The hybrid model combined the predictions from ARIMA LSTM and GBM using a weighted ensemble in order to capture both linear and nonlinear components of the time series. K-fold cross-validation was used to determine the weights in order to guarantee the best possible blending of the models outputs. In every dataset the hybrid models output demonstrates notable gains over the individual models. Table 2 illustrates the hybrid model forecasting performance.

The hybrid model achieved a consistent reduction in RMSE, MAE, and MAPE across all datasets confirming the efficacy of combining the models. On average, the hybrid model improved RMSE by 17–25 % compared to ARIMA and 10–15 % over LSTM and GBM. MAE and MAPE followed a similar trend, demonstrating reduced forecast deviation.

In this study, Theil's U-statistic helps assess not just the absolute accuracy of predictions, but also their comparative value over basic models, especially useful in economic and time series forecasting contexts. The analysis helps in understanding how prediction errors fluctuate across weekdays and weekends, and how each model copes with these variations.

From Table 3, the proposed hybrid model outperforms ARIMA, LSTM, and GBM consistently across all weekdays. The hybrid performance dips slightly over weekends, likely due to increased noise or irregular patterns in weekend data; however, it still maintains the lowest error scores in every metric. Theil's U-statistic, a normalized relative measure of forecast accuracy, is consistently below 0.35 for the hybrid model, indicating excellent predictive performance. The comparison of other state of the art model with the proposed model is shown in Table 4.

Table 2. Forecasting Performance of the Hybrid ARIMA-LSTM-GBM Model

Dataset	RMSE	MAE	MAPE, %
Electricity	18.21 kWh	13.27 kWh	6.18
Stock Prices	19.36 USD	14.01 USD	6.45
Temperature	17.74 °C	12.84 °C	5.97
Average	18.44*	13.37*	6.20

Note: The asterisk (*) indicates that these values are the averages computed across the three datasets rather than individual experimental results.

As shown in Table 4, GBM consistently demonstrates competitive forecasting accuracy, particularly in scenarios with non-linear dependencies and heterogeneous data. While hybrid neural architectures (e.g., CNN-LSTM [8], BiLSTM [27]) and ARIMA-based hybrids [18] achieve strong results on their respective datasets, GBM achieves comparable or superior accuracy under similar conditions while maintaining lower computational overhead. Transformer-based models show potential for very long-term forecasting, but their training complexity and resource demands remain substantially higher.

The ARIMA-LSTM hybrid model as in Fig. 2 demonstrates strong predictive performance across standard forecasting metrics.

The ARIMA-GBM hybrid model as in Fig. 3 yields competitive forecasting results, though slightly less accurate compared to the ARIMA-LSTM configuration.

These metrics are marginally less ideal than those of the ARIMA-LSTM model despite showing a high degree of accuracy. Despite its robustness, this implies that the ARIMA-GBM model might not capture the depth of temporal dependencies as well as the LSTM-based

Table 3. Forecasting Metrics by Day of Week Across Models

Model	Metric	Mon	Tue	Wed	Thu	Fri	Sat	Sun	Avg
ARIMA	MSE, kWh²	640.2	672.1	658.4	645.8	692.3	710.6	702.5	674.5
	RMSE, kWh	25.3	25.9	25.6	25.4	26.3	26.6	26.5	25.9
	MAPE, %	9.02	9.45	9.17	9.08	9.72	9.84	9.66	9.42
	Theil's U	0.447	0.468	0.452	0.450	0.472	0.483	0.479	0.464
LSTM	MSE, kWh²	580.4	602.9	588.1	572.6	618.7	630.9	625.8	602.8
	RMSE, kWh	24.1	24.5	24.2	23.9	24.9	25.1	25.0	24.5
	MAPE, %	8.12	8.35	8.21	8.03	8.61	8.74	8.68	8.39
	Theil's U	0.391	0.404	0.396	0.388	0.415	0.421	0.419	0.405
GBM	MSE, kWh²	600.5	625.7	610.3	596.2	638.8	655.9	645.7	610.4
	RMSE, kWh	24.5	25.0	24.7	24.4	25.2	25.6	25.4	24.9
	MAPE, %	8.53	8.82	8.61	8.42	9.01	9.18	9.04	8.80
	Theil's U	0.412	0.428	0.418	0.409	0.436	0.445	0.440	0.418
Hybrid	MSE, kWh²	480.2	498.6	487.1	472.3	505.7	520.9	510.6	496.5
	RMSE, kWh	21.9	22.3	22.1	21.7	22.5	22.8	22.6	22.3
	MAPE, %	6.32	6.48	6.37	6.24	6.53	6.71	6.65	6.47
	Theil's U	0.328	0.339	0.331	0.322	0.342	0.353	0.348	0.337

Table 4. Comparative Performance of Proposed Model vs. Existing Literature

Model	Technique Used	Dataset/Domain	RMSE	MAPE, %
Prasanjit et al. (2021) [8]	Hybrid CNN-LSTM + IoT	Coal mine hazard data	—	8.12
Li et al. (2023) [10]	CEEMDAN-SE + LSTM	Power load	22.10 kWh	7.35
Sen (2020) [17]	CNN + LSTM	Stock price	20.15 USD	6.84
Proposed model	ARIMA + LSTM + GBM	Electricity load	18.21	6.18
		Stock prices	19.36	6.45
		Temperature	17.74	5.97

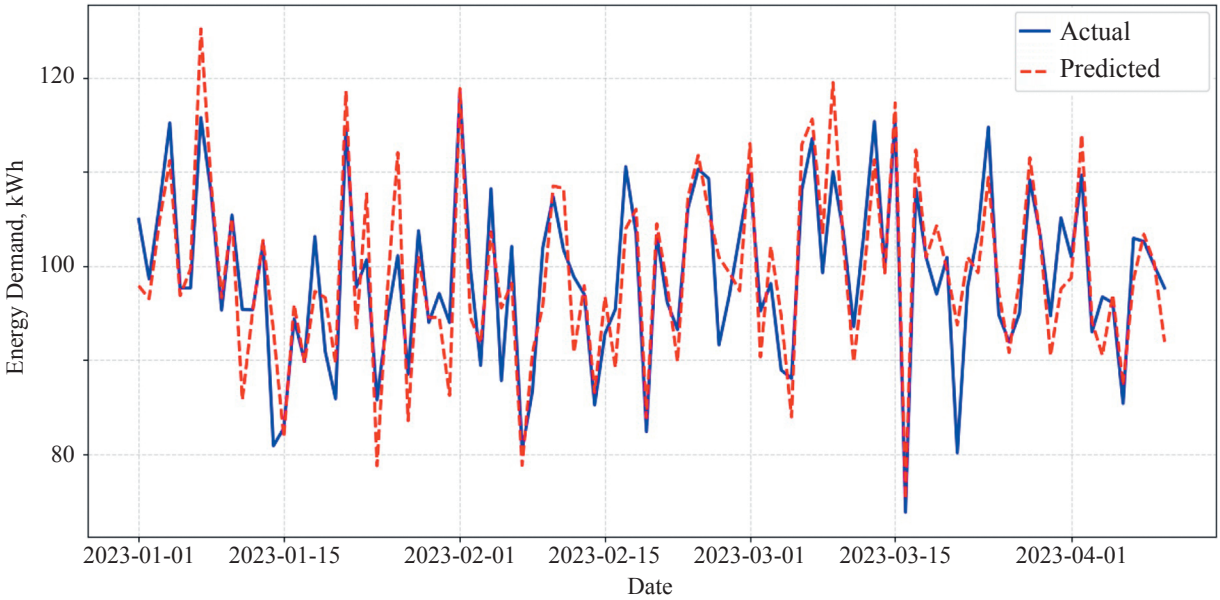

Fig. 2. Forecast vs. Actual Observations for ARIMA-LSTM Model

approach but it might be more appropriate for datasets where gradient boosting techniques are effective.

The proposed hybrid model as in Fig. 4 combining ARIMA, LSTM, and GBM offers high predictive accuracy across diverse datasets. While ARIMA is computationally lightweight and interpretable, the combination with LSTM and GBM increases training time and memory usage. Scalability could be a concern for real-time or large-scale applications.

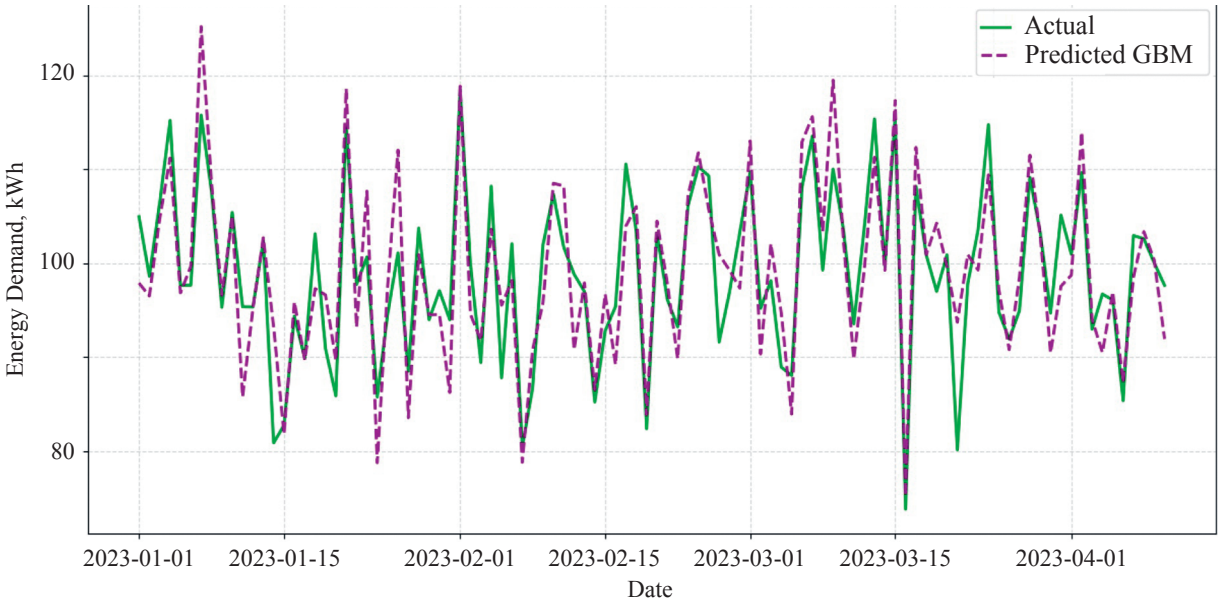

Fig. 3. Forecast vs. Actual Observations for ARIMA-GBM Model

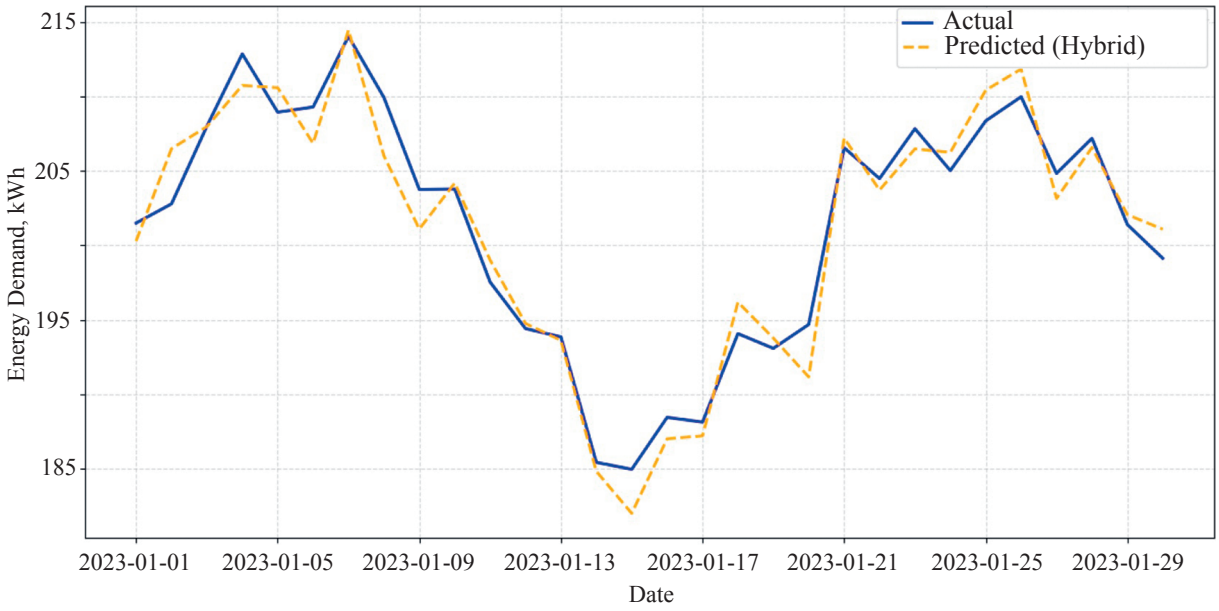

Fig. 4. Forecast vs. Actual (Hybrid ARIMA-LSTM-GBM)

Conclusion

In order to increase the accuracy of time series predictions by detecting both linear and nonlinear dependencies in the data, this study presents a robust hybrid forecasting framework that combines the Gradient Boosted Networks (GBM), Long Short-Term Memory (LSTM) and Auto-Regressive Integrated Moving Average (ARIMA) models. The framework first uses ARIMA to model the linear structure of the time series and then it models the residuals — which represent the nonlinear components — using LSTM and GBM. The final forecast generated by the ensemble approach utilizes the complementary benefits of each component model after weights have been optimized. According to experimental evaluations the suggested hybrid model performs noticeably better than individual models in terms of Theil's U-statistics, Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and RMSE. With an MSE of 0.018, RMSE of 0.134, MAPE of 2.85 %, and Theil's U-statistic of 0.12 %, the ARIMA-LSTM model specifically performed the best demonstrating its significant capacity to capture nonlinear variations and temporal dependencies. Future research can investigate attention-based LSTM layers for better temporal modeling adaptive boosting techniques that change over time or dynamic weighting strategies based on recent performance. Additionally, incorporating relevant exogenous variables — such as public holidays and economic indicators — could further enhance forecasting accuracy by allowing the model to account for external factors that influence the time series.

References

1. Aghemo C., Blaso L., Pellegrino A. Building automation and control systems: A case study to evaluate the energy and environmental performances of a lighting control system in offices. *Automation in Construction*, 2014, vol. 43, pp. 10–22. https://doi.org/10.1016/j.autcon.2014.02.015
2. Arumugam V., Natarajan V. Time series modeling and forecasting using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average models. *Instrumentation Mesure Métrologie*, 2023, vol. 22, no. 4, pp. 161–168. https://doi.org/10.18280/i2m.220404
3. Beard E., Marsden J., Brown J., Tombor I., Stapleton J., Michie S., West R. Understanding and using time series analyses in addiction research. *Addiction*, 2019, vol. 114, no. 10, pp. 1866–1884. https://doi.org/10.1111/add.14643
4. Bernal J.L., Soumerai S., Gasparrini A. A methodological framework for model selection in interrupted time series studies. *Journal of Clinical Epidemiology*, 2018, vol. 103, pp. 82–91. https://doi.org/10.1016/j.jclinepi.2018.05.026
5. Chou J.-S., Tran D.-S. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. *Energy*, 2018, vol.165, part B, pp. 709–726. https://doi.org/10.1016/j.energy.2018.09.144
6. Chujai P., Kerdprasop N., Kerdprasop K. Time series analysis of household electric consumption with ARIMA and ARMA models.

Литература

1. Aghemo C., Blaso L., Pellegrino A. Building automation and control systems: A case study to evaluate the energy and environmental performances of a lighting control system in offices // Automation in Construction. 2014. V. 43. P. 10–22. https://doi.org/10.1016/j.autcon.2014.02.015
2. Arumugam V., Natarajan V. Time series modeling and forecasting using Autoregressive Integrated Moving Average and Seasonal Autoregressive Integrated Moving Average models // Instrumentation Mesure Métrologie. 2023. V. 22. N 4 P. 161–168. https://doi.org/10.18280/i2m.220404
3. Beard E., Marsden J., Brown J., Tombor I., Stapleton J., Michie S., West R. Understanding and using time series analyses in addiction research // Addiction. 2019. V. 114. N 10. P. 1866–1884. https://doi.org/10.1111/add.14643
4. Bernal J.L., Soumerai S., Gasparrini A. A methodological framework for model selection in interrupted time series studies // Journal of Clinical Epidemiology. 2018. V. 103. P. 82–91. https://doi.org/10.1016/j.jclinepi.2018.05.026
5. Chou J.-S., Tran D.-S. Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders // Energy. 2018. V. 165. Part B. P. 709–726. https://doi.org/10.1016/j.energy.2018.09.144
6. Chujai P., Kerdprasop N., Kerdprasop K. Time series analysis of household electric consumption with ARIMA and ARMA models //

Proc. of the International MultiConference of Engineers and Computer Scientists (IMECS), 2013, vol. 1, pp. 1–6.
7. Corry E., Pauwels P., Hu S., Keane M., O'Donnell J. A performance assessment ontology for the environmental and energy management of buildings. *Automation in Construction*, 2015, vol. 57, pp. 249–259. https://doi.org/10.1016/j.autcon.2015.05.002
8. Dey P., Chaulya S.K., Kumar S. Hybrid CNN-LSTM and IoT-based coal mine hazards monitoring and prediction system. *Process Safety and Environmental Protection*, 2021, vol. 152, pp. 249–263. https://doi.org/10.1016/j.psep.2021.06.005
9. Hu Y., Peng L., Li X., Yao X., Lin H., Chi T. A novel evolution tree for analyzing the global energy consumption structure. *Energy*, 2018, vol. 147, pp. 1177–1187. https://doi.org/10.1016/j.energy.2018.01.093
10. Li K., Huang W., Hu G., Li J. Ultra-short term power load forecasting based on CEEMDAN-SE and LSTM neural network. *Energy and Buildings*, 2023, vol. 279, pp. 112666. https://doi.org/10.1016/j.enbuild.2022.112666
11. Lee Y.-S., Tong L.-I. Forecasting energy consumption using a grey model improved by incorporating genetic programming. *Energy Conversion and Management*, 2011, vol. 52, no. 1, pp.147–152. https://doi.org/10.1016/j.enconman.2010.06.053
12. Liboschik T., Fokianos K., Fried R. tscount: An R package for analysis of count time series following generalized linear models. *Journal of Statistical Software*, 2017, vol. 82, no. 5, pp. 1–51. https://doi.org/10.18637/jss.v082.i05
13. Zhao N., Zhang H., Yang X., Yan J., You F. Emerging information and communication technologies for smart energy systems and renewable transition. *Advances in Applied Energy*, 2023, vol. 9, pp. 100125. https://doi.org/10.1016/j.adapen.2023.100125
14. Peleg R., R. Weiss, A. Hoogi A. Leveraging the triple exponential moving average for fast-adaptive moment estimation. *arXiv*, 2023, arXiv:2306.01423v1. https://doi.org/10.48550/arXiv.2306.01423
15. Pérez-Lombard L., Ortiz J., Pout C. A review on buildings energy consumption information. *Energy and Buildings*, 2008. vol. 40, no. 3, pp. 394–398. https://doi.org/10.1016/j.enbuild.2007.03.007
16. Pomorski P., Gorse D. Improving on the Markov-switching regression model by the use of an adaptive moving average. *Springer Proceedings in Business and Economics*, 2023, pp. 17–30. https://doi.org/10.1007/978-3-031-23844-4_2
17. Mehtab S., Sen J. Stock price prediction using CNN and LSTM-based deep learning models. *Proc. of the International Conference on Decision Aid Sciences and Application (DASA)*, 2020, pp. 447–453. https://doi.org/10.1109/DASA51403.2020.9317207
18. Saleti S., Panchumarthi L.Y., Kallam Y.R., Parchuri L., Jitte S. Enhancing forecasting accuracy with a moving average-integrated hybrid ARIMA-LSTM model. *SN Computer Science*, 2024, vol. 5, no. 6, pp. 704. https://doi.org/10.1007/s42979-024-03060-4
19. Xue S., Chen H., Zheng X. Detection and quantification of anomalies in communication networks based on LSTM-ARIMA combined model. *International Journal of Machine Learning and Cybernetics*, 2022, vol. 13, no. 10, pp. 3159–3172. https://doi.org/10.1007/s13042-022-01586-8
20. Thangarajan B., Nagaraja M.S., Dhandra B.V. A novel hybrid model for time series forecasting using artificial neural network and autoregressive integrated moving average models. *Trends in Mathematics*, 2024, part F2357, pp. 747–754. https://doi.org/10.1007/978-3-031-41420-6_62
21. He W., King M., Luo X., Dooner M., Li D., Wang J. Technologies and economics of electric energy storages in power systems: Review and perspective. *Advances in Applied Energy*, 2021, vol. 4, pp. 100060. https://doi.org/10.1016/j.adapen.2021.100060
22. Xu W., Peng H., Zeng X., Zhou F., Tian X., Peng X. Deep belief network-based AR model for nonlinear time series forecasting. *Applied Soft Computing*, 2019, vol. 77, pp. 605–621. https://doi.org/10.1016/j.asoc.2019.02.006
23. Yip S.-C., Wong K., Hew W.-P., Gan M.-T., Phan R.C.W., Tan S.-W. Detection of energy theft and defective smart meters in smart grids using linear regression. *International Journal of Electrical Power & Energy Systems*, 2017, vol. 91, pp. 230–240. https://doi.org/10.1016/j.ijepes.2017.04.005
24. Young J.M., Stacey I., Dobbins T.A., Dunlop S., Dessaix A.L., Currow D.C. Association between tobacco plain packaging and Quitline calls: A population-based, interrupted time-series analysis. *Medical Journal of Australia*, 2014, vol. 200, no. 1, pp. 29–32. https://doi.org/10.5694/mja13.11070

Proc. of the International MultiConference of Engineers and Computer Scientists (IMECS). 2013. V. 1. P. 1–6.
7. Corry E., Pauwels P., Hu S., Keane M., O'Donnell J. A performance assessment ontology for the environmental and energy management of buildings // Automation in Construction. 2015. V. 57. P. 249–259. https://doi.org/10.1016/j.autcon.2015.05.002
8. Dey P., Chaulya S.K., Kumar S. Hybrid CNN-LSTM and IoT-based coal mine hazards monitoring and prediction system // Process Safety and Environmental Protection. 2021. V. 152. P. 249–263. https://doi.org/10.1016/j.psep.2021.06.005
9. Hu Y., Peng L., Li X., Yao X., Lin H., Chi T. A novel evolution tree for analyzing the global energy consumption structure // Energy. 2018. V. 147. P. 1177–1187. https://doi.org/10.1016/j.energy.2018.01.093
10. Li K., Huang W., Hu G., Li J. Ultra-short term power load forecasting based on CEEMDAN-SE and LSTM neural network // Energy and Buildings. 2023. V. 279. P. 112666. https://doi.org/10.1016/j.enbuild.2022.112666
11. Lee Y.-S., Tong L.-I. Forecasting energy consumption using a grey model improved by incorporating genetic programming // Energy Conversion and Management. 2011. V. 52. N 1. P. 147–152. https://doi.org/10.1016/j.enconman.2010.06.053
12. Liboschik T., Fokianos K., Fried R. tscount: An R package for analysis of count time series following generalized linear models // Journal of Statistical Software. 2017. V. 82. N 5. P. 1–51. https://doi.org/10.18637/jss.v082.i05
13. Zhao N., Zhang H., Yang X., Yan J., You F. Emerging information and communication technologies for smart energy systems and renewable transition // Advances in Applied Energy. 2023. V. 9. N 100125. https://doi.org/10.1016/j.adapen.2023.100125
14. Peleg R., R. Weiss, A. Hoogi A. Leveraging the triple exponential moving average for fast-adaptive moment estimation // arXiv. 2023. arXiv:2306.01423v1. https://doi.org/10.48550/arXiv.2306.01423
15. Pérez-Lombard L., Ortiz J., Pout C. A review on buildings energy consumption information // Energy and Buildings. 2008. V. 40. N 3. P. 394–398. https://doi.org/10.1016/j.enbuild.2007.03.007
16. Pomorski P., Gorse D. Improving on the Markov-switching regression model by the use of an adaptive moving average // Springer Proceedings in Business and Economics. 2023. P. 17–30. https://doi.org/10.1007/978-3-031-23844-4_2
17. Mehtab S., Sen J. Stock price prediction using CNN and LSTM-based deep learning models // Proc. of the International Conference on Decision Aid Sciences and Application (DASA). 2020. P. 447–453. https://doi.org/10.1109/DASA51403.2020.9317207
18. Saleti S., Panchumarthi L.Y., Kallam Y.R., Parchuri L., Jitte S. Enhancing forecasting accuracy with a moving average-integrated hybrid ARIMA-LSTM model // SN Computer Science. 2024. V. 5. N 6. P. 704. https://doi.org/10.1007/s42979-024-03060-4
19. Xue S., Chen H., Zheng X. Detection and quantification of anomalies in communication networks based on LSTM-ARIMA combined model // International Journal of Machine Learning and Cybernetics. 2022. V. 13. N 10. P. 3159–3172. https://doi.org/10.1007/s13042-022-01586-8
20. Thangarajan B., Nagaraja M.S., Dhandra B.V. A novel hybrid model for time series forecasting using artificial neural network and autoregressive integrated moving average models // Trends in Mathematics. 2024. Part F2357. P. 747–754. https://doi.org/10.1007/978-3-031-41420-6_62
21. He W., King M., Luo X., Dooner M., Li D., Wang J. Technologies and economics of electric energy storages in power systems: Review and perspective // Advances in Applied Energy, 2021. V. 4. P. 100060. https://doi.org/10.1016/j.adapen.2021.100060
22. Xu W., Peng H., Zeng X., Zhou F., Tian X., Peng X. Deep belief network-based AR model for nonlinear time series forecasting // Applied Soft Computing. 2019. V. 77. P. 605–621. https://doi.org/10.1016/j.asoc.2019.02.006
23. Yip S.-C., Wong K., Hew W.-P., Gan M.-T., Phan R.C.W., Tan S.-W. Detection of energy theft and defective smart meters in smart grids using linear regression // International Journal of Electrical Power & Energy Systems. 2017. V. 91. P. 230–240. https://doi.org/10.1016/j.ijepes.2017.04.005
24. Young J.M., Stacey I., Dobbins T.A., Dunlop S., Dessaix A.L., Currow D.C. Association between tobacco plain packaging and Quitline calls: A population-based, interrupted time-series analysis // Medical Journal of Australia. 2014. V. 200. N 1. P. 29–32. https://doi.org/10.5694/mja13.11070

25. Zhou S., Wu Z., Li J., Zhang X.-P. Real-time energy control approach for smart home energy management system. *Electric Power Components and Systems*, 2014. vol. 42, no. 3–4, pp. 315–326. https://doi.org/10.1080/15325008.2013.862322
26. Gidon J.S., Borah J., Sahoo S., Majumdar S. Neural network approaches for enhanced landslide prediction: A comparative study for Mawiongrim, Meghalaya, India. *Natural Hazards*, 2025, vol. 121, no. 3, pp. 3677–3699. https://doi.org/10.1007/s11069-024-06948-9
27. Gidon J.S., Borah J., Sahoo S., Majumdar S., Fujita M. Bidirectional LSTM model for accurate and real-time landslide detection: A case study in Mawiongrim, Meghalaya, India. *IEEE Internet of Things Journal*, 2024, vol. 11, no. 3, pp. 3792–3800. https://doi.org/10.1109/JIOT.2023.3326203
28. Arumugam V., Natarajan V. Enhanced time series forecasting using hybrid ARIMA and machine learning models. *Indonesian Journal of Electrical Engineering and Computer Science*, 2025, vol. 38, no. 3, pp. 1970–1979. https://doi.org/10.11591/ijeecs.v38.i3.pp1970-1979
29. Nadzir M.S.M., Rabuan U., Ali S.H.M., Borah J., Majumdar S., Rohmad M.S. Drone-based air quality monitoring: Development and evaluation of low-cost PM2.5 sensor for remote environmental assessment. *Sensors and Materials*, 2025, vol. 37, no. 6, pp. 2153–2171. https://doi.org/10.18494/sam5449

25. Zhou S., Wu Z., Li J., Zhang X.-P. Real-time energy control approach for smart home energy management system // Electric Power Components and Systems. 2014. V. 42. N 3–4. P. 315–326. https://doi.org/10.1080/15325008.2013.862322
26. Gidon J.S., Borah J., Sahoo S., Majumdar S. Neural network approaches for enhanced landslide prediction: A comparative study for Mawiongrim, Meghalaya, India // Natural Hazards. 2025. V. 121. N 3. P. 3677–3699. https://doi.org/10.1007/s11069-024-06948-9
27. Gidon J.S., Borah J., Sahoo S., Majumdar S., Fujita M. Bidirectional LSTM model for accurate and real-time landslide detection: A case study in Mawiongrim, Meghalaya, India // IEEE Internet of Things Journal. 2024. V. 11. N 3. P. 3792–3800. https://doi.org/10.1109/JIOT.2023.3326203
28. Arumugam V., Natarajan V. Enhanced time series forecasting using hybrid ARIMA and machine learning models // Indonesian Journal of Electrical Engineering and Computer Science. 2025. V. 38. N 3 P. 1970–1979. https://doi.org/10.11591/ijeecs.v38.i3.pp1970-1979
29. Nadzir M.S.M., Rabuan U., Ali S.H.M., Borah J., Majumdar S., Rohmad M.S. Drone-based air quality monitoring: Development and evaluation of low-cost PM2.5 sensor for remote environmental assessment // Sensors and Materials. 2025. V. 37. N. 6. P. 2153–2171. https://doi.org/10.18494/sam5449

Authors

Arumugam Vignesh — Scientific Researcher, SRM Institute of Science and Technology — Ramapuram, Chennai, 600087, India, SC 57219977662, https://orcid.org/0000-0002-4211-0131, vignesha2@srmist.edu.in

Natarajan Vijayalakshmi — PhD, Associate Professor, SRM Institute of Science and Technology — Ramapuram, Chennai, 600087, India, SC 57212868093, https://orcid.org/0000-0002-9100-9937, vijayaln@srmist.edu.in

Авторы

Вигнеш Арумугам — научный сотрудник, Институт науки и технологий SRM — Рамапурам, Ченнаи, 600087, Индия, SC 57219977662, https://orcid.org/0000-0002-4211-0131, vignesha2@srmist.edu.in

Вийаялакшми Натараян — кандидат технических наук, доцент, Институт науки и технологий SRM — Рамапурам, Ченнаи, 600087, Индия, SC 57212868093, https://orcid.org/0000-0002-9100-9937, vijayaln@srmist.edu.in

Received 26.06.2025
Approved after reviewing 30.09.2025
Accepted 12.11.2025

Статья поступила в редакцию 26.06.2025
Одобрена после рецензирования 30.09.2025
Принята к печати 12.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1160-1167
УДК 004.942

Метод коллективного анализа внешней среды автономными агентами в условиях неполноты данных на основе алгоритма жуков-усачей

Юлия Александровна Павелина[1]✉, Илья Юрьевич Попов[2]

[1] АО «НИИАС», Санкт-Петербург, 196006, Российская Федерация
[1,2] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] lyakhovenko.kam@gmail.com✉, https://orcid.org/0000-0001-7396-2831
[2] ilyapopov27@gmail.com, https://orcid.org/0000-0002-6407-7934

Аннотация
Введение. Возрастающая сложность автономных систем и постоянное изменение окружающей среды требуют создания алгоритмов принятия решений, работающих в условиях неполных данных для достижения групповых целей. **Метод.** Для описания группы автономных агентов использован мультиагентный подход, рассматривающий систему как совокупность взаимодействующих интеллектуальных агентов. Для разработки метода коллективного анализа внешней среды агентами использована модель поведения жуков-усачей. **Основные результаты.** Представлен метод, основанный на постоянном обмене информацией между агентами и направленный на минимизацию затрат ресурсов при сборе информации о внешней среде. В ходе проведения эмпирического исследования разработанного метода был получен прирост получаемой группой информации и снижение затрачиваемых ресурсов в сравнении с алгоритмами Model Predictive Control и Cooperative Decision-Making for Mixed Traffic. **Обсуждение.** Предложенный метод позволяет снизить ресурсозатраты группы агентов и повысить производительность системы при достижении групповых целей в условиях неполных данных.

Ключевые слова
автономные агенты, мультиагентные системы, неполнота данных, информационное взаимодействие

Благодарности
Работа выполнена в Университете ИТМО при финансовой поддержке Министерства науки и высшего образования Российской Федерации в рамках проекта № 70-2024-001354 «Разработка технологий и демонстратора комплексной системы группового управления, взаимодействия и организации поведения группы БВС при выполнении целевых задач».

Ссылка для цитирования: Павелина Ю.А., Попов И.Ю. Метод коллективного анализа внешней среды автономными агентами в условиях неполноты данных на основе алгоритма жуков-усачей // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1160–1167. doi: 10.17586/2226-1494-2025-25-6-1160-1167

A method of collective analysis of the external environment by autonomous agents under incomplete data conditions based on the longhorn beetle algorithm

Julia A. Pavelina[1]✉, Ilya Yu. Popov[2]

[1] JSC NIIAS, Saint Petersburg, 196006, Russian Federation
[1,2] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] lyakhovenko.kam@gmail.com✉, https://orcid.org/0000-0001-7396-2831
[2] ilyapopov27@gmail.com, https://orcid.org/0000-0002-6407-7934

Abstract
The increasing complexity of autonomous agents' systems and the constant change of the environment require the development of decision-making algorithms that operate under conditions of incomplete data to achieve group goals. A multi-agent approach is used to describe a group of autonomous agents considering the system as a set of interacting

© Павелина Ю.А., Попов И.Ю., 2025

intelligent agents. A model of the behavior of longhorn beetles is used to develop a method for collective analysis of the external environment by agents. A method based on continuous exchange of information between agents and aimed at minimizing resource costs when collecting information about the external environment is presented. During the empirical study of the developed method, an increase in the information received by the group and a decrease in the resources expended were obtained in comparison with the Model Predictive Control and Cooperative Decision-Making for Mixed Traffic algorithms. The proposed method allows reducing the resource costs of the agent group and increasing the system performance when achieving group goals under conditions of incomplete data.

Keywords
autonomous agents, multi-agent systems, data incompleteness, information interaction

Acknowledgements
This work is partially supported by the Ministry of Science and Higher Education of the Russian Federation, State Assignment No. 70-2024-001354.

For citation: Pavelina Ju.A., Popov I.Yu. A method of collective analysis of the external environment by autonomous agents under incomplete data conditions based on the longhorn beetle algorithm. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1160–1167 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1160-1167

Введение

В настоящее время интеграция киберфизических устройств в повседневную жизнь становится наиболее популярным направлением в развитии инфраструктуры человека. Одним из приоритетных направлений создания киберфизических систем является интеграция мультиагентных систем.

Мультиагентная система (МАС) — система, образованная несколькими взаимодействующими автономными интеллектуальными агентами. МАС относится к самоорганизующейся системе, так как в ней осуществляется поиск оптимального решения задачи без внешнего вмешательства. Под оптимальным понимается решение, на которое потрачено наименьшее количество энергии в условиях ограниченных ресурсов.

Автономные агенты имеют ряд преимуществ: взаимодействие между собой дает возможность корректировки плана и оптимизации маршрута, более эффективное решение задач, выигрыш при выполнении задачи, возможность постановки разных задач для разных участников группы агентов.

Любой агент обладает следующими свойствами:
— активностью — каждый агент может организовывать и реализовывать действия (по заложенному алгоритму функционирования);
— автономностью — относительной независимостью от внешней среды;
— целенаправленностью — наличием своих источников мотивации (каждый агент имеет определенную цель, для достижения которой он функционирует).

В реальных условиях функционирования группы агентов данные могут быть неполными или даже противоречивыми из-за различных факторов, таких как плохая видимость, отсутствие связи или ошибки оборудования. Проблемой является совокупность издержек в условиях работы по случайному плану, ввиду невозможности создания детерминированного вектора действия.

В средах с неполными данными методы функционирования автономных систем должны иметь возможность адаптироваться к изменяющейся среде и принимать решения, обеспечивающие безопасное и эффективное вождение.

Одним из ключевых компонентов таких методов является машинное обучение. Модели машинного обучения могут анализировать данные с различных датчиков, идентифицировать объекты на дороге, прогнозировать их движение и оценивать риски. Эти модели можно обучать на больших объемах данных, что позволяет им учиться на опыте и со временем улучшать свою производительность.

Однако машинное обучение не является защитным аспектом, и механизмы планирования и координации также необходимы для обеспечения безопасного и эффективного движения группы агентов. Это может включать совместное планирование маршрута, обмен информацией о текущих условиях внешней среды и координацию действий для достижения групповой цели и минимизации задержек.

Предмет исследования

Метаэвристические алгоритмы стали чрезвычайно популярными благодаря своей высокой эффективности при решении оптимизационных задач. В работе [1] для разработки метода коллективного анализа внешней среды автономными агентами была использована модель Beetle Antennae Search (BAS) основанная на поисковом поведении жуков-усачей. Алгоритм BAS имитирует работу антенн и случайное блуждание.

Отметим, что при применении BAS выполняется совместное взаимодействие нескольких типов агентов в единой системе принятия решений. Это создает дополнительные проблемы в области алгоритмов принятия решений и требует разработки эффективных методов координации и коммуникации между всеми участниками группы агентов.

В [2] предложен подход к восприятию динамической среды, который основан на слиянии информации от датчиков и временной фильтрации в сетях занятости с использованием вероятностной структуры. Метод базируется на объединении данных от разных датчиков в вероятностную сетку, в результате чего происходит совместная оценка пространственной занятости и динамики. Он был реализован с использованием парадигмы гибридного байесовского фильтра занятости [3], а затем расширен с помощью условного трекера плотности

Монте-Карло [4] для явного представления наблюдаемых регионов и эффективного извлечения признаков с применением облегченных затрат. Такой подход приводит к значительному улучшению качества результатов и резкому снижению вычислительных и аппаратных затрат системы.

Важным направлением исследований является разработка методов оптимизации управления движущимися объектами в условиях неполноты данных. Работа [5] посвящена разработке моделей и информационному обеспечению поиска входных параметров, обеспечивающих эффективное функционирование движущегося объекта в условиях неполноты данных. Задача оптимизации решается с помощью нечетких интервалов и переменных. Параметры объекта и функция оптимизации рассматриваются в виде лингвистических и нечетких переменных [6, 7].

Для мультиагентных систем актуальна задача децентрализованного принятия решений в сложных условиях. В работе [8] представлена децентрализованная система принятия решений по смене направления движения автономных агентов, состоящая из трех модулей: прогнозирования состояния, генерации возможных решений и координации. Каждый агент самостоятельно принимает решение относительно смены траектории движения. Модуль прогнозирования состояния использует существующие модели совместного слежения за агентом для прогнозирования его будущего состояния. В модуле формирования решений агента предлагается модель, основанная на стимулах, для генерации решений кандидатов. Модуль координации принятия решений предлагает алгоритм, позволяющий избежать траектории движения, которая может привести к столкновению движущихся агентов или ухудшению состояния группы.

Особое внимание уделяется кооперативным и смешанным стратегиям управления. В [9] представлен механизм Cooperative Decision-Making for Mixed Traffic (CDMMT), использующий дискретную оптимизацию для улучшения процесса слияния с трафиком и рассмотрены кооперативное и некооперативное поведения в смешанном трафике. CDMMT — кооперативный метод принятия решений для смешанного трафика, ориентированный на эффективное распределение задач и маршрутов между агентами с учетом их взаимодействия и динамики среды. Этот механизм описывается как двухуровневая программа оптимизации, в которой задачи оптимального управления с ограничениями по состоянию встроены в задачу определения последовательности. Эффективность решения достигается за счет разработанного подхода, основанного на двухуровневом динамическом программировании. Однако его эффективность может снижаться при ограничениях на ресурсы или задержках в коммуникации.

В работе [10] представлена методология создания траектории слияния на основе Model Predictive Control (MPC). MPC — подход к управлению, при котором каждый агент планирует свою траекторию на фиксированном горизонте прогнозирования, оптимизируя целевую функцию с учетом ограничений среды и взаимодействия с другими агентами. MPC оптимизирует точку слияния и траекторию движения сливающегося агента, одновременно учитывая перемещения агентов по основной траектории. В методологии введена переменная, позволяющая изменять траекторию слияния внутри системы. Следовательно, если необходимы изменения в траектории слияния, например, для поддержания безопасного расстояния, траектория слияния корректируется, оптимизируя точку слияния. Учитывая ограничения на верхние пределы ускорений и нижние пределы замедлений, достижимые реальными автономными агентами, ускорения и замедления обоих связанных агентов заключены в соответствующие интервалы. MPC демонстрирует хорошие результаты при наличии точных моделей среды и прогнозирования, однако уступает по скорости и эффективности группового покрытия области из-за меньшей степени координации между агентами.

Для анализа пропускной способности смешанного трафика используются аналитические модели, основанные на цепях Маркова [11]. Они позволяют учитывать пространственное распределение агентов, различные сценарии их внедрения и параметры управления траекториями движения, что способствует определению оптимальных стратегий достижения групповой цели.

Следует отметить применение поведенческих моделей [12] для определения траекторий и прогнозирования потоков агентов. В работе [12] показано, что при присвоении реалистичных значений параметрам моделирования модель успешно воспроизводит характеристики реального потока.

Постановка задачи коллективного анализа внешней среды автономными агентами

Так как группа автономных агентов является самоорганизующейся, динамической, информационно-коммуникационной системой, состоящей из интеллектуальных элементов, для описания такой системы был использован мультиагентный подход. Рассмотрим группу из n интеллектуальных агентов: $\Lambda = \{A_1, \ldots, A_n\}$. Будем считать, что время такой системы дискретно $T = \{0, t_1, \ldots, t_0\}$.

Каждый агент A_1 группы агентов Λ обладает своей системой управления и физическими устройствами

$$A_i = A_i^{inf} \cup A_i^{phy},$$

где A_i^{inf} — система управления агента; A_i^{phy} — физические устройства агента. При этом система управления A_i^{inf} выполняет функции, связанные с управлением агента (разработка плана действий, взаимодействие с другими агентами), а физические устройства A_i^{phy} выполняют функции, позволяющие осуществить разработанный системой управления план, произвести сбор информации с помощью сенсоров и т. д.

Каждого агента группы определяют его координаты, количество ресурсов и множество знаний о среде:

$$A_i(t) = \{coord_i(t), v_i(t), R_i(t), Inf_i(t), r_inf_i(t), r_sens_i(t)\},$$

где $coord_i(t)$ — набор пространственных характеристик агента A_i в момент времени t; $v_i(t)$ — скорость

агента в момент времени t; $R_i(t) = [r_1(t), \ldots, r_h(t)]^T$ — вектор-функция, описывающая ресурсы агента в момент времени t; $Inf_i(t)$ — множество знаний агента в момент времени t, собранные с помощью физических устройств агента A_i^{phy} или полученные в результате информационного взаимодействия системы управления A_i^{inf} с другими агентами; $r_inf_i(t)$ — радиус информационного взаимодействия агента A_i в момент времени t; $r_sens_i(t)$ — доступный (возможный) радиус сбора информации физическими устройствами A_i^{phy} агента A_i в момент времени t (рис. 1).

Типичная математическая модель исследования операций представлена в следующей формулировке: «Максимизация или минимизация целевой функции при условии выполнения ограничений».

Задачей функционирования группы автономных агентов является выбор такого группового алгоритма действий, который будет максимизировать сумму собранной информации всеми агентами, используя ограниченный объем ресурсов, минимизируя при этом суммарные затраты ресурсов в момент времени t:

$$\max_{\pi} \sum_{t=0}^{T} \sum_{i=1}^{n} [\alpha \Delta Inf_i(t) - \beta \Delta R_i(t)],$$

где π — алгоритм управления; $\Delta Inf_i(t)$ — информация, собранная агентом A_i в момент времени t; $\Delta R_i(t)$ — ресурсы, затраченные агентом A_i в момент времени t; α, β — коэффициенты весов.

Метод коллективного анализа внешней среды на основе алгоритма поиска жуков-усачей

Исследование окружающей среды жуками-усачами может быть использовано при построении метода анализа внешней среды и достижения некоторых целей группы автономных агентов. Действительно, агенты должны иметь эффективные алгоритмы маршрутизации и планирования, чтобы выбирать оптимальный путь и исследовать внешнюю среду. Это включает распределение и координацию агентов, оптимизацию времени в пути.

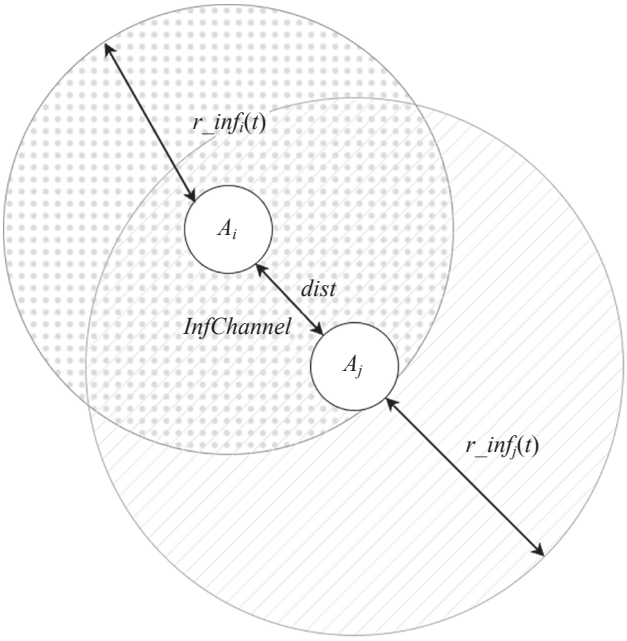

Рис. 1. Возможный информационный канал *InfChannel* между агентами A_i и A_j

Fig. 1. Possible information channel *InfChannel* between agents A_i and A_j

Важным аспектом оптимизации работы автономных агентов является связность и коммуникация. Агенты должны иметь надежные и быстрые средства связи для обмена информацией, обновления маршрутов, обнаружения препятствий и координации с другими агентами группы.

Строение антенн жуков-усачей, которые обычно состоят из множества типов обонятельных рецепторных клеток, часто уникально для конкретных видов, и функции таких сенсорных систем до сих пор являются предметом споров. Однако две основные функции таких больших антенн — улавливание запахов добычи и половых феромонов потенциально подходящего партнера. Также длинные антенны действуют как защитный механизм оповещения для жуков-усачей.

Благодаря своим большим размерам антенны могут увеличивать радиус обследования области окружающей среды вокруг себя. Такой механизм можно сравнить с механизмом, с помощью которого автономный агент исследует быстро меняющуюся среду с помощью датчиков и информационного взаимодействия внутри группы.

Жук вращает каждую антенну на одной стороне своего тела, чтобы обнаружить запах, когда он охотится или находит партнера, жук исследует окружающую область случайным образом, используя обе антенны. Когда антенны на одной стороне обнаруживают более высокую концентрацию запаха, жук поворачивается в эту сторону, в противном случае он поворачивается в другую сторону.

Для выполнения задачи коллективного анализа внешней среды агентами определим метод коллективного анализа внешней среды и построения матрицы-карты параметров группой агентов.

Метод состоит из 8 основных этапов, построенных на сборе информации агентами группы, оптимизации потребления ресурсов каждым агентом, а также повышении качества и объема собираемой информации о внешней среде.

Ниже подробно описан каждый этап метода, а также методы оптимизации на основе алгоритма поиска жуков-усачей.

Этап 0. Инициализация системы: определение скорости v и радиуса определения информации агентов $r_sens_i(t)$ для всех агентов группы Λ (рис. 2). При этом $r_sens_i(t) \geq v_i(t)\Delta t$, где Δt — принятая единица времени.

Этап 1. Сбор информации в точке $coord_i(t)$ на расстоянии $r_sens_i(t)$, построение матрицы $Inf_i(t)$.

Агент с помощью сенсоров и приборов восприятия окружающей среды A_i^{phy} собирает требуемые для анализа окружающей среды параметры.

Этап 2. Проведение информационного обмена A_i с каждого агента в радиусе $Inf_i(t)$, сопоставление и дополнение полученных множеств друг с другом:

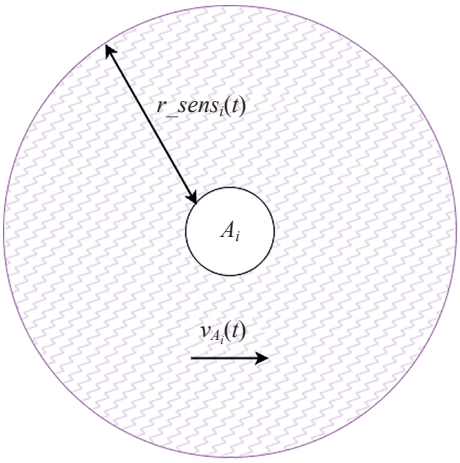

Рис. 2. Инициализация скорости и радиуса анализа области агентом A_i

Fig. 2. Initialization of the speed and radius of the agent's A_i area analysis

$Inf_i(t) = Inf_i(t-1) \cup Inf_{\Lambda_i}(t)$, где $\Lambda_i(t)$ — множество агентов в радиусе $r_inf_i(t)$ от агента A_i. На этапе 2 метода агенты, способные к взаимодействию друг с другом, на основе полученной информации дополняют карту внешней среды (рис. 3). Такой подход позволяет оптимизировать сбор данных о параметрах и объектах внешней среды.

Этап 3. Определим множество $D_{A_i} \backslash DN_{A_i} = \{d_1, \ldots, d_w\}$, где D_{A_i} — множество всех возможных направлений агента A_i; DN_{A_i} — множество направлений до ближайших агентов, с которыми произошел информационный обмен.

Полученное множество определяет список направлений, которые предотвращают повторный сбор агентом параметров, собранных другими агентами на этапе 2. Это позволяет снизить количество затрачиваемых группой агентов ресурсов.

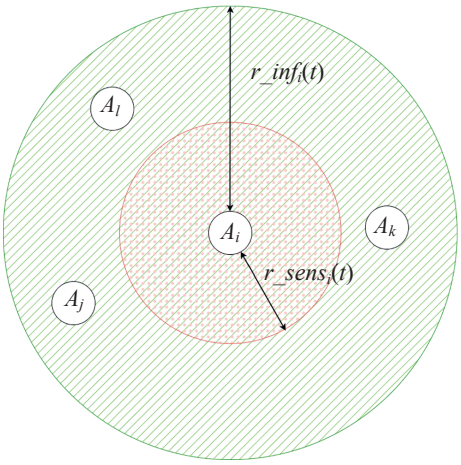

Рис. 3. Графическое отображение области информационного взаимодействия агента A_i с агентами A_j, A_l, A_k и области определения информации

Fig. 3. Graphical display of the area of information interaction of the agents A_i, A_j, A_l, A_k and the area of information definition

Этап 4. Определение вектора направления агента A_i на основе алгоритма поиска жуков-усачей:

$$coord_i(t+1) = coord_i(t) + \\ + dir(r_{sens_i}(t), coord_i(t), step(v), D_{A_i} \backslash DN_{A_i}, \\ \min\{M[Sp(coord_i(t+1))]\}),$$

где dir — функция, определяющая направление по некоторой координате $coord_i(t)$; $step(v)$ — шаг агента за единицу времени; Sp — функция минимизации, определяющая количества издержек на основе количества полученной информации $\Delta Inf_i(t+1)$ и затраченных на этот поиск ресурсов агента $\Delta R_i(t+1)$ при переходе из точки $coord_i(t)$ в точку $coord_i(t+1)$; M — математическое ожидание.

Тогда получим

$$\min\{M[Sp(coord_i(t+1))]\} = \\ = \frac{\min\{M[Sp(coord_i(t+1), \Delta R_i(t+1))]\}}{\max\{M[Sp(coord_i(t+1), \Delta Inf_i(t+1))]\}},$$

т. е. для минимизации издержек при выборе определенного направления требуется минимизации затраченных ресурсов $\Delta R_i(t+1)$ и максимизации обнаруженной новой информации об окружающей среде $\Delta Inf_i(t+1)$.

Этап 5. Информационный обмен A_i с агентами в радиусе $r_inf_i(t)$ о полученных на этапе 4 направлений векторов для момента времени $t+1$.

Этап 5 требуется для определения совпадения полученных агентами направлений, что может повысить издержки на ресурсы группы агентов, а также снизить объем полученной информации для следующего момента времени.

Этап 6. Если направления на этапе 4 совпали:

— корректировка направления при совпадении $coord_i(t+1)$, и $coord_j(t+1)$, на основе сравнения функций минимизации издержек для агента A_i и агента A_j соответственно:

$$dir(r_{sens_i}(t), coord_i(t), step(v), D_{A_i} \backslash DN_{A_i})$$

и

$$dir(r_{sens_j}(t), coord_j(t), step(v), D_{A_j} \backslash DN_{A_j});$$

— если направления не совпали, перейти к этапу 8.

Этап 7. Для агента чье значение функции издержек оказалось больше:

$$coord(t+1) = coord(t) + rnd(D_A \backslash DN_A),$$

где rnd — означает случайную функцию. Вернуться к этапу 5.

Этап 8. Переход агента в точку $coord(t+1)$, сбор информации ΔInf.

Далее этапы метода повторяются с этапа 2, пока групповая цель группы агентов не будет достигнута.

Эмпирическое исследование

Для оценки эффективности предлагаемого метода был разработан собственный программный симулятор, имитирующий перемещение группы агентов по опреде-

ленной местности. В качестве программной платформы для реализации разработанного метода использовался язык программирования Python 3. Python 3 позволяет работать с библиотекой NumPy, которая предоставляет эффективные структуры данных и операции для работы с многомерными массивами, что обеспечивает быстрые вычисления и обработку данных.

Все агенты $\Lambda = \{A_1, \ldots, A_n\}$ в симуляторе объединены в систему, где каждый элемент может взаимодействовать друг с другом для организации сбора информации.

Приведем описание области моделирования:

— область моделирования представлена в виде матрицы и разделена на квадратные элементарные сектора;
— агент может перемещаться только на одну ячейку за единицу времени: по вертикали, горизонтали и диагонали.

Для моделирования и сравнения результатов представленного метода было разработано программное обеспечение с использованием языка программирования Python 3.9, имитирующее движение агентов и обследование ими области с помощью алгоритмов MPC и CDMMT. Алгоритм MPC использует прогнозирование для оптимизации траектории движения. Алгоритм CDMMT основан на кооперативном принятии решений и использует централизованную координацию для распределения зон поиска между агентами. В эмпирическом исследовании оба алгоритма реализованы с возможностью обмена знаниями между агентами и учетом ограничений на ресурсы, что позволяет объективно сравнить их с предложенным методом на основе алгоритма BAS в условиях неполноты данных.

Входные данные моделирования следующие: область тестирования: 50 × 50 элементарных секторов; количество агентов — 10; скорость агентов постоянна и равна 1; пространственно моделирование ограничено; количество ресурсов у каждого агента в начальный момент времени одинаково и равно 100; радиус сбора информации — 5 элементарных секторов; радиус информационного взаимодействия — 10 элементарных секторов; каждый агент выполняет 20 шагов в каждом эксперименте; обмен информацией агентами происходит в радиусе — 10 элементарных секторов.

Для сравнения эффективности различных подходов анализа внешней среды группой агентов были использованы следующие метрики: суммарный объем собранной информации группой агентов за один эксперимент; суммарный объем затраченных ресурсов всеми агентами группы за один эксперимент.

В начальный момент времени агенты инициализируются в случайном месте, при этом агенты не могут появляться в одном и том же секторе. После заполнения тестовой области агентами начинается обследование этой области.

Проверка результатов:

— для оценки результатов были рассчитаны параметры объема полученной информации, собранной всей группой агентов, а также объем затраченных ресурсов;
— для каждого подхода было проведено 100 экспериментов — BAS, MPC, CDMMT;
— на каждой итерации эксперимента входные данные были одинаковыми для каждого подхода.

Основной целью моделирования была проверка возможности предложенного метода увеличить объем информации, минимизировав при этом объем затраченных ресурсов.

На рис. 4 представлены результаты эксперимента.

Как видно из рис. 4, разработанный метод позволяет повысить общую осведомленность группы агентов об анализируемой территории, а также снизить объем затраченных ресурсов. Стоит уточнить, что алгоритм CDMMT показал похожие результаты полученного объема информации, но из-за того, что агенты при обследо-

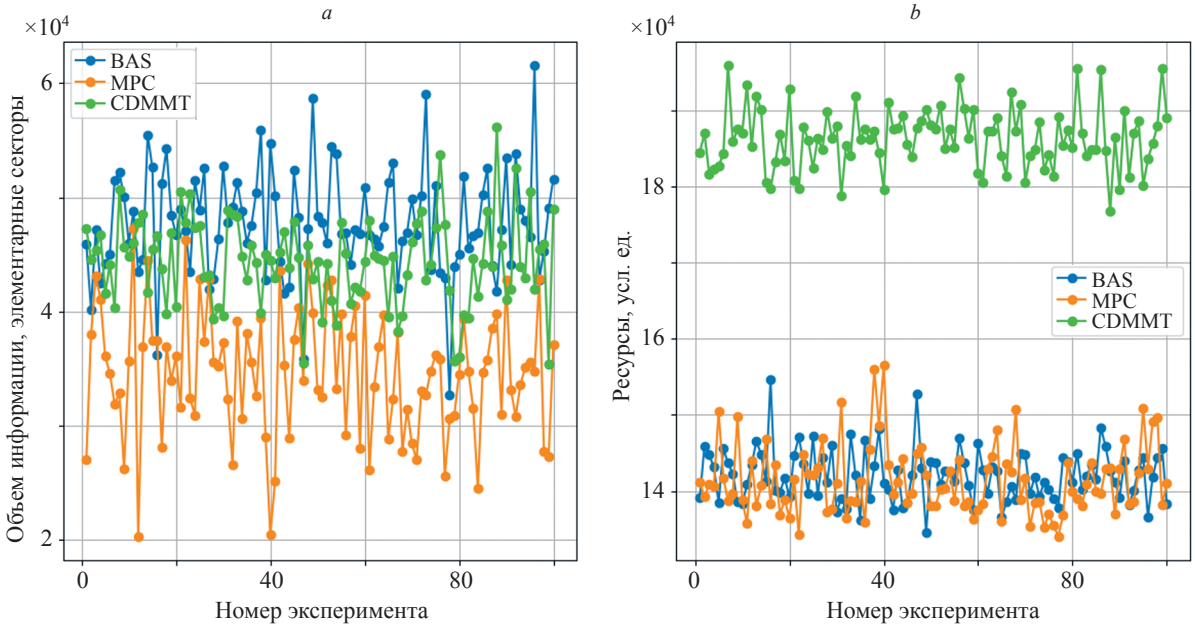

Рис. 4. Графики значений объемов полученной информации (*a*) и затраченных ресурсов (*b*) группой агентов

Fig. 4. Graphs of the values of the volumes of information received (*a*) and resources expended (*b*) by a group of agents

вании области с помощью CDMMT затрачивают большой объем ресурсов из-за работы в кооперации друг с другом. Это не соответствует постановке задачи группового управления автономных агентов, которая основывается на максимизации полученного объема информации, минимизируя при этом затраченные ресурсы.

Обсуждение

Оптимизация производительности группы автономных агентов является сложной задачей, требующей учета множества факторов, таких как эффективность маршрута, управление энергопотреблением, связность и безопасность. Разработка технологий и исследования в этой области направлены на решение этих проблем для обеспечения эффективного и безопасного использования автономных агентов и их интеграции в различные сферы жизни человека.

Эффективность и целесообразность метода были проверены с использованием как программного моделирования, результаты которого показывают возможность применения метода на практике. Результаты проведенных экспериментов демонстрируют, что предложенный метод анализа внешней среды на основе поведения жуков-усачей обеспечивает более высокую эффективность сбора информации по сравнению с Model Predictive Control и Cooperative Decision-Making for Mixed Traffic в условиях неполноты данных. Метод анализа внешней среды на основе поведения жуков-усачей достигает этого за счет самоорганизации, постоянного обмена знаниями между агентами и адаптивного поиска.

Заключение

В настоящей работе предлагается метод коллективного анализа внешней среды группой автономных агентов в условиях неполных данных. Модель группы автономных агентов построена на основе мультиагентного подхода. Метод коллективного анализа внешней среды разработан на основе алгоритма поиска жуков-усачей, основанного на поиске пищи жуками с помощью своих антенн. Представленный метод направлен на минимизацию затрат при достижении агентами групповых целей, т. е. минимизацию ресурсов, затрачиваемых на следующем этапе метода, и максимизацию получаемой информации. Кроме того, после каждого этапа и в начале инициализации системы агенты на определенном расстоянии от них обмениваются сообщениями для сравнения наборов информации.

Интеграция предлагаемого метода в процесс взаимодействия автономных агентов позволит повысить производительность такой системы, снизить уровень затрачиваемых ресурсов по сравнению со случайным выбором маршрута, а также автоматизировать работу системы в условиях неполных данных.

В качестве плана дальнейших исследований определено снижение количества допущений при моделировании и приближении условий функционирования системы к реальным, повышения нагрузки на систему за счет увеличения количества агентов.

В качестве еще одной задачи для изучения рассматривается возможность применения метода в группах беспилотных летательных аппаратов.

Литература

1. Wang J., Chen H. BSAS: Beetle swarm antennae search algorithm for optimization problems // arXiv. 2018. arXiv:1807.10470. https://doi.org/10.48550/arXiv.1807.10470
2. Laugier C., Chartre J. Intelligent perception and situation awareness for automated vehicles // Proc. of the Conference GTC Europe. 2016. hal-01428547.
3. Nègre A., Rummelhard L., Laugier C. Hybrid sampling bayesian occupancy filter // Proc. of the IEEE Intelligent Vehicles Symposium Proceedings. 2014. P. 1307–1312. https://doi.org/10.1109/ivs.2014.6856554
4. Rummelhard L., Nègre A., Laugier C. Conditional monte carlo dense occupancy tracker // Proc. of the IEEE 18th International Conference on Intelligent Transportation Systems. 2015. P. 2485–2490. https://doi.org/10.1109/itsc.2015.400
5. Finaev V.I., Ignatyev V.V., Shestova E.A., Spiridonov O.B., Zargaryan J.A., Zargaryan E.V. Optimum nominal method modification at the management of moving objects under uncertainty // ARPN Journal of Engineering and Applied Sciences. 2015. V. 10. N 16. P. 6837–6844.
6. Shapovalov I., Soloviev V., Finaev V., Beloglazov D., Zargaryan J., Kosenko E. Research of the controlled flight dynamics based on the full and simplified quadrotor models // Advances in Engineering Mechanics and Materials. 2014. P. 17–22.
7. Mamdani E.H. Application of fuzzy logic to approximate reasoning using linguistic synthesis // IEEE Transactions on Computers. 1977. V. C-26. N 12. P. 1182–1191. https://doi.org/10.1109/tc.1977.1674779
8. Nie J., Zhang J., Ding W., Wan X., Chen X., Ran B. Decentralized cooperative lane-changing decision-making for connected autonomous vehicles // IEEE Access. 2016. V. 4. P. 9413–9420. https://doi.org/10.1109/access.2017.2649567

References

1. Wang J., Chen H. BSAS: Beetle swarm antennae search algorithm for optimization problems. *arXiv*, 2018, arXiv:1807.10470. https://doi.org/10.48550/arXiv.1807.10470
2. Laugier C., Chartre J. Intelligent perception and situation awareness for automated vehicles. *Proc. of the Conference GTC Europe*, 2016, hal-01428547.
3. Nègre A., Rummelhard L., Laugier C. Hybrid sampling bayesian occupancy filter. *IEEE Intelligent Vehicles Symposium Proceedings*, 2014, pp. 1307–1312. https://doi.org/10.1109/ivs.2014.6856554
4. Rummelhard L., Nègre A., Laugier C. Conditional monte carlo dense occupancy tracker. *Proc. of the IEEE 18th International Conference on Intelligent Transportation Systems*, 2015, pp. 2485–2490. https://doi.org/10.1109/itsc.2015.400
5. Finaev V.I., Ignatyev V.V., Shestova E.A., Spiridonov O.B., Zargaryan J.A., Zargaryan E.V. Optimum nominal method modification at the management of moving objects under uncertainty. *ARPN Journal of Engineering and Applied Sciences*, 2015, vol. 10, no. 16, pp. 6837–6844.
6. Shapovalov I., Soloviev V., Finaev V., Beloglazov D., Zargaryan J., Kosenko E. Research of the controlled flight dynamics based on the full and simplified quadrotor models. *Advances in Engineering Mechanics and Materials*, 2014, pp. 17–22.
7. Mamdani E.H. Application of fuzzy logic to approximate reasoning using linguistic synthesis. *IEEE Transactions on Computers*, 1977, vol. C-26, no. 12, pp. 1182–1191. https://doi.org/10.1109/tc.1977.1674779
8. Nie J., Zhang J., Ding W., Wan X., Chen X., Ran B. Decentralized cooperative lane-changing decision-making for connected autonomous vehicles. *IEEE Access*, 2016, vol. 4, pp. 9413–9420. https://doi.org/10.1109/access.2017.2649567

9. Sun Z., Huang T., Zhang P. Cooperative decision-making for mixed traffic: A ramp merging example // Transportation Research Part C: Emerging Technologies. 2020. V. 120. P. 102764. https://doi.org/10.1016/j.trc.2020.102764
10. Cao W., Mukai M., Kawabe T., Nishira H., Fujiki N. Mild merging path generation method with optimal merging point based on MPC // IFAC Proceedings Volumes. 2013. V. 46. N 21. P. 756–761. https://doi.org/10.3182/20130904-4-jp-2042.00109
11. Ghiasi A., Hussain O., Qian Z.S., Li X. A mixed traffic capacity analysis and lane management model for connected automated vehicles: A Markov chain method // Transportation Research Part B: Methodological. 2017. V. 106. P. 266–292. https://doi.org/10.1016/j.trb.2017.09.022
12. Gipps P.G. A behavioural car-following model for computer simulation // Transportation Research Part B: Methodological. 1981. V. 15. N 2. P. 105–111. https://doi.org/10.1016/0191-2615(81)90037-0

9. Sun Z., Huang T., Zhang P. Cooperative decision-making for mixed traffic: A ramp merging example. *Transportation Research Part C: Emerging Technologies*, 2020, vol. 120, pp. 102764. https://doi.org/10.1016/j.trc.2020.102764
10. Cao W., Mukai M., Kawabe T., Nishira H., Fujiki N. Mild merging path generation method with optimal merging point based on MPC. *IFAC Proceedings Volumes*, 2013, vol. 46, no. 21, pp. 756–761. https://doi.org/10.3182/20130904-4-jp-2042.00109
11. Ghiasi A., Hussain O., Qian Z.S., Li X. A mixed traffic capacity analysis and lane management model for connected automated vehicles: A Markov chain method. *Transportation Research Part B: Methodological*, 2017, vol. 106, pp. 266–292. https://doi.org/10.1016/j.trb.2017.09.022
12. Gipps P.G. A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, 1981, vol. 15, no. 2, pp. 105–111. https://doi.org/10.1016/0191-2615(81)90037-0

Авторы

Павелина Юлия Александровна — аспирант, ведущий специалист, АО «НИИАС», Санкт-Петербург, 196006, Российская Федерация; инженер-исследователь, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57216353188, https://orcid.org/0000-0001-7396-2831, lyakhovenko.kam@gmail.com

Попов Илья Юрьевич — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 57202195632, https://orcid.org/0000-0002-6407-7934, ilyapopov27@gmail.com

Authors

Julia A. Pavelina — PhD Student, Leading Specialist, JSC NIIAS, Saint Petersburg, 196006, Russian Federation; Engineer-Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57216353188, https://orcid.org/0000-0001-7396-2831, lyakhovenko.kam@gmail.com

Ilya Yu. Popov — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 57202195632, https://orcid.org/0000-0002-6407-7934, ilyapopov27@gmail.com

Статья поступила в редакцию 03.04.2025
Одобрена после рецензирования 12.10.2025
Принята к печати 23.11.2025

Received 03.04.2025
Approved after reviewing 12.10.2025
Accepted 23.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

MODELING AND SIMULATION

doi: 10.17586/2226-1494-2025-25-6-1168-1176
УДК 66.011:519.688:004.89

Поиск оптимального температурного профиля химического процесса на основе генетического алгоритма

Эльдар Наилевич Мифтахов[1]✉, Дмитрий Владимирович Иванов[2]

[1,2] Уфимский университет науки и технологий, Уфа, 450076, Российская Федерация

[1] promif@mail.ru✉, https://orcid.org/0000-0002-0471-5949
[2] ivanov_dv@list.ru, https://orcid.org/0009-0005-7117-0191

Аннотация
Введение. Рассмотрена проблема поиска оптимального температурного профиля сложного физико-химического процесса. Наибольшие проблемы возникают при исследовании и оптимизации многокомпонентных систем. Это определяет научный и практический интерес к разработке наиболее эффективных инструментов поиска оптимальных режимов организации производства. Одним из ключевых аспектов является учет динамических ограничений, влияющих на скорость изменения управляющих параметров и обеспечивающих построение физически реализуемых траекторий изменения температуры. Для решения этой задачи предложен модифицированный генетический алгоритм, позволяющий учитывать заданные ограничения. **Метод.** Для сложного физико-химического процесса сформулирована задача оптимизации, которая заключается в поиске оптимального температурного профиля, способствующего максимизации (или минимизации) заданного целевого параметра в условиях задаваемых ограничений на скорость изменения температуры. Метод основан на дискретном разбиении всего времени протекания процесса и представлении температурного профиля в виде кусочно-линейной функции, значения которой на каждом из интервалов определяются с использованием генетического алгоритма оптимизации. Основные этапы выполнения генетического алгоритма были модифицированы и представлены в виде адаптированной схемы эволюционного поиска, учитывающей допустимые границы изменения управляющих параметров. Внесенные модификации позволили повысить устойчивость алгоритма к локальным экстремумам и обеспечить более точное соблюдение задаваемых ограничений. **Основные результаты.** Эффективность алгоритма, функциональность программного модуля и механизм взаимодействия были апробированы путем организации вычислительного эксперимента по исследованию кинетики реакции диметилкарбоната со спиртами в присутствии дикобальтоктакарбонила. Результаты численного моделирования продемонстрировали, что температурный режим оказывает значительное влияние на кинетику процесса, а проведенные вычислительные испытания позволили однозначно идентифицировать оптимальный температурный профиль, поиск которого проводился в условиях ограничений на рост температуры и дополнительного требования линейного изменения концентрации целевого продукта. **Обсуждение.** Предложенная модификация генетического алгоритма позволила значительно повысить его устойчивость к локальным экстремумам и обеспечить более точное соблюдение технологических ограничений. В частности, анализ полученных профилей показал, что представленный метод позволяет находить решения, обеспечивающие более равномерное распределение концентраций целевого продукта, что особенно важно при проектировании реакционных систем с высокой чувствительностью к изменению параметров. Разработанный метод оптимизации может быть полезен при проектировании и масштабировании химико-технологических процессов, а проведенное исследование подтверждает эффективность использования численных методов и эволюционных алгоритмов для оптимизации условий химических реакций.

Ключевые слова
эвристические методы, генетический алгоритм, оптимальный температурный профиль, реакция диметилкарбоната со спиртами

Благодарности
Исследование выполнено за счет гранта Российского научного фонда № 24-21-00380, https://rscf.ru/project/24-21-00380/.

© Мифтахов Э.Н., Иванов Д.В., 2025

Ссылка для цитирования: Мифтахов Э.Н., Иванов Д.В. Поиск оптимального температурного профиля химического процесса на основе генетического алгоритма // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1168–1176. doi: 10.17586/2226-1494-2025-25-6-1168-1176

Optimization of the temperature profile in a chemical process using a genetic algorithm

Eldar N. Miftakhov[1]✉, Dmitry V. Ivanov[2]

[1,2] Ufa University of Science and Technology, Ufa, 450076, Russian Federation

[1] promif@mail.ru✉, https://orcid.org/0000-0002-0471-5949
[2] ivanov_dv@list.ru, https://orcid.org/0009-0005-7117-0191

Abstract
This study addresses the problem of finding an optimal temperature profile for a complex physico-chemical process. The greatest difficulties arise in the study and optimization of multicomponent systems, which determines both scientific and practical interest in developing the most effective tools for identifying optimal production regimes. One of the key aspects is the consideration of dynamic constraints that affect the rate of change of control parameters and ensure the construction of physically feasible trajectories of temperature variation. To solve this problem, a modified genetic algorithm is proposed, allowing for the incorporation of predefined constraints. An optimization problem is formulated for a complex physico-chemical process, aiming to determine the optimal temperature profile that maximizes (or minimizes) a given target parameter while satisfying constraints on the rate of temperature change. The method is based on discretizing the total process duration and representing the temperature profile as a piecewise linear function, with segment values determined using a genetic optimization algorithm. The main stages of the genetic algorithm have been modified and presented as an adaptive evolutionary search scheme that accounts for permissible control parameter variations. These modifications enhance the algorithm robustness against local extrema and ensure more precise adherence to predefined constraints. The efficiency of the algorithm, the functionality of the software module, and the interaction mechanism were tested through a computational experiment investigating the kinetics of the reaction of dimethyl carbonate with alcohols in the presence of dicobalt octacarbonyl. Numerical simulations demonstrated that the temperature regime significantly influences reaction kinetics, and computational trials enabled the unambiguous identification of the optimal temperature profile under constraints on temperature increase and an additional requirement for the linear variation of the target product concentration. The proposed modification of the genetic algorithm significantly improved its robustness against local extrema and ensured stricter compliance with technological constraints. In particular, an analysis of the obtained profiles showed that the proposed method allows for solutions that ensure a more uniform distribution of the target product concentration, which is especially important in the design of reaction systems highly sensitive to parameter variations. This optimization approach can be useful for the design and scaling of chemical-technological processes, and the conducted study confirms the effectiveness of numerical methods and evolutionary algorithms for optimizing chemical reaction conditions.

Keywords
heuristic methods, genetic algorithm, optimal temperature profile, dimethyl carbonate reaction with alcohols

Acknowledgements
The study was supported by the grant from the Russian Science Foundation No. 24-21-00380, https://rscf.ru/en/project/24-21-00380/.

For citation: Miftakhov E.N., Ivanov D.V. Optimization of the temperature profile in a chemical process using a genetic algorithm. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1168–1176 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1168-1176

Введение

Оптимизация параметров химических реакций является важнейшим этапом в повышении эффективности производственных процессов и снижении их затрат. Одним из ключевых факторов, определяющих протекание процесса, является температурный профиль, который оказывает прямое влияние как на скорость элементарных реакций, так и на выход целевых и побочных продуктов. Оптимизация температурного профиля позволяет снизить энергозатраты, повысить селективность и минимизировать образование нежелательных примесей, что особенно важно для таких высокотехнологичных отраслей, как, например, фармацевтическая и нефтехимическая промышленность [1, 2].

Наибольшие сложности в решении задач управления температурным режимом возникают при исследовании и оптимизации многокомпонентных систем, реакции в которых протекают в несколько стадий. Классические подходы, в основе которых лежат многократные эмпирические испытания и экспериментальный подбор, требуют значительных затрат времени и ресурсов и не могут гарантировать нахождение оптимального решения. Все это определяет огромный научный и практический интерес к разработке наиболее эффективных инструментов поиска оптимальных режимов организации производства.

В основе классического метода моделирования сложных химических процессов и систем лежит кинетический подход [3, 4], в рамках которого набор химических реакций описывается системой дифференциальных уравнений, отражающих изменение материального баланса по каждому компоненту реакции. В дальнейшем полученная система расширяется уравнениями

теплопередачи и дополнительными балансовыми соотношениями. В условиях использования кинетического подхода к моделированию для определения оптимального режима производства предлагается разбиение всего времени протекания процесса на равные промежутки и представление температурного профиля в виде кусочно-линейной функции. Задача оптимизации в этом случае сводится к идентификации вектора, определяющего значения температуры в каждой точке дискретизации с использованием соответствующих методов и алгоритмов оптимизации. При этом открытым остается вопрос выбора соответствующих численных методов, способных решать задачи идентификации достаточно большого числа неизвестных. В частности, классические методы (градиентного спуска, Лагранжа [5] и т. д.) позволяют достигать высокой точности в случае исследования систем с достаточно простой структурой и небольшим количеством переменных. Кроме того, классические методы требуют знания аналитического представления целевой функции и недостаточно эффективны для задач глобальной оптимизации.

В последнее время широкое распространение благодаря универсальности и способности решать задачи глобальной оптимизации получили эвристические методы [6], принципиально отличающиеся от классических. Эти методы не предъявляют строгих требований к целевой функции, способны эффективно работать с нелинейными и многомерными данными, допускают возможность адаптации для решения разных типов задач. Однако, эвристические методы также не лишены недостатков, поскольку могут требовать тонкой настройки своих параметров и значительных вычислительных ресурсов. В работе [7] предложено решение задачи определения оптимального начального состава реакционной смеси методом, в основе которого лежит генетический алгоритм [8]. Полученные результаты подтвердили высокую эффективность алгоритма и его потенциал для решения задач глобальной оптимизации.

Кроме того, в доминирующей части научных исследований при решении подобных задач часто ограничиваются теоретической оптимизацией, когда найденные режимы характеризуются резкими изменениями температуры. В таких случаях не учитываются тепловая инерция системы, ограничения на скорость нагрева и охлаждения, а также процессы теплообмена, которые неизбежно сопровождают любой химический процесс. Пренебрежение физическими ограничениями приводит к нереализуемым решениям, которые не могут быть воспроизведены в экспериментальных или промышленных условиях. Реальная динамика изменения температуры определяется уравнениями теплопередачи и теплоемкостью реакционной среды, что требует более комплексного подхода к оптимизации. В настоящей работе предлагается рассмотреть влияние динамических ограничений на температурный профиль с целью построения физически реализуемых траекторий.

Цель работы — разработка алгоритма поиска оптимального температурного режима сложного физико-химического процесса с учетом задаваемых динамических ограничений на скорость изменения температуры.

Постановка задачи

Сформулируем задачу идентификации оптимального температурного профиля в общем виде. Пусть химическая реакция на промежутке времени $[t_{start}, t_{end}]$ описывается системой обыкновенных дифференциальных уравнений с начальными условиями $\mathbf{X}(t_{start}) = \mathbf{X}^0$:

$$\frac{d\mathbf{X}}{dt} = \mathbf{F}(t, \mathbf{T}, \mathbf{X}(t, \mathbf{T})),$$

где $\mathbf{X}(t, \mathbf{T}) = (x_1(t, \mathbf{T}), \ldots, x_n(t, \mathbf{T}))^T$ — вектор концентраций веществ; t — время; $\mathbf{T} = (T_1, \ldots, T_m)$ — вектор, определяющий температуру реакционной смеси в дискретные моменты времени; $\mathbf{F}$ — непрерывная вектор-функция; $\mathbf{X}^0 = (x_1^0, x_2^0, \ldots, x_n^0)^T$ — вектор начальных концентраций; m — количество точек разбиения исходного отрезка $[t_{start}, t_{end}]$.

В качестве критерия оптимальности примем функционал общего вида:

$$H(\mathbf{T})(t_{end}) = H(x_1(t_{end}, \mathbf{T}), \ldots, x_n(t_{end}, \mathbf{T})) \to extr, \quad (1)$$

который может выражать выход целевого или побочного продукта, конверсию, селективность и т. д.

С целью построения физически реализуемых траекторий введем ограничение на изменение температуры в виде:

$$|T_{i+1} - T_i| \le \Delta T_{\max}, \; i = 1, \ldots, m-1, \quad (2)$$

где T_i — температура реакционной смеси в момент времени t_i; $\Delta T_{\max}$ — заданное максимально допустимое изменение температуры за один шаг. В силу технологических ограничений температура реакционной смеси может меняться в диапазоне

$$T_{\min} \le T_i \le T_{\max}, \; i = 1, \ldots, m. \quad (3)$$

Требуется найти оптимальный температурный профиль $\mathbf{T}^*$ размерностью m, доставляющий оптимум функционалу (1) при выполнении ограничений (2) и (3).

В работах [9, 10] были разработаны вариации эвристических методов, успешно примененные для решения задач оптимизации различного типа. Однако в условиях заданных ограничений на искомый температурный профиль требуется модифицировать ряд шагов алгоритма, чтобы гарантированно найти глобальное решение задачи.

Описание генетического алгоритма с учетом ограничений на скорость изменения температуры

В основе предлагаемого подхода к решению задачи оптимизации лежит генетический алгоритм, основное наполнение которого представляет собой итерационное выполнение четырех основных операторов — селекции, скрещивания, мутации и создания нового поколения [11]. В условиях задаваемых ограничений на скорость изменения температуры модифицируем операторы генетического алгоритма и представим его основные этапы.

Этап 1. Инициализация популяции. Задается начальная популяция из K возможных решений, представляющая собой матрицу размерностью $K \times m$, каждая строка которой представляет собой один из возможных температурных профилей. С учетом имеющихся ограничений вида (2) и (3) рекомендуется вектор начальной популяции определить случайным образом на заданном отрезке:

$$T_{i,1} = \mathrm{rnd}(T_{\min}, T_{\max}),$$
$$T_{i,j} = T_{i,j-1} + \mathrm{rnd}(-\Delta T_{\max}, \Delta T_{\max}), j = 2, \ldots, m,$$

для всех $i = 1, \ldots, K$ при условии, что $T_{\min} \le T_{i,j} \le T_{\max}$. Функция $\mathrm{rnd}(a, b)$ — возвращает случайное число из диапазона $[a, b]$.

На этапе 1 алгоритма также задается параметр $iter = 1$, определяющий порядковый номер формируемого поколения, максимальное число итераций $iMax$ и точность решения задачи $\varepsilon > 0$, значение которой используется для остановки алгоритма.

Этап 2. Оценка. Проводится оценка степени приспособленности найденных «особей» путем расчета значения целевой функции для каждого предложенного на этапе 1 решения. В случае минимизации целевой функции проверяется условие $H(\mathbf{T}_i^{iter}) \le \varepsilon$, и если оно выполняется хотя бы для одного $i = 1, \ldots, K$, то алгоритм следует остановить и вывести решение в качестве оптимального. Алгоритм следует также остановить в случае достижения предельного числа итераций $iter = iMax$. Если целью решения задачи является достижение максимального значения, то в качестве условия окончания итерационного поиска принимается критически малое изменение значения функции приспособленности на протяжении нескольких итераций.

Этап 3. Селекция. На основании найденных значений $H(\mathbf{T}_i^{iter})$, определяющих степень приспособленности каждой «особи», проводится процедура селекции — отбора потенциальных родителей, для выполнения которой рассчитываются соответствующие вероятности. Если хотя бы один из показателей $H(\mathbf{T}_i^{iter})$ превышает значение 1, то требуется предварительно нормализовать полученные оценки. В случае минимизации целевой функции для оценки вероятностей используется выражение:

$$p(\mathbf{T}_i^{iter}) = \frac{1 - H(\mathbf{T}_i^{iter})}{\sum_{i=1}^{K}(1 - H(\mathbf{T}_i^{iter}))}.$$

Затем рассчитанные значения располагаются последовательно на отрезке от 0 до 1, образуя полную группу. Далее, с помощью процедуры генерации случайных чисел, равномерно распределенных на этом отрезке, и их попадания в соответствующие интервалы, выбираются особи для формирования родительского пула.

Этап 4. Скрещивание. Из выбранного родительского пула с помощью оператора кроссинговера с некоторой вероятностью $p_x > 0$ формируется новое поколение. Выбор оператора играет ключевую роль в обеспечении достаточной вариативности, что влияет на скорость сходимости алгоритма и качество найденных решений. Классический арифметический кроссинговер [12, 13] основан на взвешенном усреднении значений родительских особей, обеспечивая постепенную адаптацию и способствуя плавному изменению параметров. С целью повышения генетического разнообразия в настоящей работе предлагается использовать двухточечный кроссинговер [14], в основе выполнения которого лежит механизм разбиения родительских хромосом в двух случайных точках с последующим обменом центральных сегментов между ними. Это приводит к более резким и разнообразным изменениям в популяции, что особенно важно при поиске сложных оптимальных решений.

В частности, если на некоторой итерации случайным образом были выбраны две родительские особи $\mathbf{T}^1 = (T_1^1, T_2^1, \ldots, T_m^1)$ и $\mathbf{T}^2 = (T_1^2, T_2^2, \ldots, T_m^2)$, то в случае использования двухточечного кроссинговера генерируются две случайные точки разреза $1 \le q_1, q_2 \le m$, на основании которых значения вектора потомка образуются по правилу

$$\tilde{T}_i = \begin{cases} T_i^1, \text{ если } 1 \le i < q_1 \text{ или } q_2 \le i \le m, \\ T_i^2, \text{ если } q_1 \le i < q_2. \end{cases}$$

Поскольку данный тип оператора может привести к резким изменениям температуры, то проводится дополнительный контроль, позволяющий оценить разницу между соседними значениями и если условие $|T_{i+1} - T_i| \le \Delta T_{\max}$ не выполняется хотя бы для одного из $1 \le i < m$, то требуется принудительно изменить значение T_{i+1}, ограничив разницу величиной $\Delta T_{\max}$.

Этап 5. Мутация. Процедура мутации проводится с некоторой вероятностью $p_m > 0$ и играет ключевую роль в поддержании генетического разнообразия, позволяющего избежать сходимости популяции к локальному экстремуму. В связи с необходимостью плавных и контролируемых изменений температурного профиля требуется, чтобы мутация происходила в пределах разумных отклонений, избегая резких скачков, которые могли бы нарушить физическую реалистичность модели. В связи с этим на этапе 5 предлагается использовать оператор мутации по Гауссу [15, 16], согласно которому каждое значение T_i (ген) изменяется по формуле

$$\tilde{T}_i = T_i + N(\mu, \sigma).$$

Функция $N(\mu, \sigma)$ возвращает случайное число из нормального (Гауссова) распределения, где μ — среднее значение; σ — отклонение.

Поскольку ограничение на изменение температуры известно и составляет $\Delta T_{\max}$, то контролируемые изменения температурного профиля достигаются при значениях параметров $\mu = 0$ и $\sigma = \Delta T_{\max}$. По окончанию этапа 5 также проводится дополнительная проверка условия $|T_{i+1} - T_i| \le \Delta T_{\max}$, $1 \le i < m$, после которой увеличивается счетчик итераций $iter = iter + 1$ и осуществляется переход к этапу 2 с целью оценки полученных решений.

Вычислительный эксперимент

Продемонстрируем работу генетического алгоритма для решения задачи нахождения оптимального температурного режима на примере реакции диметилкарбоната со спиртами [17] в присутствии дикобальтоктакарбонила. Данная реакция играет ключевую роль в развитии «зеленой» химии, поддерживая промышленный переход к более экологичным и устойчивым технологиям [18].

Брутто-реакция диметилкарбоната $(MeO)_2CO$ со спиртами ROH в присутствии дикобальтоктакарбонила $Co_2(CO)_8$ имеет вид

$$ROH + (MeO)_2CO \xrightarrow{Co_2(CO)_8} ROCO_2Me + ROMe + CO_2 + MeOH.$$

Каталитический цикл в схеме химических превращений реакции включает следующие стадии [19]:

$$Co_2(CO)_8 + ROH \xrightarrow{k_1} CO^{2+}(ROH) + Co(CO)_4^- + 4CO,$$

$$Co(CO)_4^- + (MeO)_2CO \xrightarrow{k_2} Me^+[Co(CO)_4^-] + CO_2 + MeO^-,$$

$$Co(CO)_4^- + (MeO)_2CO \underset{k_8}{\overset{k_3}{\rightleftharpoons}} Co(CO)_4CO_2Me + MeO^-,$$

$$Co(CO)_4CO_2Me + ROH \xrightarrow{k_4} HCo(CO)_4 + CO_2 + ROMe,$$

$$Co(CO)_4CO_2Me + ROH \xrightarrow{k_5} ROCO_2Me + HCo(CO)_4,$$

$$Me^+[Co(CO)_4^-] + ROH \xrightarrow{k_6} ROMe + HCo(CO)_4,$$

$$HCo(CO)_4 + MeO^- \xrightarrow{k_7} MeOH + Co(CO)_4,$$

где $k_1, \ldots, k_8$ — константы, характеризующие скорости элементарных стадий.

В рамках кинетического подхода к моделированию данная реакция описывается системой дифференциальных уравнений

$$\begin{cases} \dfrac{dy_1}{dt} = -k_1y_3y_1 - k_4y_1y_{12} - k_5y_1y_{12} - k_6y_1y_{11}, \\ \dfrac{dy_2}{dt} = -k_2y_2y_{10} - k_3y_2y_{10} + k_8y_9y_{12}, \\ \dfrac{dy_3}{dt} = -k_1y_3y_1;\ \dfrac{dy_4}{dt} = k_6y_1y_{11} + k_4y_1y_{12}, \\ \dfrac{dy_5}{dt} = k_5y_1y_{12};\ \dfrac{dy_6}{dt} = k_2y_2y_{10} + k_4y_1y_{12}, \\ \dfrac{dy_7}{dt} = k_7y_8y_9, \\ \dfrac{dy_8}{dt} = k_4y_1y_{12} + k_5y_1y_{12} + k_6y_1y_{11} - k_7y_8y_9, \\ \dfrac{dy_9}{dt} = k_2y_2y_{10} + k_3y_2y_{10} - k_8y_9y_{12} - k_7y_8y_9, \\ \dfrac{dy_{10}}{dt} = k_1y_3y_1 - k_2y_2y_{10} - k_3y_2y_{10} + k_8y_9y_{12} + k_7y_8y_9, \\ \dfrac{dy_{11}}{dt} = k_2y_2y_{10} - k_6y_1y_{11}, \\ \dfrac{dy_{12}}{dt} = k_3y_2y_{10} - k_8y_9y_{12} - k_4y_1y_{12} - k_5y_1y_{12}, \\ \dfrac{dy_{13}}{dt} = 4k_1y_3y_1;\ \dfrac{dy_{14}}{dt} = k_1y_3y_1. \end{cases} \quad (4)$$

Система уравнений (4) имеет следующие начальные условия:

$$y_1(0) = y_1^0,\ y_2(0) = y_2^0,\ y_3(0) = y_3^0,\ y_i(0) = 0,\ i = 4, \ldots, 14,$$

где y_i — концентрации (моль/л) соответствующих веществ: y_1 — ROH, y_2 — $(MeO)_2CO$, y_3 — $Co_2(CO)_8$ — исходные реагенты; y_4 — ROMe, y_5 — $ROCO_2Me$, y_6 — CO_2, y_7 — MeOH — продукты реакции; y_8 — $HCo(CO)_4$, y_9 — MeO, y_{10} — $Co(CO)_4$, y_{11} — $Me^+[Co(CO)_4^-]$, y_{12} — $Co(CO)_4CO_2Me$, y_{13} — CO, y_{14} — $Co^{2+}(ROH)$ — промежуточные реагенты; k_j — константы, характеризующие скорости элементарных стадий, которые зависят от температуры, согласно уравнению Аррениуса:

$$k_j = k_j^0 \exp\left(-\frac{E_j}{RT}\right),$$

где k_j^0 — предэкспоненциальные множители; E_j — энергия активации; R — универсальная газовая постоянная (8,314 Дж/(моль·К)); T — абсолютная температура. Значения кинетических параметров [20] реакции представлены в таблице.

Таблица. Значения кинетических параметров

Table. Values of kinetic parameters

Параметры	Номер элементарной стадии j							
	1	2	3	4	5	6	7	8
$\ln k_j^0$	22,0	11,9	20,7	18,2	12,0	20,1	15,9	10,9
E_j, ккал/моль	24,1	8,1	13,4	15,0	5,6	21,5	10,6	9,0

Для обеспечения выполнения закона сохранения вещества в любой момент времени необходимо учитывать следующие балансовые соотношения [20]:

$$\begin{gathered} y_9 + y_7 + y_2 = y_2^0,\ y_{14} + y_3 = y_3^0,\ y_{13} + 4y_3 = 4y_3^0, \\ y_1 - y_3 + y_8 + y_7 = y_1^0 - y_3^0,\ y_4 - y_8 + y_5 - y_7 = 0, \\ y_6 + y_2 + y_{12} + y_5 = y_2^0,\ y_{10} + y_3 - y_2 - y_7 = y_3^0 - y_2^0, \\ y_{11} + y_2 + y_{12} + y_8 + y_7 = y_2^0. \end{gathered} \quad (5)$$

Использование соотношений (5) позволяет уменьшить количество независимых переменных и сократить систему (4) до 6 уравнений, сохранив при этом полноту описания процесса.

Требуется найти вектор $\mathbf{T}^* = (T_1, \ldots, T_m)$, определяющий температуру реакционной смеси в дискретные моменты времени и способствующий достижению максимальной концентрации продукта y_5 в реакционной массе в конечный момент моделирования $t_{end} = 300$ мин при начальных условиях $y_1(0) = 5{,}882$ моль/л, $y_2(0) = 17{,}647$ моль/л, $y_3(0) = 0{,}059$ моль/л

$$H(\mathbf{T}^*)(t_{end}) = y_5(t_{end}) \to \max. \quad (6)$$

Для решения данной задачи генетический алгоритм, представленный в разделе «Описание генетического алгоритма с учетом ограничений на скорость изменения температуры», был запрограммирован с использованием языка Python (версия 3.10.12). Вычислительные испытания проводились в облачной среде с использованием следующих параметров генетического алгоритма: размер начальной популяции — 30, допустимая точность решения задачи $\varepsilon = 10^{-3}$, предельное число итераций $iMax = 10^3$, вероятность скрещивания $p_x = 0{,}7$, вероятность мутации $p_m = 0{,}2$. Температурный профиль определялся $m = 10$ точками на промежутке 0–300 мин и максимально допустимым изменением температуры $\Delta T_{\max} = 5$ °C. Также учитывалось, что на температуру наложены ограничения 150 °C $\le T_i \le$ 200 °C, $i = 1, \ldots, m$.

Многократно проведенные вычислительные испытания показали, что алгоритм достаточно быстро сходится и позволяет достичь максимального значения $y_5(t_{end}) = 5{,}519$ моль/л, однако выделить единственный температурный профиль не удается, поскольку задача обладает целой совокупностью решений, способствующих выполнению критерия (6). Один из найденных температурных профилей и соответствующее ему изменение концентрации продукта y_5 представлены на рис. 1 и 2.

Из полученных зависимостей на рис. 2 видно, что для данного конкретного температурного профиля максимальное значение целевого продукта достигается уже после 75 мин протекания процесса. Полученные результаты позволяют сформировать дополнительное требование линейности, согласно которому максимальный выход целевого продукта y_5 достигается при условии его линейного роста во времени. Данное требование обусловлено необходимостью минимизации побочных реакций и обеспечения стабильности процесса. Линейное изменение концентрации целевого продукта свидетельствует о равномерности расхода исходных реагентов и

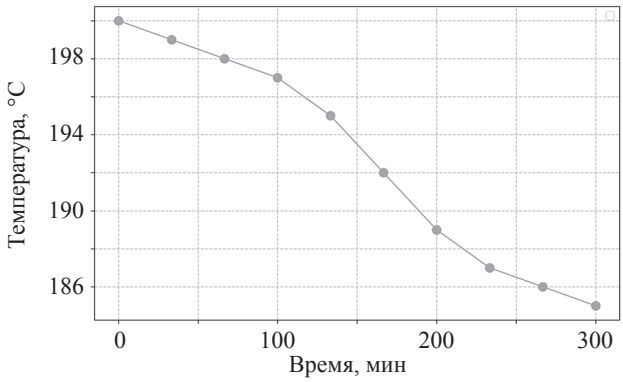

Рис. 1. Найденный температурный профиль при температурных ограничениях — 150–200 °C, $\Delta T_{\max} = 5$ °C; параметры генетического алгоритма: размер популяции — 30, $\varepsilon = 10^{-3}$, $iMax = 10^3$, $p_x = 0{,}7$, $p_m = 0{,}2$

Fig. 1. Calculated temperature profile under temperature constraints of 150–200 °C, $\Delta T_{\max} = 5$ °C; genetic algorithm parameters: population size — 30, $\varepsilon = 10^{-3}$, $iMax = 10^3$, $p_x = 0{,}7$, $p_m = 0{,}2$

устойчивости кинетики, что важно для промышленной реализации процесса. Кроме того, равномерный рост y_5 позволяет обеспечить предсказуемость выхода продукта, что упрощает разработку системы автоматического управления температурным режимом.

Нахождение оптимального температурного профиля, способствующего максимальному равномерному выходу целевого продукта, требует изменения вида целевой функции. Для этого на основании найденного максимального значения $y_5^{\max}$ рассчитываются дискретные точки вектора $y_5^{line}(t_i)$, определяющего линейный тренд. Для того чтобы оценить отклонение реальных значений от целевого тренда рассчитывается величина среднеквадратичной ошибки Mean Squared Error (MSE), определяющая среднее отклонение в квадратичной форме

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (y_5(t_i) - y_5^{line}(t_i))^2.$$

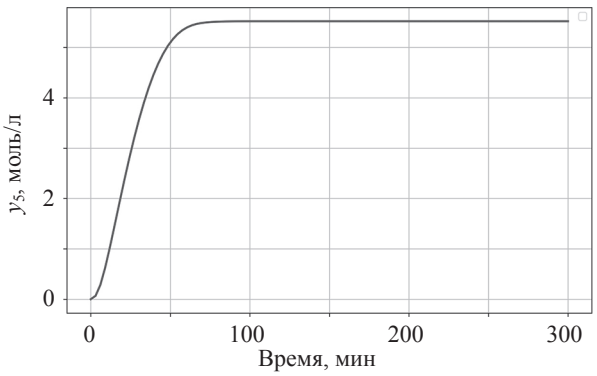

Рис. 2. Динамика концентрации $ROCO_2Me$ при $[ROH]_0 = 5{,}882$ моль/л, $[(MeO)_2CO]_0 = 17{,}647$ моль/л, $[Co_2(CO)_8]_0 = 0{,}059$ моль/л; кинетические параметры (таблица)

Fig. 2. Dynamics of $ROCO_2Me$ concentration at $[ROH]_0 = 5.882$ mol/l, $[(MeO)_2CO]_0 = 17.647$ mol/l, $[Co_2(CO)_8]_0 = 0{,}059$ mol/l; kinetic parameters (Table)

В таком случае целевая функция будет состоять из двух частей — отклонения конечного значения от максимального и MSE между реальным и линейным трендами:

$$H(\mathbf{T}^*)(t_{end}) = (y_5(t_{end}) - y_5^{\max})^2 + MSE \to \min. \quad (7)$$

Параметры выполнения генетического алгоритма и ограничения на температуру для обновленной целевой функции (7) не менялись. Многочисленные вычислительные испытания позволили однозначно идентифицировать температурный профиль **T*** =[165,1, 168,1, 164,3, 159,3, 162,7, 167,7, 172,6, 177,6, 182,6, 187,6], способствующий достижению максимального значения целевого продукта y_5 = 5,38 моль/л в конечный момент моделирования t_{end} = 300 мин в условиях линейного роста концентрации целевого продукта. Алгоритм остановил свою работу при достижении предельного числа итераций 10^3, при этом минимальное значение целевой функции (7) составило 0,037, а *MSE* = 0,02. Запуск алгоритма проводился многократно, однако не приводил к изменению внешнего вида и поведения профиля, визуализация которого представлена на рис. 3, *a*.

График изменения мольной концентрации продукта y_5 (моль/л) во времени при найденном температурном профиле представлен на рис. 3, *b* в виде сплошной линии. Пунктирная линия на рисунке демонстрирует линейный тренд, позволяющий оценить отклонение реальной динамики от идеализированного линейного роста. Полученный рисунок позволяет проанализировать влияние температуры на процесс и эффективность выбранного температурного режима.

Все стадии каталитического цикла протекают во всем исследуемом температурном диапазоне (150–200 °C). Температурный профиль в модели является внешним управляющим параметром, определяющим относительные скорости стадий через температурную зависимость констант скоростей. При этом в области относительно низких температур преобладает образование комплекса $HCo(CO)_4$, тогда как при более высоких температурах возрастает вклад стадий, ведущих к формированию $Co(CO)_4CO_2Me$. Таким образом, оптимальные температурные профили не исключают отдельные стадии механизма, а лишь перераспределяют их относительный вклад в общий ход процесса.

С целью оценки корректности получаемых расчетных результатов задача также решалась с использованием классических подходов к оптимизации, в частности методом Нелдера–Мида, не требующего аналитического представления производной. Для соблюдения задаваемых ограничений на рост температуры целевая функция была модифицирована путем введения штрафных коэффициентов, обеспечивающих направление поиска оптимального решения в область допустимых значений. Проведенные вычислительные испытания постоянно приводили к отличающимся результатам, при этом значение MSE кратно превышало ранее найденное значение, что позволило убедиться в эффективности разработанного генетического алгоритма.

Очевидно, что минимизировать ошибку, определяемую целевой функцией вида (7), можно как в случае снятия ограничений на изменение температуры, так и изменения требований к росту целевого компонента y_5. Требование линейности в данном случае введено для оценки эффективности разработанного генетического алгоритма, позволяющего решать задачи поиска оптимального температурного профиля, обеспечивающего равномерный рост целевого компонента. Подобные требования способствуют нахождению таких режимов нагрева, при которых реакция протекает наиболее эффективно, избегая резких скачков скорости.

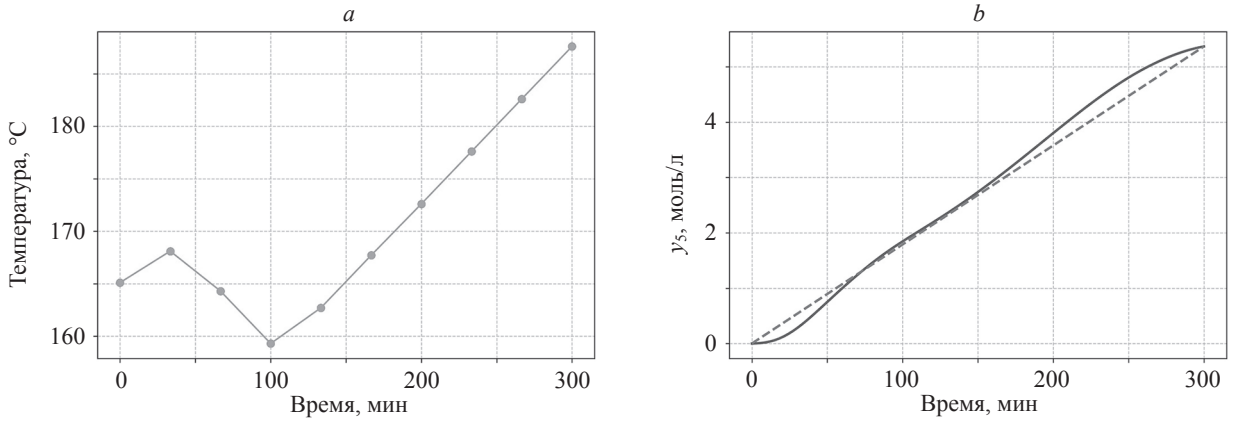

Рис. 3. Оптимальный температурный профиль при условии линейного роста концентрации $ROCO_2Me$ (*a*) и соответствующая динамика (*b*) с наложенным линейным трендом (пунктир) при тех же начальных условиях, температурных ограничениях и параметрах алгоритма (рис. 1–2); *MSE* = 0,02

Fig. 3. Optimal temperature profile under the condition of linear growth of $ROCO_2Me$ concentration (*a*) and the corresponding dynamics (*b*) with a superimposed linear trend (dashed line), obtained under the same initial conditions, temperature constraints, and algorithm parameters as in Figs. 1 and 2; *MSE* = 0.02

Заключение

Разработан алгоритм поиска оптимального температурного режима сложного физико-химического процесса с целью управления динамикой концентрации целевого вещества. Оптимизация температурного профиля играет ключевую роль в повышении эффективности химико-технологических процессов, влияя на скорость реакции, селективность и выход целевого продукта.

С целью построения физически реализуемых траекторий изменения температурного режима рассмотрено влияние динамических ограничений на изменение управляющих параметров. В основе предлагаемого подхода к решению задачи оптимизации лежит генетический алгоритм, который способствует нахождению глобального оптимума благодаря случайному характеру при поиске решений и возможности поддерживать их разнообразие в популяции. В условиях задаваемых ограничений на скорость изменения температуры основные этапы выполнения генетического алгоритма были модифицированы и представлены в виде адаптированной схемы эволюционного поиска, учитывающей допустимые границы изменения управляющих параметров. Внесенные модификации позволили повысить устойчивость алгоритма к локальным экстремумам и обеспечить более точное соблюдение технологических ограничений.

Эффективность алгоритма, функциональность программного модуля и механизм взаимодействия были проверены на примере вычислительного эксперимента по исследованию кинетики реакции диметилкарбоната со спиртами в присутствии дикобальтоктакарбонила. Результаты численного моделирования продемонстрировали, что температурный режим оказывает значительное влияние на кинетику процесса. В ходе проведенных вычислительных испытаний дополнительно было введено требование линейности концентрации целевого продукта, обусловленное необходимостью равномерного расхода исходных реагентов и устойчивости кинетики, что важно для промышленной реализации процесса. Результаты многократно проведенных вычислений позволили однозначно идентифицировать оптимальный температурный профиль, удовлетворяющий заданным условиям.

Предложенный метод оптимизации может быть полезен при проектировании и масштабировании химико-технологических процессов, а проведенное исследование подтверждает эффективность использования численных методов и эволюционных алгоритмов для оптимизации условий химических реакций.

Литература

1. Marpaung T.J. Optimization of temperature on the concentration of a mixture of substances in liquid waste treatment process palm oil // Journal of Physics: Conference Series. 2019. V. 1235. N 1. P. 012123. https://doi.org/10.1088/1742-6596/1235/1/012123
2. Шевчук В.П., Шаровина С.О. Управление температурным профилем ректификационной колонны тарельчатого типа // Приборы и системы. Управление, контроль, диагностика. 2013. № 3. С. 39–47.
3. Слинько М.Г. Кинетическая модель как основа математического моделирования каталитических процессов // Теоретические основы химической технологии. 1976. Т. 10. № 1. С. 137–146.
4. Быков В.И., Цыбенова С.Б. Нелинейные модели химической кинетики. М.: Ком-Книга, 2010. 350 с.
5. Васильев Ф.П. Численные методы решения экстремальных задач. М.: Наука, 1988. 552 с.
6. Пантелеев А.В., Скавинская Д.В., Алёшина Е.А. Метаэвристические алгоритмы поиска оптимального программного управления. М.: Инфра-М, 2020. 396 с.
7. Мифтахов Э.Н., Кашникова А.П., Иванов Д.В. Использование генетических алгоритмов для решения задачи поиска оптимального состава реакционной смеси // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24. № 4. С. 637–644. https://doi.org/10.17586/2226-1494-2024-24-4-637-644
8. Katoch S., Chauhan S.S., Kumar V. A review on genetic algorithm: past, present, and future // Multimedia Tools and Applications. 2021. V. 80. N 5. P. 8091–8126. https://doi.org/10.1007/s11042-020-10139-6
9. Мифтахов Э.Н., Иванов Д.В. Поиск оптимального состава реакционной смеси с использованием алгоритма имитации отжига // Автометрия. 2024. Т. 60. № 5. С. 120–128. https://doi.org/10.15372/AUT20240514
10. Мифтахов Э.Н. Алгоритм поиска оптимального режима подачи регулятора в процессе получения полимерных продуктов // Инженерные технологии и системы. 2024. Т. 34. № 4. С. 597–614. https://doi.org/10.15507/2658-4123.034.202404.597-614
11. Awad A., Hawash A., Abdalhaq B. A Genetic Algorithm (GA) and swarm-based Binary Decision Diagram (BDD) reordering optimizer reinforced with recent operators // IEEE Transactions on Evolutionary Computation. 2023. V. 27. N 3. P. 535–549. https://doi.org/10.1109/tevc.2022.3170212

References

1. Marpaung T.J. Optimization of temperature on the concentration of a mixture of substances in liquid waste treatment process palm oil. *Journal of Physics: Conference Series*, 2019, vol. 1235, no. 1, pp. 012123. https://doi.org/10.1088/1742-6596/1235/1/012123
2. Shevchuk V.P., Sharovina S.O. The control of the temperature profile of the rectification column dish type. *Instruments and Systems: Monitoring, Control, and Diagnostics*, 2013, no. 3, pp. 39–47. (in Russian)
3. Slinko M.G. Kinetic model as a basis for mathematical modeling of catalytic processes. *Teoreticheskie Osnovy Khimicheskoi Tekhnologii*, 1976, vol. 10, no. 1, pp. 137–146. (in Russian)
4. Bykov V.I., Tcybenova S.B. *Nonlinear Models of Chemical Kinetics*. Moscow, Kom-Kniga Publ., 2010, 350 p. (in Russian)
5. Vasilev F.P. *Numerical Methods for Extreme Problems Solving*. Moscow: Nauka Publ., 1988, 552 p. (in Russian)
6. Panteleev A.V., Skavinskaia D.V., Aleshina E.A. *Metaheuristic Algorithms for Searching of the Optimal Program Control*. Moscow: Infra-M Publ., 2020, 396 p. (in Russian)
7. Miftakhov E.N., Kashnikova A.P., Ivanov D.V. Using genetic algorithms to solve the problem of finding the optimal composition of the reaction mixture. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 4, pp. 637–644. (in Russian). https://doi.org/10.17586/2226-1494-2024-24-4-637-644
8. Katoch S., Chauhan S.S., Kumar V. A review on genetic algorithm: past, present, and future. *Multimedia Tools and Applications*, 2021, vol. 80, no. 5, pp. 8091–8126. https://doi.org/10.1007/s11042-020-10139-6
9. Miftakhov E.N., Ivanov D.V. Searching for the optimal composition of the reaction mixture using a simulated annealing algorithm. *Avtometriya*, 2024, vol. 60, no. 5, pp. 120–128. (in Russian). https://doi.org/10.15372/AUT20240514
10. Miftakhov E.N. Algorithm for searching the optimal regulator supply mode in the process of manufacturing polymer products. *Engineering Technologies and Systems*, 2024, vol. 34, no. 4, pp. 597–614. (in Russian). https://doi.org/10.15507/2658-4123.034.202404.597-614
11. Awad A., Hawash A., Abdalhaq B. A Genetic Algorithm (GA) and swarm-based Binary Decision Diagram (BDD) reordering optimizer reinforced with recent operators. *IEEE Transactions on Evolutionary Computation*, 2023, vol. 27, no. 3, pp. 535–549. https://doi.org/10.1109/tevc.2022.3170212

12. Yang D., Rao K., Xu B., Sheng W. PIR sensors deployment with the accessible priority in smart home using genetic algorithm // International Journal of Distributed Sensor Networks. 2015. V. 11. N 11. P. 146270. https://doi.org/10.1155/2015/146270
13. Антипина Е.В., Мустафина С.А., Антипин А.Ф. Автоматизация поиска оптимальных значений параметров процесса олигомеризации этилена // Научно-технический вестник информационных технологий, механики и оптики. 2024. Т. 24. № 4. С. 563–570. https://doi.org/10.17586/2226-1494-2024-24-4-563-570
14. Zhou T., Fang W. Two-point crossover operator in genetic algorithm for deep learning compiler // Proc. of the Companion Conference on Genetic and Evolutionary Computation. 2023. P. 347–350. https://doi.org/10.1145/3583133.3590536
15. Bell O. Applications of Gaussian Mutation for self adaptation in evolutionary genetic algorithms // arXiv. 2022. 10.48550/arXiv.2201.00285. https://doi.org/10.48550/arXiv.2201.00285
16. Tinós R., Yang S. Evolution strategies with q-Gaussian mutation for dynamic optimization problems // Proc. of the 11th Brazilian Symposium on Neural Networks. 2010. P. 223–228. https://doi.org/10.1109/sbrn.2010.46
17. Koledina K.F., Koledin S.N., Schadneva N.A., Mayakova Y.Yu., Gubaydullin I.M. Kinetic model of the catalytic reaction of dimethylcarbonate with alcohols in the presence $Co_2(CO)_8$ and $W(CO)_6$ // Reaction Kinetics, Mechanisms and Catalysis. 2017. V. 121. N 2. P. 425–438. https://doi.org/10.1007/s11144-017-1181-3
18. Fiorani G., Perosa A., Selva M. Dimethyl carbonate: a versatile reagent for a sustainable valorization of renewables // Green Chemistry. 2018. V. 20. N 2. P. 288–322. https://doi.org/10.1039/c7gc02118f
19. Коледина К.Ф., Коледин С.Н., Щаднева Н.А., Губайдуллин И.М. Кинетика и механизм каталитической реакции спиртов с диметилкарбонатом // Журнал физической химии. 2017. Т. 91. № 3. С. 422–428. https://doi.org/10.7868/S004445371703013X
20. Коледина К.Ф. Многокритериальная оптимизация и оптимальное управление химическими процессами на основе детализированной кинетической модели: диссертация на соискание ученой степени доктора физико-математических наук. Уфа: Уфимский федеральный исследовательский центр Российской академии наук, 2020. 334 с.

12. Yang D., Rao K., Xu B., Sheng W. PIR sensors deployment with the accessible priority in smart home using genetic algorithm. *International Journal of Distributed Sensor Networks*, 2015, vol. 11, no. 11, pp. 146270. https://doi.org/10.1155/2015/146270
13. Antipina E.V., Mustafina S.A., Antipin A.F. Automation of search for optimal values of the ethylene oligomerization process parameters. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2024, vol. 24, no. 4, pp. 563–570. (in Russian). https://doi.org/10.17586/2226-1494-2024-24-4-563-570
14. Zhou T., Fang W. Two-point crossover operator in genetic algorithm for deep learning compiler. *Proc. of the Companion Conference on Genetic and Evolutionary Computation*, 2023, pp. 347–350. https://doi.org/10.1145/3583133.3590536
15. Bell O. Applications of Gaussian Mutation for self adaptation in evolutionary genetic algorithms. *arXiv*, 2022, 10.48550/arXiv.2201.00285. https://doi.org/10.48550/arXiv.2201.00285
16. Tinós R., Yang S. Evolution strategies with q-Gaussian mutation for dynamic optimization problems. *Proc. of the 11th Brazilian Symposium on Neural Networks*, 2010, pp. 223–228. https://doi.org/10.1109/sbrn.2010.46
17. Koledina K.F., Koledin S.N., Schadneva N.A., Mayakova Y.Yu., Gubaydullin I.M. Kinetic model of the catalytic reaction of dimethylcarbonate with alcohols in the presence $Co_2(CO)_8$ and $W(CO)_6$. *Reaction Kinetics, Mechanisms and Catalysis*, 2017, vol. 121, no. 2, pp. 425–438. https://doi.org/10.1007/s11144-017-1181-3
18. Fiorani G., Perosa A., Selva M. Dimethyl carbonate: a versatile reagent for a sustainable valorization of renewables. *Green Chemistry*, 2018, vol. 20, no. 2, pp. 288–322. https://doi.org/10.1039/c7gc02118f
19. Koledina K.F., Koledin S.N., Shchadneva N.A., Gubaidullin I.M. Kinetics and mechanism of the catalytic reaction between alcohols and dimethyl carbonate. *Russian Journal of Physical Chemistry A*, 2017, vol. 91, no. 3, pp. 442–447. https://doi.org/10.1134/s003602441703013x
20. Koledina K.F. *Multi-Criteria Optimization and Optimal Control of Chemical Processes Based on a Detailed Kinetic Model*. Ufa, UFRC RAS, 2020, 334 p. (in Russian)

Авторы

Мифтахов Эльдар Наилевич — доктор физико-математических наук, старший научный сотрудник, Уфимский университет науки и технологий, Уфа, 450076, Российская Федерация, sc 56178153800, https://orcid.org/0000-0002-0471-5949, promif@mail.ru

Иванов Дмитрий Владимирович — кандидат физико-математических наук, доцент, Уфимский университет науки и технологий, Уфа, 450076, Российская Федерация, sc 57197070892, https://orcid.org/0009-0005-7117-0191, ivanov_dv@list.ru

Authors

Eldar N. Miftakhov — D.Sc. (Physics & Mathematics), Senior Researcher, Ufa University of Science and Technology, Ufa, 450076, Russian Federation, sc 56178153800, https://orcid.org/0000-0002-0471-5949, promif@mail.ru

Dmitry V. Ivanov — PhD (Physics & Mathematics), Associate Professor, Ufa University of Science and Technology, Ufa, 450076, Russian Federation, sc 57197070892, https://orcid.org/0009-0005-7117-0191, ivanov_dv@list.ru

Статья поступила в редакцию 04.04.2025
Одобрена после рецензирования 29.09.2025
Принята к печати 25.11.2025

Received 04.04.2025
Approved after reviewing 29.09.2025
Accepted 25.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1177-1184
УДК 531.746, 550.832

Использование магнитного градиентометра в скважинном инклинометре с целью компенсации внешних помех

Олег Валерьевич Гасанов[1], Виталий Николаевич Ситников[2], Вадим Рафисович Хаматдинов[3], Илья Викторович Гринев[4]✉, Александр Борисович Королев[5]

[1,2,3,4,5] ООО «Нефтегазгеофизика», Тверь, 170033, Российская Федерация

[1] gasanov.o@karotazh.ru, https://orcid.org/0009-0004-6275-688X
[2] v.sitnikov@karotazh.ru, https://orcid.org/0009-0009-0226-3297
[3] hvadim@karotazh.ru, https://orcid.org/0009-0006-3257-0317
[4] i.grinev@karotazh.ru✉, https://orcid.org/0009-0003-9078-4792
[5] korolev@karotazh.ru, https://orcid.org/0009-0002-1290-574X

Аннотация
Введение. В процессе измерений магнитного азимута оси ствола скважины расчеты проводятся на основе компонент суперпозиции магнитного поля Земли и паразитных полей остаточной намагниченности сборки геофизических приборов и бурильной колонны. В условиях высоких широт горизонтальная составляющая геомагнитного поля критически мала. Это приводит к тому, что даже слабые по величине, около 1 % от величины геомагнитного поля, паразитные поля приводят к погрешностям измерения магнитного азимута оси скважины в 4° и более. В известных исследованиях представлено большое количество методов борьбы с данным влиянием. Однако, почти все они требуют либо дополнительного оборудования и предварительных измерений, либо знания точных величин модуля и наклонения геомагнитного поля для места измерений. В результате возникает проблема создания метода компенсации, который бы не требовал предварительных замеров параметров паразитного поля или модуля и наклонения геомагнитного поля. **Метод.** Предложен метод использования в инклинометре дополнительного магнитометра для измерения градиента суперпозиционного магнитного поля. На основе имитационного моделирования по величине полученного градиента подбирается эквивалентный по влиянию источник магнитного поля в виде кругового витка с током. Далее из показаний опорного магнитометра вычитается влияние от подобранного витка. В лабораторных экспериментах в качестве источников паразитных полей использовались кольцевые неодимовые магниты (три варианта источников с разной магнитной индукцией), располагающиеся на оси инклинометра. В качестве магнитного градиентометра применены два магнитометрических датчика с расстоянием между магнитометрами 0,307 м. **Основные результаты.** В экспериментах, при помощи разработанного алгоритма, были подобраны параметры витков с током, эквивалентных источникам по магнитному влиянию. Это позволило компенсировать показания опорного магнитнометра и повысить точность измерений магнитного азимута с: –1°15′36″ (источник 1), –3°9′36″ (источник 2) и 12°30′36″ (источник 3) до ±0°39′ для случая использования всех источников. В эксперименте величины магнитной индукции источников в месте расположения опорного магнитометра были равны: 0,42 %, 1,59 % и 5,60 % от величины геомагнитного поля соответственно. **Обсуждение.** Предложенный метод позволяет повысить точность измерений азимута без замера на скважине параметров паразитного и геомагнитного полей. Кроме того, использование метода дает возможность уменьшить длины охранного кожуха и немагнитных вставок по обе стороны от инклинометра в случае бурения. Таким образом, метод может быть реализован в составе датчика, осуществляющего вычисление и компенсацию паразитных полей непосредственно во время каротажа или бурения.

Ключевые слова
инклинометр, геонавигация, каротаж, навигация во время бурения, геофизическая сборка, бурильная колонна, поле остаточной намагниченности, увеличение точности расчета азимута

Ссылка для цитирования: Гасанов О.В., Ситников В.Н., Хаматдинов В.Р., Гринев И.В., Королев А.Б. Использование магнитного градиентометра в скважинном инклинометре с целью компенсации внешних помех // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1177–1184. doi: 10.17586/2226-1494-2025-25-6-1177-1184

© Гасанов О.В., Ситников В.Н., Хаматдинов В.Р., Гринев И.В., Королев А.Б., 2025

Using a magnetic gradiometer in a borehole inclinometer to compensate for external magnetic interference

Oleg V. Gasanov[1], Vitaly N. Sitnikov[2], Vadim R. Khamatdinov[3], Ilya V. Grinyov[4]✉, Alexander B. Korolyov[5]

[1,2,3,4,5] LTD Neftegazgeofizika, Tver, 170033, Russian Federation

[1] gasanov.o@karotazh.ru, https://orcid.org/0009-0004-6275-688X
[2] v.sitnikov@karotazh.ru, https://orcid.org/0009-0009-0226-3297
[3] hvadim@karotazh.ru, https://orcid.org/0009-0006-3257-0317
[4] i.grinev@karotazh.ru✉, https://orcid.org/0009-0003-9078-4792
[5] korolev@karotazh.ru, https://orcid.org/0009-0002-1290-574X

Abstract

In measurements of the magnetic azimuth of the borehole axis, calculations are based on the superposition of the Earth's magnetic field and parasitic fields from the remanent magnetization of the geophysical tool assembly and the drill string. At high latitudes the horizontal component of the geomagnetic field is very small. As a result, even weak parasitic fields — on the order of 1 % of the geomagnetic field — can cause azimuth errors of 4° or more. Many methods to mitigate this effect have been reported in the literature. However, almost all of them require either additional equipment and preliminary measurements, or knowledge of the exact values of the magnitude and inclination of the geomagnetic field at the survey location. In connection with all of the above, there is a problem of creating a compensation method that would not require preliminary measurements of the parameters of the parasitic field or the modulus and inclination of the geomagnetic field. This paper proposes using an additional magnetometer in the inclinometer to measure the gradient of the superpositional magnetic field. From simulation, and using the measured gradient, an equivalent magnetic source in the form of a circular current loop is determined. The calculated field of this loop is then subtracted from the reference magnetometer readings. In the laboratory experiments, ring neodymium magnets (three variants with different magnetic flux densities) placed on the inclinometer axis were used as parasitic-field sources. A magnetic gradiometer was formed by two magnetometer sensors spaced 0.307 m apart. In experiments, the developed algorithm identified parameters of current loops equivalent to the sources in terms of magnetic effect. This enabled compensation of the reference magnetometer readings and improved azimuth accuracy from −1°15′36″ (source 1), −3°9′36″ (source 2) and +12°30′36″ (source 3) to ±0°39′ for all sources. In the experiment the field magnitudes at the reference magnetometer location were 0.42 %, 1.59 % and 5.60 % of the geomagnetic field, respectively. The proposed method increases azimuth measurement accuracy without requiring measurements of parasitic or geomagnetic field parameters. In addition, the use of the method allows reducing the length of nonmagnetic collars on both sides of the inclinometer during drilling. Thus, the method can be implemented in a sensor that computes and compensates for parasitic fields in real time during logging or drilling.

Keywords

inclinometer, geosteering, logging, measurement-while-drilling (MWD) navigation, geophysical tool assembly, drill string, remanent magnetization, azimuth accuracy improvement

For citation: Gasanov O.V., Sitnikov V.N., Khamatdinov V.R., Grinyov I.V., Korolyov A.B. Using a magnetic gradiometer in a borehole inclinometer to compensate for external magnetic interference. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1177–1184 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1177-1184

Введение

В задачах скважинной навигации, опирающейся на гравитационное и геомагнитное (ГМП) поля, широко распространена проблема, связанная с влиянием на измеряемый магнитный азимут паразитных полей остаточной намагниченности (ПОН) сборки геофизических приборов и бурильных труб, которые используются в комплексе со скважинным инклинометром. Данное влияние особенно выражено в наклонно-направленных и горизонтальных скважинах.

Приведенная проблема, в силу своей актуальности, решается большим количеством аппаратных, программных и комбинированных методов. В случае навигации во время бурения инклинометр разносится в пространстве с источниками ПОН при помощи немагнитных труб необходимой длины [1]. В ряде случаев может использоваться размагничивание бурильного оборудования до [2] или непосредственно в процессе [3] проведения работ в скважине. Применяется также метод компенсации помех посредством съема девиационных кривых влияния на азимут ПОН компоновки низа бурильной колонны [4, 5], а также через измерение, моделирование и компенсацию поля компоновки низа бурильной колонны [6]. Измерение осевой составляющей суммарного ПОН компоновки низа бурильной колонны и бурильной колонны до и после наращивания сборки предлагается в работе [7].

В большинстве методов компенсации ПОН, представленных в научных работах, учитывается только осевая составляющая ПОН. В частности, подобный подход, на основе известных для места проведения работ модуля и угла наклонения ГМП, используется в широко известном «методе коротких утяжеленных бурильных труб» [8–10]. Однако, как известно, глобальные магнитные модели, применяемые в данном методе, могут давать существенные ошибки в определении среднесуточных локальных значений параметров ГМП [11]. Более того, суточные вариации, в том числе модуля и угла наклонения ГМП, существенно возрастают с увеличением широты места исследований [12] и могут доходить до сотен нанотесла и десятых градуса

даже в спокойные геомагнитные дни в условиях севера Сибири и Арктического шельфа.

Учет только осевой составляющей ПОН является существенным недостатком указанных методов, так как, по результатам практических наблюдений, вектор ПОН в месте расположения магнитометров часто не направлен по оси сборки, а в некоторых случаях угол расхождения может доходить до десятков градусов [13].

В ряде работ предложено минимизировать влияние ПОН путем комбинирования магнитометрических и гироскопических датчиков [14–16]. Данное направление является перспективным, однако, в силу высокой стоимости гироскопов, существенно увеличивает стоимость инклинометрической аппаратуры.

Также решением проблемы является исследование скважины отдельным инклинометром, но и в данном случае происходит удорожание геофизических работ в связи с увеличением объема спуско-подъемных операций. В случае измерений на бурильных трубах данный подход не только не работает, но может приводить к еще большим влияниям, так как инклинометр приближается к источнику ПОН.

Перспективным методом решения проблемы представляется апостериорная обработка данных инклинометрии [17]. Недостатком же данного метода является то, что скорректированные данные становятся доступными только после сбора необходимого количества данных, а полный анализ возможен только после окончания скважинных работ.

В то же время в геофизике широко используется двухзондовый метод измерений, когда электрическое, электромагнитное или радиационное поля измеряются на двух расстояниях от источника. При фиксированном расстоянии между точками измерений и известных параметрах источника, по градиенту поля вычисляются параметры окружающих прибор пород.

В том случае, если два магнитометра применяются в инклинометре, по градиенту магнитного поля необходимо вычислить не параметры слагающих стенки скважины пород, а параметры самих источников суперпозиционного поля. Важно заметить, что данный подход осуществим только в том случае, если стенки скважины слагают породы со слабо выраженными магнитными свойствами. Кроме того, как в случае транзитного скважинного инклинометра, так и в случае забойного датчика, источников ПОН — два: от нижней и верхней частей сборки.

В работах [18, 19] предложено использование двух магнитометров с измерением градиента и последующей компенсацией магнитного азимута. Однако, как и в случае «метода коротких утяжеленных бурильных труб» [8–10], корректируется только осевая компонента поля, измеряемого опорным магнитометром.

В настоящей работе представлен метод, который не требует знания модуля и угла наклонения ГМП в месте измерений. Кроме того, проводится расчет не только осевой, но и радиальной составляющей ПОН. Подход базируется на том предположении, что в качестве источника ПОН можно рассматривать одиночный круговой виток с током.

Метод

Постановка задачи. Располагающийся в сборке геофизических приборов инклинометр измеряет не ГМП, а суперпозицию ГМП и ПОН. Это приводит к существенным погрешностям измерения инклинометром магнитного азимута оси скважины. Отметим, что ПОН можно с необходимой для практических целей точностью аппроксимировать магнитным влиянием одиночного кольцевого витка с током (рис. 1). Определение ориентации и силы тока в таком эквивалентном витке — задача, которая может быть решена на основе показаний двух трехосевых магнитометров (M_1 и M_2 (рис. 1 и рис. 2, *a*, *c*)), располагающихся на известном расстоянии друг от друга на оси прибора. Ближайший к нижнему концу инклинометра магнитометр (M_2) является опорным, так как его показания корректируются на основе вычисленных параметров источника ПОН и используются при расчете магнитного азимута.

Таким образом, цель моделирования — повышение точности расчета магнитного азимута инклинометра.

Для достижения указанной цели необходимо согласование показаний магнитометров — минимизация разности показаний при разных ориентациях прибора без внешних магнитных воздействий (отдельная серьезная задача, которая не рассматривается в настоящей работе). Далее требуется вычисление параметров витка с током, который имитирует влияние ПОН, по измеренной разности показаний магнитометров. В результате проводится корректировка показаний опорного магнитометра и перерасчет на их основе магнитного азимута (погрешность не должна превышать ±1°30′ при комнатной температуре).

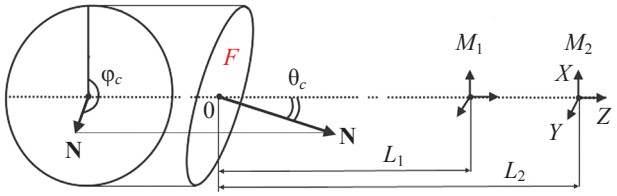

Рис. 1. Эквивалентный круговой виток F и его расположение относительно магнитометров M_1 и M_2 прибора.

Центр витка располагается в точке 0 (центре системы координат). Черными стрелками показан вектор нормали **N** к плоскости витка, в левой части рисунка — осевая проекция витка, в правой части рисунка — направления осей в системе координат.

θ_c — угол между вектором нормали **N** и осью Z системы координат; φ_c — угол между проекцией вектора нормали **N** на плоскость $X0Y$ и осью X; L_1 и L_2 — расстояния от центра витка до магнитометров

Fig. 1. Equivalent circular current loop F and its position relative to the device magnetometers M_1 and M_2.

The black arrow indicates the loop normal vector **N**. The left part of the figure shows the loop axial projection.

In the picture: θ_c — angle between the normal vector **N** and the Z-axis of the coordinate system; φ_c — angle between the projection of the normal vector **N** onto the $X0Y$ plane and the X-axis; L_1 and L_2 — distances from the center of the coil to the magnetometers

Проекции вектора магнитной индукции произвольно ориентированного кругового витка с током на оси координат в произвольной точке. Индукцию магнитного поля в вакууме проводника с постоянным током в точке с радиус-вектором $\mathbf{r}_0 = \{r_{x0}; r_{y0}; r_{z0}\}$ можно вычислить по закону Био–Савара–Лапласа

$$\mathbf{B}(\mathbf{r}_0) = \frac{\mu_0}{4\pi}\int_\gamma \frac{I[\mathbf{dr}(\mathbf{r}_0 - \mathbf{r})]}{|\mathbf{r}_0 - \mathbf{r}|^3}, \quad (1)$$

где μ_0 — магнитная постоянная; $\mathbf{r}$ — вектор, соответствующий точке интегрирования на контуре γ с током I; $\mathbf{dr}$ — вектор элемента контура.

Рассмотрим в качестве источника паразитного ПОН (1) круговой виток с током (рис. 1). Примем центр витка за начало координат. Зададим ориентацию плоскости витка в пространстве через вектор нормали $\mathbf{N}$ к его плоскости.

В таком случае вектор, соответствующий точке интегрирования на контуре витка, можно задать как

$$\mathbf{r} = \mathbf{r}(\varphi) = \{(R\cos\varphi; R\sin\varphi; 0)\mathbf{\Omega}\},$$

где R — радиус витка; φ — угол положения точки контура относительно оси X в плоскости $X0Y$; $\mathbf{\Omega} = \mathbf{\Omega}_Y(\theta_c)\mathbf{\Omega}_Z(\varphi_c)$ — матрица поворотов вокруг оси Y на угол θ_c:

$$\mathbf{\Omega}_Y(\theta_c) = \begin{pmatrix} \cos\theta_c & 0 & \sin\theta_c \\ 0 & 1 & 0 \\ -\sin\theta_c & 0 & \cos\theta_c \end{pmatrix};$$

вокруг оси Z на угол φ_c:

$$\mathbf{\Omega}_Z(\varphi_c) = \begin{pmatrix} \cos\varphi_c & -\sin\varphi_c & 0 \\ \sin\varphi_c & \cos\varphi_c & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Вектор элемента контура можно вычислить по формуле (ток течет по направлению вектора)

$$\mathbf{dr} = \{(r(\varphi + d\varphi) - r(\varphi))\mathbf{\Omega}\} = \{(-R\sin\varphi d\varphi; R\cos\varphi d\varphi; 0)\mathbf{\Omega}\}.$$

В общем виде $\mathbf{dr} = \{dr_x; dr_y; dr_z\}$. Вектор между точкой на проводнике и точкой измерений

$$\mathbf{r}_0 - \mathbf{r} = \{(r_{x0}; r_{y0}; r_{z0}) - (R\cos\varphi; R\sin\varphi; 0)\mathbf{\Omega}\}.$$

В общем виде $\mathbf{r}_0 - \mathbf{r} = \{r_x; r_y; r_z\}$.

Используя приведенные величины и формулу (1), можно ввести проекции вектора индукции от кругового витка на оси координат в точке измерений:

$$B_x(\mathbf{r}_0) = \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{r_y dr_z - r_z dr_y}{|\mathbf{r}_0 - \mathbf{r}|^3},$$
$$B_y(\mathbf{r}_0) = \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{r_z dr_x - r_x dr_z}{|\mathbf{r}_0 - \mathbf{r}|^3}, \quad (2)$$
$$B_z(\mathbf{r}_0) = \frac{\mu_0 I}{4\pi}\int_0^{2\pi} \frac{r_x dr_y - r_y dr_x}{|\mathbf{r}_0 - \mathbf{r}|^3}.$$

Проекции вектора индукции произвольно ориентированного кругового витка с током на оси двух магнитометров. Рассмотрим проекции поля кругового витка с током на оси двух магнитометров, располагающихся на оси Z в точках $M_1 = (0, 0, L_1)$ и $M_2 = (0, 0, L_2)$ (рис. 1).

Введем обозначения: H_x, H_y, H_z — проекции поля, измеряемого магнитометрами; верхние индексы (M_1) и (M_2) — принадлежность компонент соответствующим магнитометрам. Тогда разности показаний по осям магнитометров, возникающие при внешнем воздействии, аппроксимированном витком с током (2), запишем в виде:

$$H_x^{(M_1)} - H_x^{(M_2)} = B_x(M_1) - B_x(M_2),$$
$$H_y^{(M_1)} - H_y^{(M_2)} = B_y(M_1) - B_y(M_2), \quad (3)$$
$$H_z^{(M_1)} - H_z^{(M_2)} = B_z(M_1) - B_z(M_2).$$

Если подобраны параметры витка, удовлетворяющие (3), данные опорного (M_2) магнитометра могут быть скомпенсированы на величины влияния,

$$\hat{H}_x^{(M_2)} = H_x^{(M_2)} - B_x(M_2),$$
$$\hat{H}_y^{(M_2)} = H_y^{(M_2)} - B_y(M_2), \quad (4)$$
$$\hat{H}_z^{(M_2)} = H_z^{(M_2)} - B_z(M_2).$$

С учетом полученных (4) значений $\hat{H}_x^{(M_2)}$, $\hat{H}_y^{(M_2)}$ и $\hat{H}_z^{(M_2)}$ вычислим скорректированный магнитный азимут прибора

$$\alpha = \mathrm{arctg}\left(\frac{\hat{H}_x^{(M_2)}\sin\varphi_{Inc} + \hat{H}_y^{(M_2)}\cos\varphi_{Inc}}{\hat{H}_x^{(M_2)}\cos\theta_{Inc}\cos\varphi_{Inc} - \hat{H}_y^{(M_2)}\cos\theta_{Inc}\sin\varphi_{Inc} - \hat{H}_z^{(M_2)}\sin\theta_{Inc}}\right), \quad (5)$$

где θ_{Inc} и φ_{Inc} — зенитный и визирный углы инклинометра, вычисляемые на основе показаний акселерометров.

Проекции векторов индукции двух произвольно ориентированных круговых витков с током на оси двух магнитометров. В том случае, когда инклинометр используется в транзитном исполнении (не в конце сборки), на магнитометры будут оказывать влияние не один, а два источника ПОН (F_1 и F_2 на рис. 2, *c*, *d*). Источник 2 располагается в точке $S = (0; 0; L_3)$. В таком случае необходимо найти параметры двух эквивалентных источников, располагающихся сверху и снизу от магнитного градиентометра:

$$H_x^{(M_1)} - H_x^{(M_2)} = B_x^{(F_1)}(M_1) + B_x^{(F_2)}(M_1 - S) - B_x^{(F_1)}(M_2) - B_x^{(F_2)}(M_2 - S);$$
$$H_y^{(M_1)} - H_y^{(M_2)} = B_y^{(F_1)}(M_1) + B_y^{(F_2)}(M_1 - S) - B_y^{(F_1)}(M_2) - B_y^{(F_2)}(M_2 - S); \quad (6)$$
$$H_z^{(M_1)} - H_z^{(M_2)} = B_z^{(F_1)}(M_1) + B_z^{(F_2)}(M_1 - S) - B_z^{(F_1)}(M_2) - B_z^{(F_2)}(M_2 - S).$$

Случай двух источников ПОН соответствует использованию забойного датчика, в связи с чем приве-

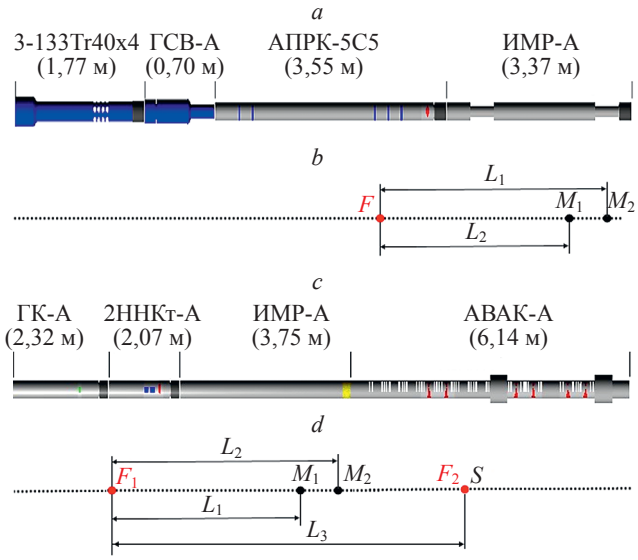

Рис. 2. Варианты расположения источников полей остаточной намагниченности (на примере геофизических сборок на буровом инструменте): концевое (*a*), (*b*) и транзитное (*c*), (*d*) использования инклинометра магниторезистивного (ИМР).

Нижний конец сборки — справа, буровой инструмент — слева от указанных сборок.
3-133Tr40x4 — переводник со сборки на буровой инструмент; ГСВ — головка свободного вращения. Приборы: АПРК-5С5 — радиоактивного; ГК — гамма; 2ННКт — двойного нейтронного; АВАК — волнового акустического каротажа. Добавка «-А» — прибор автономный

Fig. 2. Variants of RMF (Remanent Magnetization Field) source locations (example: geophysical tool assemblies on a drilling string): end placement (*a*), (*b*); transit placement of magnetoresistive inclinometer (MRI) (*c*), (*d*).

In the assemblies shown, the lower end is at the right and the drilling tool is at the left. Legend: 3-133Tr40x4 — sub from the assembly to the drilling tool; ГСВ — free-rotating head; АПРК-5С5 — radioactive tool; ГК — gamma; 2ННКт — dual neutron; АВАК — sonic logging. Suffix '-A' denotes an autonomous tool

денная (6) модель может быть использована во время бурения.

Описание средств для проведения эксперимента. Эксперимент проводился с использованием двух инклинометрических датчиков БДИ-32 производства ООО «Нефтегазгеофизика», расположенных в одной трубе, размещенной на калибровочной установке. Расстояние между магнитометрами составляло 0,307 м ($L_2 - L_1$ на рис. 1).

Проверку проходил вариант с источником с одной стороны (рис. 1 и рис. 2, *a*, *b*). В качестве источников ПОН применены кольцевые неодимовые магниты, располагающиеся на оси трубы с датчиками. Использовалось три магнита со следующими радиусами и расстояниями до точки M_1 (рис. 1): источники 1 и 2 — $R = 0{,}013$ м, $L_1 = 0{,}515$ м; источник 3 — $R = 0{,}023$ м, $L_1 = 0{,}481$ м.

Таким образом, в расчетном процессе осуществлялся подбор только ориентации (θ_c и φ_c (рис. 1)), направления и силы тока I витка. Ток витка подбирался из условия его минимального значения, удовлетворяющего критерию (3).

Основы алгоритма расчета параметров эквивалентного витка. Параметры ориентации и сила тока имитирующего ПОН витка подбирались методом золотого сечения таким образом, чтобы разности показаний реальных магнитометров в M_1 и M_2 совпали с моделируемыми разностями (3) в тех же точках. При этом вычисление проекций индукции поля от витка на оси магнитометров (2) производилось путем численного интегрирования.

Основные результаты

Параметры витков с током, имитирующих источники ПОН, приведены в таблице. Параметры были получены при помощи предложенного алгоритма.

Для демонстрации эффективности метода на рис. 3 приведены разности расчетных (5) и заданных на калибровочной установке значений магнитного азимута. Зенитные углы во время измерений: 30° и 90°. Значения азимутов 90° и 270° обусловлены тем, что именно при данных азимутах вектор ПОН направлен перпендикулярно вектору ГМП и, следовательно, его влияние ярко выражено [20]. Измерения проводились при разных визирных углах (углах разворота корпуса инклинометра вокруг своей оси).

Зелеными квадратами на рис. 3 указаны погрешности расчетных значений азимута без влияния ПОН. Как видно из графиков (синие круги), скорректированные на влияние ПОН, значения магнитного азимута (5) не отклонялись от заданных значений более чем на 0°39′ в случае трех источников паразитного поля. Фактически, диапазон погрешностей после корректировки не отличался от диапазона погрешностей без влияния ПОН. В то же время максимальные погрешности измерения азимута без корректировки (красные треугольники) были следующими: –1°15′36″ (источник 1), –3°9′36″ (источник 2), 12°30′36″ (источник 3).

Несложно заметить, что погрешности при зенитных углах 30° (рис. 3, *a*, *c*, *e*) меньше погрешностей при зенитных углах 90° (рис. 3, *b*, *d*, *f*). Это полностью укладывается в представленную модель влияния ПОН [20].

Таблица. Расчетные параметры витков с током, эквивалентных источникам ПОН

Table. Calculated parameters of current loops equivalent to Remanent Magnetization Field (RMF) sources

Номер источника	Модуль магнитной индукции источника в точке M_2, нТл	Доля от ГМП, %	θ_c,°	φ_c,°
1	216,87	0,42	7°12′36″	283°45′36″
2	811,34	1,59	1°39′36″	241°0′36″
3	2965,57	5,80	5°37′48″	284°39′

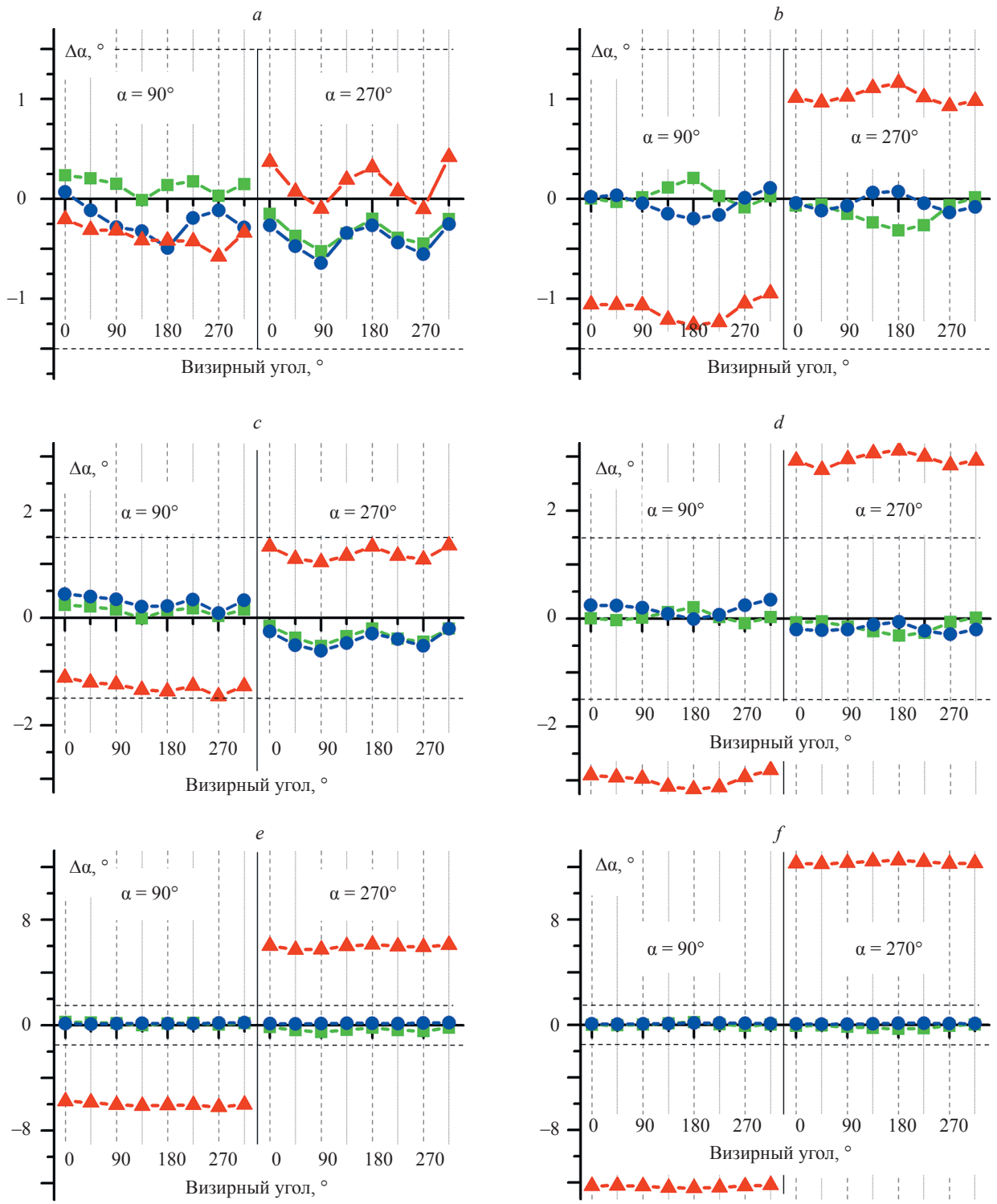

Рис. 3. Разности Δα расчетных и заданных на калибровочной установке значений магнитного азимута при разных визирных углах: без влияния поля остаточной намагниченности (зеленые квадраты), с влиянием поля остаточной намагниченности (красные треугольники) и с влиянием, скомпенсированным по алгоритму из настоящей работы (синие круги). Зенитные углы во время измерений: 30° (*a*, *c*, *e*); 90° (*b*, *d*, *f*). Пунктирными линиями указаны рамки допуска для погрешности измерения азимута (±1°30′). Исследуемые источники: 1 (*a*, *b*); 2 (*c*, *d*); 3 (*e*, *f*)

Fig. 3. Differences (Δα) between calculated and reference magnetic azimuth values on the calibration rig for different sighting inclination angles: without RMF influence (green squares), with RMF influence (red triangles), and with RMF influence compensated by the algorithm proposed here (blue circles). Inclination angles during measurements: 30° (*a*, *c*, *e*); 90° (*b*, *d*, *f*). Dotted lines show the azimuth error tolerance limits (±1°30′). RMF source indices: 1 (*a*, *b*); 2 (*c*, *d*); 3 (*e*, *f*)

Обсуждение

Опыт обработки реальных каротажных данных показывает, что значение модуля магнитной индукции ПОН в месте расположения магнитометров, обычно, находится в диапазоне от 1 % до 6 % от величины ГМП [13, 17]. Кроме того, вектора ПОН не направлены строго по оси сборки [13]. Таким образом, выбранные источники (таблица) отражают свойства полей реальных источников.

Однако стоит отметить, что данные погрешности измерения азимута получены для г. Твери, где угол наклонения ГМП равен 71°25′48″. Из-за меньшей горизонтальной составляющей вектора ГМП, в случае высоких широт значения погрешностей были бы больше [20] и могли доходить до 4–7° даже в случае источника 2. В связи с этим рассматриваемую технологию можно считать особенно актуальной для использования на месторождениях, располагающихся на высоких широтах, включая перспективные месторождения Арктического шельфа.

Помимо корректировки магнитного азимута непосредственно во время бурения или каротажа, предложенный в работе метод компенсации влияния ПОН позволяет удешевить охранный кожух инклинометра вследствие уменьшения длины дорогостоящей немагнитной части, а также уменьшить длину немагнитных вставок по обе стороны от датчика при бурении.

Из минусов методики можно отметить удорожание инклинометрического датчика вследствие необходимости установки второго магнитометра. Тем не менее, в силу сравнительно небольшой стоимости магнитометров, удешевление корпуса и отсутствие необходимости дополнительной апостериорной обработки существенно снижают издержки. Также к минусам можно отнести усложнение алгоритмов калибровки вследствие необходимости согласования данных магнитометров. В то же время стоит отметить, что усложнение алгоритмов не усложняет работу оператора.

Заключение

Представленный в работе метод вычисления параметров поля остаточной намагниченности на основе измерений градиента суперпозиционного (суммы поля остаточной намагниченности и геомагнитного поля) магнитного поля, позволяет скорректировать погрешность измерения магнитного азимута в рамки допуска. Причем, для данной коррекции не требуются измеренные или эталонные значения параметров геомагнитного поля для места проведения каротажа. Кроме того, метод позволяет вычислять влияние не только осевой, но и радиальной составляющей, а также подстраиваться под изменения параметров поля остаточной намагниченности в случае перемагничивания или иных причин.

Таким образом, данный метод может быть использован при инклинометрии скважин непосредственно во время проведения работ. Для этого необходимо реализовать инклинометр-градиентометр в одном корпусе, а в процесс наладки внести процедуру согласования магнитометров. Данный датчик позволит сократить расходы на немагнитную часть корпуса, а также снизить требования по части наличия немагнитных вставок рядом с инклинометром.

Литература

1. Jamieson A. Introduction to Wellbore Positioning. University of the Highlands & Islands, 2012. 164 p.
2. Бармаков Ю.Н., Герасимчук О.А., Козырев П.В., Неуструев В.В., Филатов М.М., Юрков Д.И. Исследование и разработка методов для устранения намагниченности оборудования буровой установки // Каротажник. 2019. № 1 (295). С. 98–104.
3. Шафигуллин Р.И., Еромасов В.Г., Андиряков В.Ф., Никулин О.В. Устройство для размагничивания бурового инструмента. Патент RU № 2591056C1. Бюл. 2016. № 19.
4. Сидоров А.А., Харбаш В.Я., Шурыгин С.В. Способ определения и компенсации магнитной девиации инклинометра. Патент RU № 2186966C2. Бюл. 2002. № 16.
5. Рыжков И.В., Ковшов Г.Н. Расчет магнитной девиации инклинометра в условиях буровой // Вестник Приднепровской государственной академии строительства и архитектуры. 2011. № 11-12 (164-165). С. 86–91.
6. Чупров В.П., Васильев А.В., Кудряшов А.А. Коррекция влияния намагниченности компоновки низа бурильной колонны на измерения азимута при навигации скважин с помощью телесистем // Каротажник. 2016. № 7 (265). С. 131–135.
7. Дмитрюков Ю.Ю., Исмагилов М.А. Способ измерения магнитного азимута в процессе бурения. Патент RU № 2349938C1. Бюл. 2009. № 8.
8. Горичка М.В., Кузнецов А.Б., Абзалов З.З., Бевзенко В.А. Замеры высокого разрешения // Бурение и нефть. 2018. № 9. С. 24–30.
9. Кейн С.А., Трохов В.В. Разработка технико-технологических рекомендаций по повышению качества выполнения проектной траектории наклонно-направленных скважин // Строительство нефтяных и газовых скважин на суше и на море. 2015. № 1. С. 4–7.
10. Priest J., Quinn T., Frost E.Jr. Magnetic interference effects on accelerometer and magnetometer data: detection, quality control and

References

1. Jamieson A. *Introduction to Wellbore Positioning*. University of the Highlands & Islands, 2012, 164 p.
2. Barmakov Yu.N., Gerasimchuk O.A., Kozyrev P.V., Neustruev V.V., Filatov M.M., Yurkov D.I. Researches and working out methods to eliminate drilling-rig magnetization. *Karotazhnik*, 2019, no. 1 (295), pp. 98–104. (in Russian)
3. SHafigullin R.I., Eromasov V.G., Andiryakov V.F., Nikulin O.V. Device for demagnetisation of drilling tool. *Patent RU № 2591056C1,* 2016. (in Russian)
4. Sidorov A.A., Kharbash V.Ja., Shurygin S.V. Method of determination and compensation of inclinometer magnetic deviation. *Patent RU № 2186966C2*, 2002. (in Russian)
5. Ryzhkov I.V., Kovshov G.N. The calculation ofthe magneticdeviation ofthe inclinometerinthe drilling. *Bulletin of the pridnipro state academy of construction and architecture*, 2011, no. 11-12 (164-165), pp. 86–91. (in Russian)
6. Chuprov V.P., Vasilev A.V., Kudryashov A.A. Correction for the effect of bottomhole assembly magnetization on azimuth measurements in borehole navigation by telesystems. *Karotazhnik*, 2016, no. 7 (265), pp. 131–135. (in Russian)
7. Dmitrjukov J.J., Ismagilov M.A. Method of magnetic azimuth measurement while drilling. *Patent RU № 2349938C1*, 2009. (in Russian)
8. Gorichka M.V., Kuznetcov A.B., Abzalov Z.Z., Bevzenko V.A. High definition surveys. *Burenie i neft*, 2018, no. 9, pp. 24–30. (in Russian)
9. Keyn S.A., Trokhov V.V. Development of technical and technological recommendations on quality enhancement of project trajectory performance of directional wells. *Construction of Oil and Gas Wells on Land and Sea*, 2015, no. 1, pp. 4–7. (in Russian)
10. Priest J., Quinn T., Frost E.Jr. Magnetic interference effects on accelerometer and magnetometer data: detection, quality control and

correction // Proc. of the 55th Annual Logging Symposium SPWLA. 2014. P. 1–13.
11. Гринев И.В., Королев А.Б., Ситников В.Н. Контроль качества инклинометрических измерений. Учет суточных вариаций магнитного поля // Каротажник. 2015. № 12 (258). С. 99–108.
12. Buchanan A., Finn C., Love J., Worthington W., Lawson F., Maus S., Okewunmi S., Poedjono B. Geomagnetic referencing - the real-time compass for directional drillers // Oilfield Review. 2013. V. 25. N 3. P. 32–47.
13. Гринев И.В., Королев А.Б., Ситников В.Н., Гасанов О.В. Влияние радиальной составляющей поля остаточной намагниченности на показания инклинометра // Каротажник. 2025. № 1 (333). С. 147–153.
14. Li Z., Geng Y., Yang Y., Wang W., Hang S. Estimation of geomagnetic components under unknown interferences for drilling tools // Measurement Science and Technology. 2024. V. 35. N 5. P. 056310. https://doi.org/10.1088/1361-6501/ad2b47
15. El Gizawy M., Lowdon R., Aklestad D.L. Combining magnetic and gyroscopic surveys provides the best possible accuracy // SPE Drilling and Completion. 2023. V. 38. N 4. P. 586–593. https://doi.org/10.2118/212547-PA
16. Zhang Q., Pang H., Wan C. Magnetic interference compensation method for geomagnetic field vector measurement // Measurement. 2016. V. 91. P. 628–633. https://doi.org/10.1016/j.measurement.2016.05.081
17. Гринев И.В., Королев А.Б., Ситников В.Н. Компенсация влияния поля остаточной намагниченности бурильной колонны и сборки геофизических приборов на показания инклинометра // Каротажник. 2020. № 1 (301). С. 104–111.
18. Биндер Я.И., Вольфсон Г.Б., Гаспаров П.М., Клюшкин П.А., Розенцвейн В.Г. Компенсация магнитных помех в феррозондовом инклинометре // Гироскопия и навигация. 2005. № 1 (48). С. 68–75.
19. Биндер Я.И., Клюшкин П.А., Тихонов А.Г. Экспериментальное исследование магнитометрической системы ориентации ствола скважины с компенсацией магнитных помех // Каротажник. 2010. № 1 (190). С. 63–69.
20. Гринев И.В., Королев А.Б., Ситников В.Н. Влияние остаточной намагниченности бурильной колонны и сборки геофизических приборов на показания инклинометра // Каротажник. 2019. № 4 (298). С. 87–95.

correction. *Proc. of the 55th Annual Logging Symposium SPWLA*, 2014, pp. 1–13.
11. Grinev I.V., Korolev A.B., Sitnikov V.N. Inclinometry measurements quality control. Corrections for daily variations in the magnetic field. *Karotazhnik*, 2015, no. 12 (258), pp. 99–108. (in Russian)
12. Buchanan A., Finn C., Love J., Worthington W., Lawson F., Maus S., Okewunmi S., Poedjono B. Geomagnetic referencing - the real-time compass for directional drillers. *Oilfield Review*, 2013, vol. 25, no. 3, pp. 32–47.
13. Grinev I.V., Korolev A.B., Sitnikov V.N., Gasanov O.V. The effect of the radial component of the remnant magnetization field on the inclinometer readings. *Karotazhnik*, 2025, no. 1 (333), pp. 147–153. (in Russian)
14. Li Z., Geng Y., Yang Y., Wang W., Hang S. Estimation of geomagnetic components under unknown interferences for drilling tools. *Measurement Science and Technology*, 2024, vol. 35, no. 5, pp. 056310. https://doi.org/10.1088/1361-6501/ad2b47
15. El Gizawy M., Lowdon R., Aklestad D.L. Combining magnetic and gyroscopic surveys provides the best possible accuracy. *SPE Drilling and Completion*, 2023, vol. 38, no. 4, pp. 586–593. https://doi.org/10.2118/212547-PA
16. Zhang Q., Pang H., Wan C. Magnetic interference compensation method for geomagnetic field vector measurement. *Measurement*, 2016, vol. 91, pp. 628–633. https://doi.org/10.1016/j.measurement.2016.05.081
17. Grinev I.V., Korolev A.B., Sitnikov V.N. Compensation of the effect of the field of remanent magnetization of the drillstem and well logging tools string on the inclinometer readings. *Karotazhnik*, 2020, no. 1 (301), pp. 104–111. (in Russian)
18. Binder Ya.L., Volfson G.B., Gasparov P.M., Klushkin P.A., Rozentsvein V.G. Magnetic disturbances compensation in flux-gate inclinometer. *Gyroscopy and Navigation*, 2005, no. 1 (48), pp. 68–75. (in Russian)
19. Binder Ya.I., Klyushkin P.A., Tikhonov A.G.. Experimental investigation on a magnetometric system for borehole orientation with magnetic noise compensation. *Karotazhnik*, 2010, no. 1 (190), pp. 63–69. (in Russian)
20. Grinev I.V., Korolev A.B., Sitnikov V.N. The effect of the remanent magnetization of the drillstem and logtool set on the inclinometer readings. *Karotazhnik*, 2019, no. 4 (298), pp. 87–95. (in Russian)

Авторы

Гасанов Олег Валерьевич — заведующий отделом, ООО «Нефтегазгеофизика», Тверь, 170033, Российская Федерация, sc 56548899800, https://orcid.org/0009-0004-6275-688X, gasanov.o@karotazh.ru

Ситников Виталий Николаевич — инженер, ООО «Нефтегазгеофизика», Тверь, 170033, Российская Федерация, sc 57190258699, https://orcid.org/0009-0009-0226-3297, v.sitnikov@karotazh.ru

Хаматдинов Вадим Рафисович — кандидат технических наук, генеральный директор, ООО «Нефтегазгеофизика», Тверь, 170033, Российская Федерация, https://orcid.org/0009-0006-3257-0317, hvadim@karotazh.ru

Гринев Илья Викторович — кандидат физико-математических наук, заведующий лабораторией, ООО «Нефтегазгеофизика», Тверь, 170033, Российская Федерация, sc 37661261600, https://orcid.org/0009-0003-9078-4792, i.grinev@karotazh.ru

Королев Александр Борисович — эксперт-консультант, ООО «Нефтегазгеофизика», Тверь, 170033, Российская Федерация, https://orcid.org/0009-0002-1290-574X, korolev@karotazh.ru

Authors

Oleg V. Gasanov — Head of Department, LTD Neftegazgeofizika, Tver, 170033, Russian Federation, sc 56548899800, https://orcid.org/0009-0004-6275-688X, gasanov.o@karotazh.ru

Vitaly N. Sitnikov — Engineer, LTD Neftegazgeofizika, Tver, 170033, Russian Federation, sc 57190258699, https://orcid.org/0009-0009-0226-3297, v.sitnikov@karotazh.ru

Vadim R. Khamatdinov — PhD, General Manager, LTD Neftegazgeofizika, Tver, 170033, Russian Federation, https://orcid.org/0009-0006-3257-0317, hvadim@karotazh.ru

Ilya V. Grinyov — PhD (Physics & Mathematics), Head of Laboratory, LTD Neftegazgeofizika, Tver, 170033, Russian Federation, sc 37661261600, https://orcid.org/0009-0003-9078-4792, i.grinev@karotazh.ru

Alexander B. Korolyov — Expert Advisor, LTD Neftegazgeofizika, Tver, 170033, Russian Federation, https://orcid.org/0009-0002-1290-574X, korolev@karotazh.ru

Статья поступила в редакцию 19.02.2025
Одобрена после рецензирования 02.09.2025
Принята к печати 22.11.2025

Received 19.02.2025
Approved after reviewing 02.09.2025
Accepted 22.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1185-1196

DeFs-CBDE: Clustering-guided binary mutation in multi-objective differential evolution for microarray gene selection

Mohamed Djellal Serandi[1]✉, Fatma Boufera[2], Amina Houari[3], Farid Flitti[4]

[1,2,3] University Mustapha Stambouli, LISYS laboratory, Mascara, 29000, Algeria
[4] Higher Colleges of Technology in Dubai, Dubai, 500001, United Arab Emirates

[1] mohamed.djellalserandi@univ-mascara.dz✉, https://orcid.org/0009-0009-5775-7956
[2] fboufera@univ-mascara.dz, https://orcid.org/0000-0002-5733-586X
[3] amina.houari@univ-mascara.dz, https://orcid.org/0000-0002-3628-7483
[4] fflitti@hct.ac.ae, https://orcid.org/0000-0002-2480-2580

Abstract
DNA microarray technology produces high-dimensional gene expression data, where many genes are irrelevant to disease. Effective feature selection is thus essential to mitigate the curse of dimensionality and enhance classification performance. This study introduces a multi-objective feature selection approach employing a Clustering-Based Binary Differential Evolution (CBDE) mutation to identify a compact set of disease-relevant genes. The proposed DeFs-CBDE algorithm was assessed on four gene expression datasets, i.e. brain, breast, lung, and central nervos system cancer by selecting informative feature subsets and evaluating them using five state-of-the-art classifiers, i.e., Support Vector Machine (SVM), Naive Bayes, K-Nearest Neighbors, Decision Tree (DT), and Random Forest. The DeFs-CBDE method achieved of 100 % accuracy on the brain dataset with three classifiers. On the lung dataset, DeFs-CBDE reached 97.56 % accuracy with SVM and DT. For the breast dataset, DeFs-CBDE attained 93.33 % accuracy very close to the highest score of 93.81 % accuracy. The CNS dataset proved the most challenging, where it achieved 91.67 % accuracy with SVM. Across all datasets, DeFs-CBDE consistently achieved high classification performance.

Keywords
microarray data, feature selection, optimization, differential evolution, mutation, multi-objective

For citation: Djellal Serandi M., Boufera F., Houari A., Flitti F. DeFs-CBDE: Clustering-guided binary mutation in multi-objective differential evolution for microarray gene selection. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1185–1196. doi: 10.17586/2226-1494-2025-25-6-1185-1196

УДК 004.93'14

DeFs-CBDE: бинарная мутация, управляемая кластеризацией, в многокритериальной дифференциальной эволюции для отбора генов с помощью микрочипов

Мохамед Джеллаль Серанди[1]✉, Фатма Буфера[2], Амина Хуари[3], Фарид Флитти[4]

[1,2,3] Университет Мустафы Стамбули, Маскара, 29000, Алжир
[4] Высший колледж технологии, колледж в Дубае, Дубай, 500001, Объединенные Арабские Эмираты

[1] mohamed.djellalserandi@univ-mascara.dz✉, https://orcid.org/0009-0009-5775-7956
[2] fboufera@univ-mascara.dz, https://orcid.org/0000-0002-5733-586X
[3] amina.houari@univ-mascara.dz, https://orcid.org/0000-0002-3628-7483
[4] fflitti@hct.ac.ae, https://orcid.org/0000-0002-2480-2580

Аннотация
Технология ДНК-микрочипов позволяет получать высокоразмерные данные об экспрессии генов, многие из которых не имеют отношения к заболеванию. Эффективный отбор признаков, таким образом, необходим для смягчения «проклятия размерности» и повышения эффективности классификации. В данном исследовании

© Djellal Serandi M., Boufera F., Houari A., Flitti F., 2025

представлен многокритериальный подход к отбору признаков с использованием мутации на основе кластеризации для идентификации компактного набора генов, связанных с заболеванием. Предложенный алгоритм DeFs-CBDE был оценен на четырех наборах данных об экспрессии генов: рак головного мозга, молочной железы, легких и центральной нервной системы, путем отбора информативных подмножеств признаков и их оценки с использованием пяти современных классификаторов, а именно: метода опорных векторов, наивного байесовского алгоритма, метода K-ближайших соседей, дерева решений и случайного леса. Метод DeFs-CBDE достиг 100 % точности на наборе данных о мозге с тремя классификаторами. В наборе данных по легким DeFs-CBDE достиг точности 97,56 % с использованием метода опорных векторов и дерева решений. В наборе данных по молочной железе DeFs-CBDE достиг точности 93,33 %, что очень близко к максимальному результату в 93,81 %. Набор данных по центральной нервной системе оказался самым сложным, где точность составила 91,67 % с использованием. Во всех наборах данных DeFs-CBDE стабильно демонстрировал высокую эффективность классификации.

Ключевые слова
данные микрочипов, выбор признаков, оптимизация, дифференциальная эволюция, мутация, многокритериальный анализ

Ссылка для цитирования: Джеллаль Серанди М., Буфера Ф., Хуари А., Флитти Ф. DeFs-CBDE: бинарная мутация, управляемая кластеризацией, в многокритериальной дифференциальной эволюции для отбора генов с помощью микрочипов // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1185–1196 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-6-1185-1196

Introduction

A challenge faced in the explosion of microarray gene expression data is its high dimensionality (thousands of genes) combined with a limited sample number (tens or hundreds), leading to the well-known "curse of dimensionality" [1]: As the dimensionality increases, the number of samples required for inference grows rapidly. This problem is addressed by Feature Selection (FS) which aims at selecting a pertinent subset of genes to enhance model performance and interpretability. The development of high-throughput genomics, such as microarrays, has indeed transformed our understanding of diseases such as breast cancer. However, although microarray data provide a wealth of information, appropriate feature selection is necessary to identify genes that significantly influence disease classification.

Choosing the optimal subset of the original features is NP-Hard and requires a lot of time and resources. As a result, selecting an appropriate search algorithm is crucial for feature selection methods. A feature selection algorithm should be assessed from the perspective of effectiveness and efficiency. The effectiveness of a feature selection method is determined by the time it takes to find the final set of [2].

In general, it is possible to represent feature selection as a multi-objective [3] combinatorial optimization problem with two key objectives: minimizing the number of selected features and the Mean Squared Residual (MSR) [4]. Multi-objective evolutionary algorithms enable the search for multiple solutions located on the Pareto front in a single run. These methods are frequently used to solve feature selection problems [5].

One of the most widely used evolutionary algorithms for FS problems is Differential Evolution (DE) which is renowned for its ease of use and effectiveness. The application of DE to the multi-objective scenario has not received much attention, and most current research focuses on maximizing classification accuracy, or the single-objective case [5]. A multi-objective FS approach was initially developed using multi-objective DE and applied to the feature selection task. The results highlight the effectiveness of DE in addressing feature selection challenges.

This research suggests a novel multi-objective feature selection method based on DE for microarray data to improve the model performance and interpretability. The MSR is optimized simultaneously with limiting the number of selected genes. Additionally, a novel mutation operator, named Clustering-Based Binary Differential Evolution (CBDE) is introduced to adapt the algorithm for better optimization in feature selection.

The highlights of the DeFs-CBDE method are:

— A new FS approach for gene expression datasets is proposed.
— A novel CBDE mutation operator is proposed.
— The DeFs-CBDE is extensively validated on four datasets (Brain, Breast, Lung, and Central Nervous System (CNS)).
— The performance obtained with the proposed method is compared to the performance of the FS technique in [6] and with other existing works [7–9].

Related work

Many researchers have focused on FS to improve classification performance on gene expression data. Among the most commonly used approaches are Particle Swarm Optimization (PSO), DE, and hybrid methods, which have shown promising results across various cancer-related datasets. In the following, we present related studies based on these techniques.

Zhang et al. [5] proposed a Multi-Objective Feature Selection method called the Binary Differential Evolution (BDE) with self-learning (MOFS-BDE). The method introduces novel operators, i.e., the probability difference based binary mutation, the One-bit Purifying Search and the efficient non-dominated sorting with crowding distance. Experimental results show that the MOFS-BDE achieves balance between global exploration and local exploitation, and is highly competitive compared to state-of-the-art genetic algorithm, particle swarm, differential evolution, and artificial bee colony-based feature selection algorithms.

Ali et al. [6] proposed a feature selection method based on a hybrid filter-genetic algorithm to enhance the classification of cancer in high dimensional microarray datasets. The method first discards irrelevant features using filter-based techniques including Information Gain (IG), Information Gain Ratio (IGR), and Chi-Squared (CS). Then, a Genetic Algorithm (GA) is used to further optimize the selected features for cancer classification. The approach was tested on four microarray datasets (Brain, Breast, Lung, and CNS) and compared to six other FS methods. The experimental results showed that the GA in the proposed method was able to reduce further 50 % of irrelevant features from the subset selected by filtering.

Ahadzadeh et al. [7] introduced an algorithm called Simple, Fast, and Efficient FS to search a vast area of high-dimensional data, eliminating irrelevant characteristics from the search space in the first phase. Next, the smaller search space is explored using the PSO method to identify significant and pertinent features.

Prajapati et al. [8] proposed a feature selection approach based on DE to reduce the dimensionality of microarray datasets. The method improves classification accuracy by combining DE with Random Forest (RF), Decision Tree (DT), and LR (Logistic Regression), and it outperforms models without feature selection.

Hamla and Ghanem [9] proposed a hybrid feature selection method that combines Fisher score filtering and Support Vector Machine Recursive Feature Elimination (SVM-RFE) to improve gene selection in microarray data.

Paul et al. [10] introduced a multi-objective PSO-based feature selection method designed for multi-label classification with online arrival of features. The approach automatically identifies the most relevant subset of features through a structured three-phase filtering process. In the first phase, a multi-objective PSO selects candidate features from each incoming group. The second phase eliminates redundancy by comparing newly selected features with those previously chosen. In the final phase, the method prunes previously selected features that become insignificant due to the arrival of new data. This adaptive and evolutionary framework effectively maintains a compact and relevant feature set, improving multi-label classification.

Han et al. [11] suggested a Multi-Objective PSO with Adaptive Strategies (MOPSO-ASFS) for FS. The approach uses an adaptive penalty of Penalty Boundary Interaction mechanism to enhance the selection pressures of the archive. An adaptive leading particle selection based on feature information combines the opposite mutation and the feature frequencies to improve the selection pressure of each particle. The experimental results from 14 UCI datasets and 6 gene expression datasets demonstrated the effectiveness of MOPSO-ASFS in high-dimensional space compared to classical multi-objective FS algorithms, the multi-dimensional regions.

Dashtban et al. [12] proposed a bioinspired multi-objective algorithm that improves the Bat Algorithm with refined formulations, effective multi-objective operators, and social learning-based random walks. Fisher criterion is used as initial filtering to reduce irrelevant genes. Applied to three cancer microarray datasets, results revealed new combinations of important biomarkers associated with the development of three challenging cancers. Also, the algorithm showed strong classification performance.

Ghosh et al. [13] introduced a two-phase feature selection framework designed for microarray data analysis. In the first phase, they applied an ensemble filtering strategy that combined the top-ranked features identified by symmetrical uncertainty, chi-square, and ReliefF methods using union and intersection operations. The refined feature subsets were then further optimized with a genetic algorithm to select the most informative features.

According to this state-of-the-art, GA, PSO, and DE have demonstrated strong potential in addressing the challenges of feature selection, particularly in high-dimensional and noisy data environments. Among them, GA and DE are widely recognized for their effectiveness. DE, in particular, is distinguished by its simplicity, adaptability, and its ability to incorporate advanced mutation strategies such as CBDE.

Proposed method

In this paper, an improved multi-objective differential evolution approach for FS (DeFs-CBDE) method, based on a CBDE mutation operator, is presented. This novel method uses a Non-dominated Sorting Genetic Algorithm (NSGA-II).

Differential evolution

Differential evolution is a technique that has been effectively used in many scientific and technical domains in recent years to solve continuous optimization challenges. DE was first proposed by [14] as a straightforward but effective heuristic for global optimization in continuous spaces. The DE algorithm consists of five main steps: initialization, mutation, crossover, selection, and termination [15].

Architecture of the proposed method

As shown in Fig. 1, the data set is first preprocessed and then divided into training and testing sets. The training set is then used to apply the proposed DeFs-CBDE algorithm resulting in a selected feature subset (New subset training with FS). The same features, without applying the DeFs-CBDE algorithm, are extracted to form the corresponding test set (New subset testing with FS). These new generated training and testing subsets are then fed into classifiers to evaluate the performance of selected features.

DeFs-CBDE algorithm description

By incorporating DE into the feature selection procedure, the suggested method described in Algorithm 1 guides the search for the best feature subsets.

Algorithm 1. The pseudocode of DeFs-CBDE approach.

Input: A partial training matrix of size $n \times m$ (n genes and m samples), Initial Population size p, Initial crossover probability C, Initial factor F;

Output: A matrix of size $s \times m$ (the dataset with s optimally selected features);

Begin

Initialize population P by generating p solutions, each gene random in [0, 1];

Compute fitness value ($f1$, $f2$) for each solution in P;

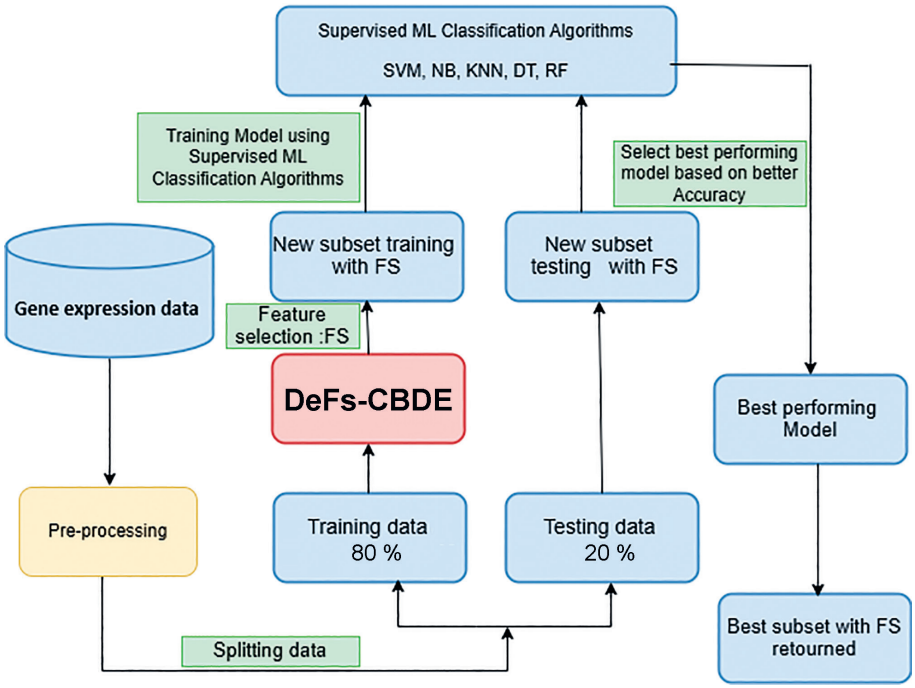

Fig. 1. Illustration of DeFs-CBDE scheme

Initialize array C and array F for all individuals in P using initial C and F;
While termination criteria not met **do**
  For each solution i in P **do**
    Select 3 randomly chosen solutions (x, y, z) from P ranked from best to worst based on fitness, ensuring $x \neq y \neq z \neq i$;
    Generate a mutant vector using **CBDE** operator: **CBDE** (x, y, z, Fi);
    Create a trial vector by performing binomial crossover between the mutant vector and the current solution i using crossover probability C;
    Evaluate the fitness of the trial vector using $f1$ and $f2$;
    if the trial vector is better than the current solution then Replace the current solution i with the trial vector;
    else Ci = normrnd (mean(C), 0.1); Fi = normrnd (mean (F), 0.1);
  End for
  Merge current population P with offspring population;
  Perform Non-Dominated Sorting using NSGA-II based on Pareto dominance;
  Calculate crowding distance for solutions within each rank and select top p solutions for the next generation based on rank and crowding distance;
end while
Return the final set of selected features;
Function CBDE (x, y, z, F) w = **GENEINSERTION** (x, **GENEREMOVAL** (y, z, F), F);
Function GENEINSERTION (x, y, F) **if** $x_i = 1$ or ($y_i = 1$ and $randi < F$) then return 1; **else** return x_i;
Function GENEREMOVAL (x, y, F) **if** $x_i = 0$ or ($y_i = 1$ and $randi < F$) then return 0; **else** return x_i.

Binary representation of features

The method uses BDE to perform efficient FS in a binary search space where each gene is represented by a bit. A variation of DE called BDE was created to address optimization issues with binary variables [15]. The aim of BDE is to effectively search discrete solution spaces where decision variables are binary. In BDE, the crossover and mutation operators have been adapted to work with binary variables. For every potential solution, a binary string of size n (genes) is used. Every bit in the string corresponds to a feature: a value of 1 means that the feature has been selected, whereas a value of 0 means that it has not. As illustrated in Fig. 2.

Population initialization

A parent population is initialized by randomly generating binary strings, each representing a potential subset of features.

Multi-Objective approach

A multi-objective approach uses many objective functions, as the name suggests. Many objectives or various criteria are prevalent in the majority of practical decision-making difficulties in many real-world or practical scenarios.

— Pareto-based evaluation: the Pareto-based rank assignment is used to evaluate individuals [16].
— NSGA-II: the fitness-sharing concept is applied within NSGA-II to promote solution diversity [17].

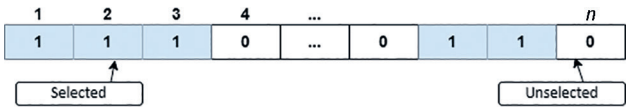

Fig. 2. Representation of binary string encoding for FS solution

CBDE mutation operator

The proposed CBDE operator is inspired by the Biclustering Binary Differential Evolution (BBDE) [15] strategy but introduces key improvements to better suit gene clustering for feature selection. Unlike BBDE, which simultaneously handles row and column structures, CBDE eliminates biclustering constraints and focuses exclusively on clustering genes into compact and discriminative subsets, a requirement more relevant to genomic feature selection.

Furthermore, CBDE operator is designed to enhance the mutation process by prioritizing solutions with better fitness scores, ensuring more effective exploration of the search space. Specifically, three vectors ($\mathbf{x_{r1}}$, $\mathbf{x_{r2}}$, and $\mathbf{x_{r3}}$) are randomly selected and ranked by fitness, with the best ($\mathbf{x_{r1}}$) strongly influencing mutation and the worst ($\mathbf{x_{r3}}$) being suppressed, thus biasing the mutation toward high-quality solutions and increasing the chance of producing better offspring. After ordering, the mutation operator combines the vectors based on the chosen strategy such as difference-based perturbation. A factor F (initially 0.65) controls gene insertion or removal by comparing a random value *randi* ∈ [0, 1], where i specifies the position of the current bit in the individual. This mechanism prevents solutions from degenerating into trivial cases (e.g., all 0s or all 1s), preserving population diversity and enabling efficient gene clustering. The CBDE method performs gene insertion or removal if the generated number *randi* is less than F; otherwise, the gene remains unchanged.

Empirically, our results originally presented in Tables 3 to 7 (provided in link[1]) demonstrate that this mechanism improves both classification accuracy and stability. Although these tables are no longer included directly in the manuscript, the provided link ensures access to the full experimental details. The consistent superiority of CBDE over other state-of-the-art methods further confirms its practical relevance.

The proposed CBDE operator introduces several innovations, including a clustering-oriented design for compact and non-redundant gene subsets, a fitness-guided ranking mechanism for knowledge-driven exploration, an adaptive dual insertion-removal strategy to preserve diversity, a perturbation tailored to gene clustering for biological interpretability and empirical validation, showing statistically significant improvements over conventional binary DE approaches.

$$CBDE = \text{GENEINSERTION}(\mathbf{x_{r1}}, \text{GENEREMOVAL}(\mathbf{x_{r2}}, \mathbf{x_{r3}}, F), F). \quad (1)$$

The new individual is evaluated using the fitness function.

$$Fi = \text{normrnd}(\text{mean}(F), 0.1), \quad (2)$$

$$CRi = \text{normrnd}(\text{mean}(CR), 0.1). \quad (3)$$

Fi and CRi represent the adapted versions of the mutation factor (F) and crossover probability (CR) for individual i, adjusting the mutation intensity and crossover rate respectively only when the individual shows no improvement, thereby ensuring effective self-adaptation and enhancing the algorithm's exploration capability.

If its score improves, the values of F and CR remain unchanged; otherwise, they are slightly adjusted using predefined Eqs. (2) and (3).

A summary of the CBDE mutation operation, using the adapted gene insertion and removal mechanisms tailored for clustering tasks is as follows:

$$\text{GENEINSERTION}(\mathbf{x}, \mathbf{y}, F) = \begin{cases} 1, \text{ if } x_i = 1 \text{ or } (y_i = 1 \text{ and } randi < F) \\ x_i, \text{ otherwise} \end{cases}, \quad (4)$$

$$\text{GENEREMOVAL}(\mathbf{x}, \mathbf{y}, F) = \begin{cases} 0, \text{ if } x_i = 0 \text{ or } (y_i = 1 \text{ and } randi < F) \\ x_i, \text{ otherwise} \end{cases}. \quad (5)$$

Illustrative example of the CBDE mutation operator

Consider the gene expression matrix $\mathbf{M} \in R^{5\times5}$, where rows represent genes and columns represent samples, and the initial population $\mathbf{P}$, consisting of four individuals represented by binary vectors (5 bits, one per gene):

$$\mathbf{M} = \begin{pmatrix} 42 & 15\,36 & 71\,24 \\ 33 & 58\,17 & 39\,65 \\ 13 & 45\,92 & 12\,26 \\ 57 & 81\,40 & 59\,31 \\ 21 & 14\,29 & 76\,54 \end{pmatrix}; \mathbf{P} = \begin{pmatrix} \mathbf{a:}\ 1\,0 & 1\,0 & 1 \\ \mathbf{b:}\ 0\,1 & 1\,0 & 0 \\ \mathbf{c:}\ 1\,1 & 0\,1 & 0 \\ \mathbf{d:}\ 0\,0 & 1\,1 & 1 \end{pmatrix}.$$

We apply the CBDE ($\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$) mutation using Eq. (1) with scaling factor $F = 0.65$ and a random vector: $\mathbf{rand} = [0.8, 0.9, 0.2, 0.5, 0.6]$.

Step 1: Gene Removal. We compute GENEREMOVAL ($\mathbf{c}$, $\mathbf{d}$, F) using Eq. (5). Applying it to $\mathbf{c} = [1, 1, 0, 1, 0]$ and $\mathbf{d} = [0, 0, 1, 1, 1]$, we obtain $\mathbf{x'} = [1, 1, 0, 0, 0]$.

If the resulting fitness does not improve, Eqs. (2) and (3) are applied to adjust F and CR.

Step 2: Gene Insertion. We then compute GENEINSERTION ($\mathbf{b}$, $\mathbf{x}'$, F) using Eq. (4). Applying it to $\mathbf{b} = [0, 1, 1, 0, 0]$ and $\mathbf{x'} = [1, 1, 0, 0, 0]$, we obtain CBDE ($\mathbf{b}$, $\mathbf{c}$, $\mathbf{d}$) = [1, 1, 1, 0, 0].

Interpretation: The resulting solution selects genes 1, 2, and 3. Therefore, the gene cluster constructed is:

$$\begin{pmatrix} 42 & 15\,36 & 71\,24 \\ 33 & 58\,17 & 39\,65 \\ 13 & 45\,92 & 12\,26 \end{pmatrix}.$$

This cluster is then evaluated using the fitness function which combines the number of selected genes and the MSR.

Fitness evaluation

The algorithm evaluates the quality of each solution using two objective functions: the size of the selected features and the MSR, aiming to guide the search toward solutions that balance minimal feature size with high classification accuracy.

In this study, the MSR [18] is used as the second objective function to assess the coherence of gene clusters

[1]Available at: https://doi.org/10.5281/zenodo.17201993 (accessed: 16.11.2025).

by measuring the similarity of gene expression patterns across all samples; a lower MSR indicates more consistent and biologically relevant gene groups. The MSR for a cluster C that consists of I rows and J columns is defined as follows:

$$MSR(C) = \frac{1}{|I| \times |J|} \sum_{i=1}^{|I|} \sum_{j=1}^{|J|} (c_{ij} - c_{iJ} - c_{Ij} + c_{IJ})^2, \quad (6)$$

where c_{ij} is the expression value of gene i under condition j; $c_{iJ} = \frac{1}{|J|} \sum_{j \in J} c_{ij}$ is the average expression of gene i across all conditions; $c_{Ij} = \frac{1}{|I|} \sum_{i \in I} c_{ij}$ is the average expression of all genes under condition j; $c_{IJ} = \frac{1}{|I| \times |J|} \sum_{i \in I} \sum_{j \in J} c_{ij}$ is the overall average expression within the cluster.

Example

Consider a gene cluster C composed of 2 genes and their expression values across 3 samples (conditions) as follows:

$$C = \begin{pmatrix} 10 & 15 & 30 \\ 25 & 20 & 40 \end{pmatrix}.$$

We compute the MSR as defined in Eq. (6). For example, when $i = 1, j = 1$, we have:

$$c_{11} = 10, c_{IJ} \approx 23.333,$$
$$c_{iJ} \approx 18.333, c_{Ij} = 17.5.$$

We then substitute these into the MSR equation and iterate for all values in the matrix:

$$MSR(C) = \frac{1}{6} \sum_{i=1}^{2} \sum_{j=1}^{3} (c_{ij} - c_{iJ} - c_{Ij} + c_{IJ})^2.$$

After plugging in all values and performing the calculations: $MSR(C) \approx 22.22$.

Selection process

After evaluating fitness, the algorithm chooses the best individuals to create the next generation based on dominance (NSGA-II), ensures diversity and convergence.

Iteration and convergence

The process is repeated (mutation, crossover, fitness evaluation, and selection) the predetermined number of iterations or until the convergence criteria are completed.

Experiments and Evaluation

Experimental Settings

To get the best results, the study suggested DeFs-CBDE feature selection method optimal parameters were chosen by trial and error. The parameter settings for the suggested DeFs-CBDE approach on each experimental dataset are displayed in Table 1.

Description of considered datasets

In this work, we evaluated the effectiveness of the suggested DeFs-CBDE feature selection strategy using four high-dimensional carcinogenic microarray datasets. These four datasets include brain cancer [19], breast cancer [20], lung cancer [21], and CNS [21], as illustrated in Table 2. We focused on these four widely adopted benchmark datasets because they are well-curated, publicly available, and have been extensively used in previous studies on feature selection and classification of gene expression data. This choice allows us to ensure fair comparison with prior work and to validate our method under standard experimental conditions. However, we acknowledge that relying only on these datasets may limit the generalizability of the conclusions. Therefore, as discussed in the Conclusion and Future Work, we plan to extend the evaluation to larger and more diverse datasets, to confirm the robustness of DeFs-CBDE across different cancer types and experimental platforms.

Table 1. Hyperparameters used in the proposed DeFs-CBDE

Hyper Parameter	Value
Initial F (Factor)	0.65
Crossover rate	0.6
Number of generations	50
Number of Individuals	20

The 5-fold cross-validation method is employed to divide the dataset into training and testing sets. Each algorithm is executed independently 20 times to ensure statistical reliability. All experiments were conducted on a Windows 10 PC with a 2.40 GHz Xeon E5620 8-core processor and 40 GB of RAM, using the parameter settings shown in Table 1.

Experimental results and discussion

Evaluation of the suggested DeFs-CBDE Performance

The performance of machine learning models using features selected by the proposed DeFs-CBDE method was compared to the FS techniques presented in [6]. Fig. 3 shows the classification results using all features, features selected by individual filter methods (Chi-squared (CS), Information Gain (IG), and Gain Ratio (IGR)), hybrid filter-GA methods (CS-GA, IG-GA, and IGR-GA) from [6], and those selected by DeFs-CBDE.

For the brain dataset, as shown in Fig. 3, *a*, the DeFs-CBDE method shows superior performance across all classifiers, achieving 100 % in accuracy, recall, precision, and F-measure with SVM, K-Nearest Neighbors (KNN), and RF. It remains competitive with NB (85.71 %) and matches or exceeds other methods with DT. These results confirm its robustness and effectiveness in FS.

Table 2. Description of the high-dimensional microarray datasets used in this study

Dataset	#Features	#Samples	#Classes
Brain	5,597	42	5
Breast	24,481	97	2
Lung	12,600	203	5
CNS	7,129	60	2

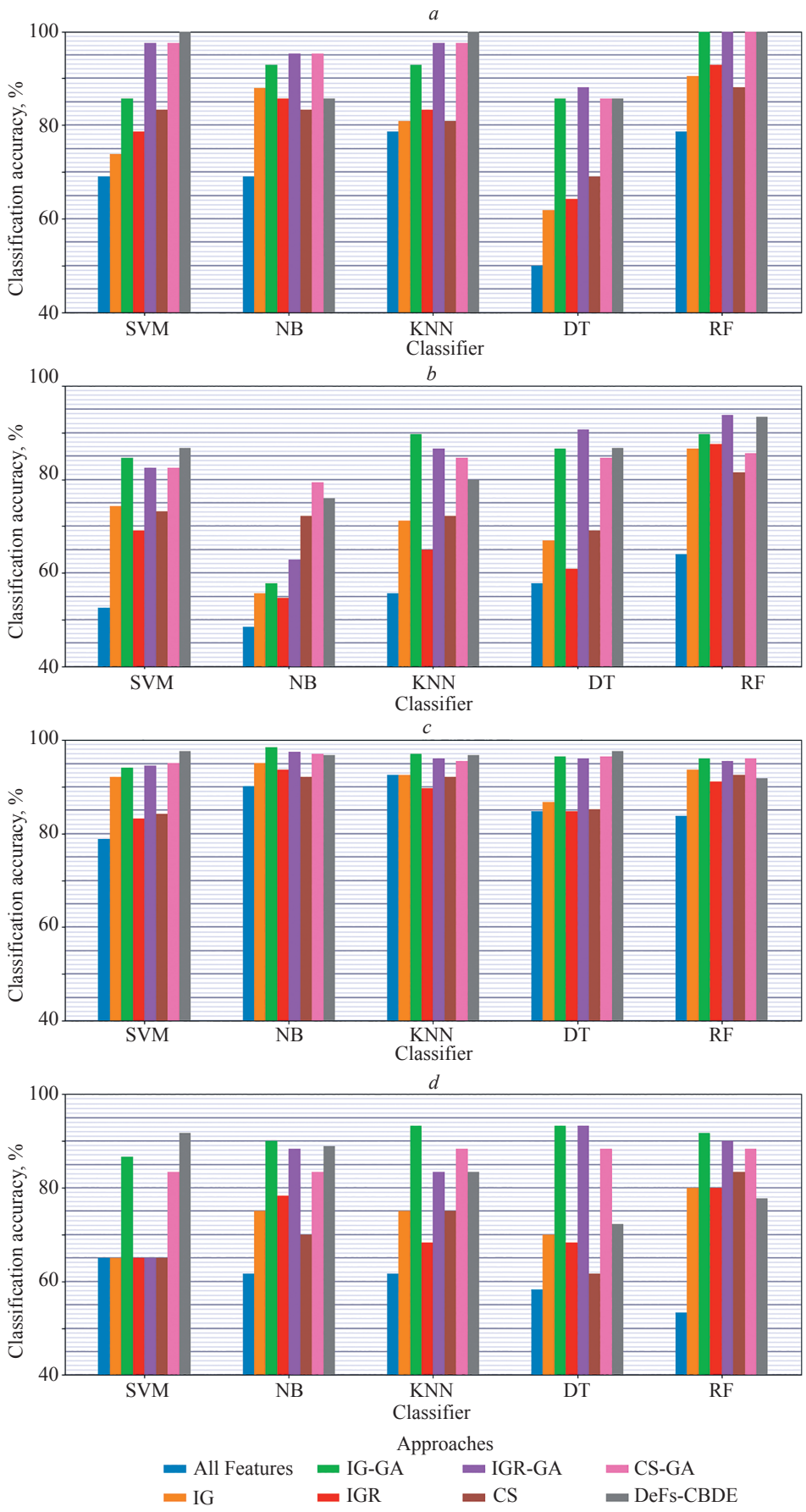

Fig. 3. Comparison of classifier accuracies on the data sets: Brain (*a*); Breast (*b*); Lung (*c*); CNS (*d*) for DeFs-CBDE method

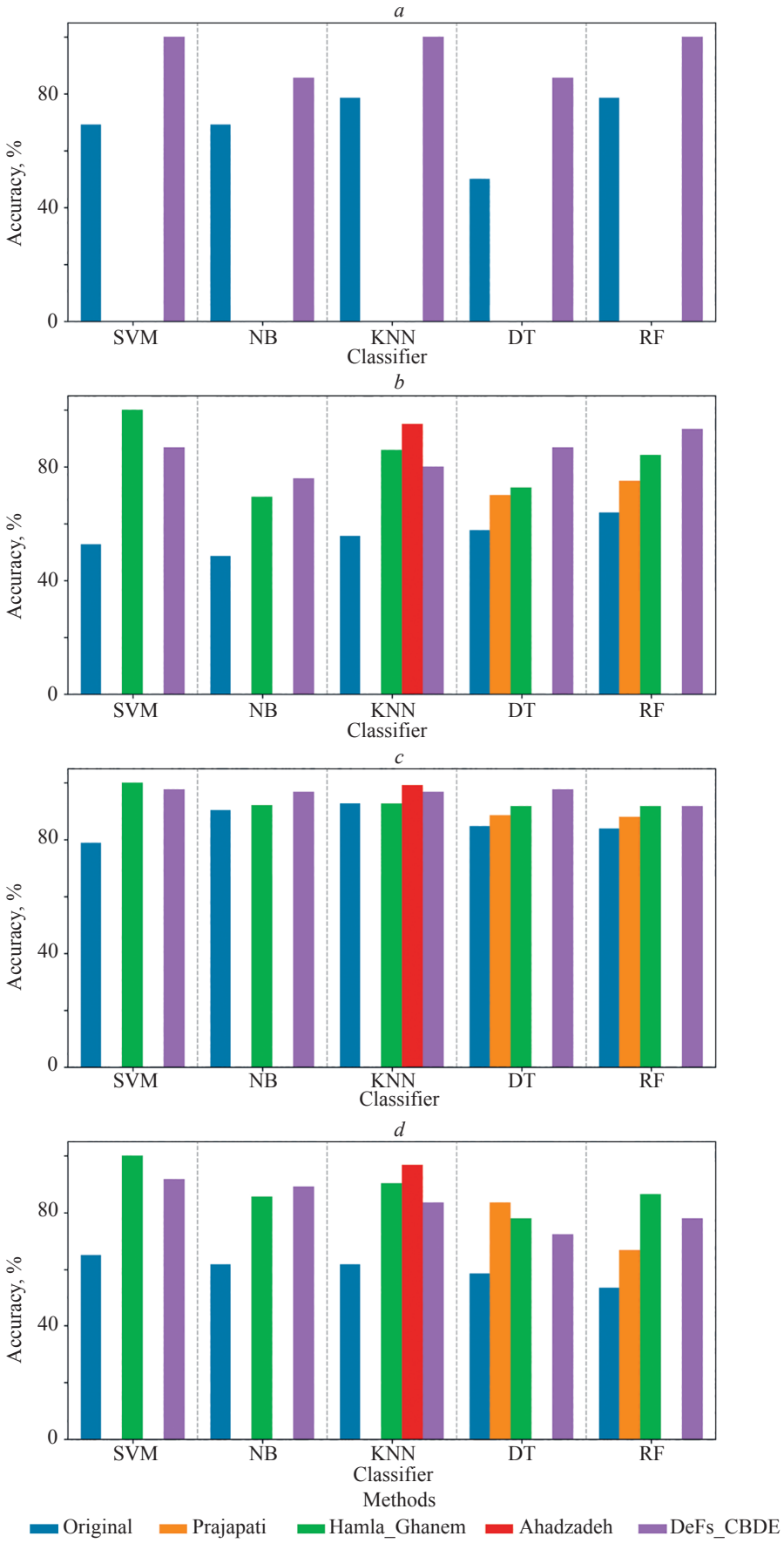

Fig. 4. Comparison of classifier accuracies on the data sets: Brain (*a*); Breast (*b*); Lung (*c*); CNS (*d*) with existing works

For the Breast dataset, as shown in Fig. 3, *b*, the DeFs-CBDE method shows consistently strong performance across classifiers. It outperforms other methods with SVM (86.67 %) and RF (93.33 %), and delivers stable results with NB (76 %) and DT (86.67 %). Although slightly below IGR-GA with KNN, it remains competitive, confirming its robustness in FS.

For the Lung dataset, as shown in Fig. 3, *c*, DeFs-CBDE shows strong and stable performance across classifiers, achieving top accuracy with SVM and DT (97.56 %) and solid results with NB and KNN (96.72 %). While RF reaches 91.80 %, slightly below other methods, the approach remains competitive overall, confirming its reliability and versatility in FS.

For the CNS dataset, as illustrated in Fig. 3, *d*, DeFs-CBDE demonstrates competitive performance across all classifiers. It achieves 91.67 % accuracy with SVM and 88.89 % with NB, outperforming or matching most existing methods. While slightly lower with KNN, DT, and RF, the method maintains balanced and reliable metrics, confirming its robustness in FS.

Comparison of DeFs-CBDE with existing works

In addition to the main comparison with several state of the art FS algorithms in [6] reported in the previous section, we have compared the proposed method to other existing works [7–9]. Across all four datasets, the proposed DeFs-CBDE approach achieves notably superior accuracies. On the Brain dataset, it reaches 100 % with SVM, KNN, and RF, exceeding by far the baseline of 78.57 %. For the Breast dataset, it achieves 93.33 % with RF, outperforming the best competing result (84.09 % by Hamla and Ghanem). On the Lung dataset, DeFs-CBDE obtains 97.56 % with SVM and DT, very close to the top scores while maintaining consistency. Finally, on the CNS dataset, it reaches 91.67 % with SVM, compared to 65 % without FS, confirming its effectiveness across diverse data. A comprehensive comparison of classification accuracies between DeFs-CBDE and the methods reported in [7–9] is presented in Fig. 4.

The initial data for the diagrams (Fig. 4) are provided in link[1].

Complexity of computation

The computational cost of the DeFs-CBDE method can be approximated based on its main parameters, as summarized below. *T*: Number of iterations, *p*: Population size, *n*: Number of genes (features), *m*: Number of samples, *s*: Number of selected features, *Tfitness*: Time complexity of computing the fitness function (*f*1,*f*2).

The overall time complexity is:

$$O(T \times (p \times (n + Tfitness)) + p^2).$$

According to this analysis, the main factors influencing computational effort are the population size (*p*), the number of iterations (*T*), and the number of genes (*n*). By striking an effective balance among these variables, our method can handle high-dimensional data at a computationally affordable cost.

Biological Relevance of Selected Features (FS) — Lung Dataset

Fig. 4, *a* shows the functional enrichment results obtained from the genes selected by the FS method applied to the Lung dataset. The analysis, performed with the g:Profiler tool [22], highlights a strong biological significance of the retained genes. Indeed, among the selected features, about 70 % are mapped to significantly enriched pathways (p-value < 0.05 after multiple testing correction). The most enriched categories include cell cycle regulation, cell proliferation, and immune response which are all closely related to lung cancer mechanisms. Additionally, several enriched terms from Gene Ontology (GO) were identified, particularly within *Biological Process* and *Molecular Function*. For instance, genes involved in apoptosis regulation and intracellular signaling pathways are significantly represented (adjusted p-value ≈ 10^{-5}). These findings demonstrate that the FS-selected genes are not random but rather capture key biological mechanisms of lung cancer, thus providing strong biological validation for the feature selection process.

The initial data for the diagrams (Fig. 5) are provided in link[2].

[1] Available at: https://doi.org/10.5281/zenodo.17201993 (accessed: 16.11.2025).

[2] Available at: https://doi.org/10.5281/zenodo.17201993 (accessed: 16.11.2025).

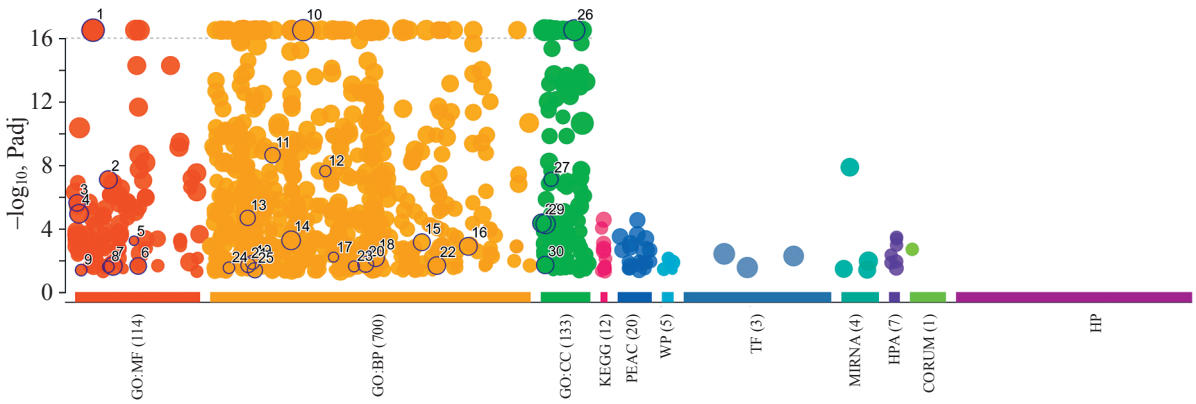

Fig. 5. Functional enrichment analysis of FS-selected genes on the Lung dataset using, g:Profiler

We now provide concrete examples of top-ranked genes selected by CBDE and highlight their known involvement in lung cancer. For example:

— Epidermal Growth Factor Receptor (EGFR): frequently mutated in non-small cell lung cancer (NSCLC) and a target of tyrosine kinase inhibitors [23].
— Kirsten Rat Sarcoma Viral Oncogene: one of the most common driver mutations in lung adenocarcinoma, associated with poor prognosis [24].
— Tumor Protein p53: the most frequently mutated tumor suppressor gene in lung cancer, playing a central role in cell cycle regulation and apoptosis [24].
— Anaplastic Lymphoma Kinase (ALK): gene rearrangements involving ALK are important biomarkers guiding targeted therapy in lung adenocarcinoma [23].
— Mucin 1: overexpressed in lung cancer and associated with tumor progression and metastasis [25].

Furthermore, the functional enrichment analysis performed on the broader set of genes selected by CBDE (e.g., pathways such as 'p53 signaling pathway' and 'EGFR tyrosine kinase inhibitor resistance') provides additional statistical support that our method captures not only individual key players but also biologically coherent pathways and processes critically involved in lung carcinogenesis.

Conclusion

This paper presents a new and effective technique for using Multi-Objective Differential Evolution as a feature selection strategy to find high-quality features. A novel mutation operator, called Clustering-Based Binary Differential Evolution (CBDE), which blends gene insertion and removal, is introduced. This allows for better solutions. Throughout the experimentation phase, datasets on the brain cancer, breast cancer, lung cancer, and the Central Nervous System were used to illustrate the benefits of our strategy over current approaches. The outcomes were remarkable, and the chosen features showed a great deal of variability.

We used five classifiers Support Vector Machine, Naive Bayes, K-Nearest Neighbors, Decision Tree and Random Forest across several datasets to evaluate the effect of DeFs-CBDE. We compared the proposed method with three feature selection algorithms: Chi-Squared (CS), Information Gain (IG), and Gain Ratio (IGR), as well as three hybrid filter-GA feature selection techniques: IG-GA, IGR-GA, and CS-GA.

The proposed algorithm was used on high-dimensional datasets described in the section 'Description of Considered Datasets' to improve the performance of Machine Learning methods by eliminating irrelevant features and retaining only the most relevant subsets. Experiments showed that reducing the dimensionality up to 50 % has improved Machine Learning performance. Despite its effectiveness demonstrated in the results section, the proposed method has some limitations. First, the CBDE increases operator exploration to escape local optima. However, this increases the calculation burden and the search time specifically for large datasets with high dimensionality. Second, the algorithm parameters require careful tuning that may not generalize well across all types of datasets. Additionally, adding new unseen data may require re-running the selection algorithm.

Future work will focus on optimizing the computational efficiency and robustness of DeFs-CBDE by exploring hybrid and adaptive approaches. In particular, we plan to incorporate recent metaheuristic optimization techniques, such as the Whale Optimization Algorithm, Shark Smell Optimization, and Grey Wolf Optimizer, which have shown strong performance in high-dimensional and nonlinear optimization problems. In addition, we intend to evaluate DeFs-CBDE on a wider range of datasets, to further demonstrate its generalizability across different cancer types and experimental platforms. Finally, we plan to extend this line of research with a formal ablation analysis to better quantify the contribution of each component of the CBDE operator.

References

1. Trunk G.V. A problem of dimensionality: A simple example. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1979, vol. PAMI-1, no. 3, pp. 306–307. https://doi.org/10.1109/tpami.1979.4766926
2. Saberi-Movahed F., Rostami M., Berahmand K., Karami S., Tiwari P., Oussalah M., Band S.S. Dual regularized unsupervised feature selection based on matrix factorization and minimum redundancy with application in gene selection. *Knowledge-Based Systems*, 2022, vol. 256 pp. 109884. https://doi.org/10.1016/j.knosys.2022.109884
3. Deb K. Multi-objective optimisation using evolutionary algorithms: an introduction. *Multi-objective Evolutionary Optimisation for Product Design and Manufacturing*, 2011, pp. 3–34.
4. Cheng Y., Church G.M. Biclustering of expression data. *Proc. of the 8th International Conference on Intelligent Systems for Molecular Biology*, 2000, pp. 93–103.
5. Zhang Y., Gong D., Gao X., Tian T., Sun X. Binary differential evolution with self-learning for multi-objective feature selection. *Information Sciences*, 2020, vol. 507, pp. 67–85. https://doi.org/10.1016/j.ins.2019.08.040
6. Ali W., Saeed F. Hybrid filter and genetic algorithm-based feature selection for improving cancer classification in high-dimensional

Литература

1. Trunk G.V. A problem of dimensionality: A simple example // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1979. V. PAMI-1. N 3. P. 306–307. https://doi.org/10.1109/tpami.1979.4766926
2. Saberi-Movahed F., Rostami M., Berahmand K., Karami S., Tiwari P., Oussalah M., Band S.S. Dual regularized unsupervised feature selection based on matrix factorization and minimum redundancy with application in gene selection // Knowledge-Based Systems. 2022. V. 256. P. 109884. https://doi.org/10.1016/j.knosys.2022.109884
3. Deb K. Multi-objective optimisation using evolutionary algorithms: an introduction // Multi-objective Evolutionary Optimisation for Product Design and Manufacturing. 2011. P. 3–34.
4. Cheng Y., Church G.M. Biclustering of expression data // Proc. of the 8th International Conference on Intelligent Systems for Molecular Biology. 2000. P. 93–103.
5. Zhang Y., Gong D., Gao X., Tian T., Sun X. Binary differential evolution with self-learning for multi-objective feature selection // Information Sciences. 2020. V. 507. P. 67–85. https://doi.org/10.1016/j.ins.2019.08.040
6. Ali W., Saeed F. Hybrid filter and genetic algorithm-based feature selection for improving cancer classification in high-dimensional

microarray data. *Processes*, 2023, vol. 11, no. 2, p. 562. https://doi.org/10.3390/pr11020562
7. Ahadzadeh B., Abdar M., Safara F., Khosravi A., Menhaj M.B., Suganthan P.N. SFE: A simple, fast, and efficient feature selection algorithm for high-dimensional data. *IEEE Transactions on Evolutionary Computation*, 2023, vol. 27, no. 6, pp. 1896–1911. https://doi.org/10.1109/tevc.2023.3238420
8. Prajapati S., Das H., Gourisaria M.K. Feature selection using differential evolution for microarray data classification. *Discover Internet of Things*, 2023. vol. 3, no. 1, p. 12. https://doi.org/10.1007/s43926-023-00042-5
9. Hamla H., Ghanem K. A hybrid feature selection based on fisher score and svm-rfe for microarray data. *Informatica*, 2024, vol. 48, no. 1, pp. 57–68. https://doi.org/10.31449/inf.v48i1.4759
10. Paul D., Jain A., Saha S., Mathew J. Multi-objective pso based online feature selection for multi-label classification. *Knowledge-Based Systems*, 2021, vol. 222, pp. 106966, https://doi.org/10.1016/j.knosys.2021.106966
11. Han F., Chen W.-T., Ling Q.-H., Han H. Multi-objective particle swarm optimization with adaptive strategies for feature selection. *Swarm and Evolutionary Computation*, 2021, vol. 62, pp. 100847. https://doi.org/10.1016/j.swevo.2021.100847
12. Dashtban M., Balafar M., Suravajhala P. Gene selection for tumor classification using a novel bio-inspired multi-objective approach. *Genomics*, 2018, vol. 110, no. 1, pp. 10–17. https://doi.org/10.1016/j.ygeno.2017.07.010
13. Ghosh M., Adhikary S., Ghosh K.K., Sardar A., Begum S., Sarkar R. Genetic algorithm based cancerous gene identification from microarray data using ensemble of filter methods. *Medical and Biological Engineering and Computing*, 2019, vol. 57, no. 1, pp. 159–176. https://doi.org/10.1007/s11517-018-1874-4
14. Storn R., Price K. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces. *International Computer Science Institute*, 1995, vol. 95, no. 12, pp. 1–12.
15. Charfaoui Y., Houari A., Boufera F. AMoDeBic: An adaptive multi-objective differential evolution biclustering algorithm of microarray data using a biclustering binary mutation operator. *Expert Systems with Applications*, 2024, vol. 238, part B, pp. 121863, https://doi.org/10.1016/j.eswa.2023.121863
16. Goldberg D.E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989, 412 p.
17. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 2002, vol. 6, no. 2, pp. 182–197. https://doi.org/10.1109/4235.996017
18. Noronha M.D., Henriques R., Madeira S.C., Zárate L.E. Impact of metrics on biclustering solution and quality: a review. *Pattern Recognition*, 2022, vol. 127, pp. 108612. https://doi.org/10.1016/j.patcog.2022.108612
19. Pomeroy S.L., Tamayo P., Gaasenbeek M., Sturla L.M., Angelo M., McLaughlin M.E., et al. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, 2002, vol. 415, no. 6870, pp. 436–442. https://doi.org/10.1038/415436a
20. van't Veer L.J., Dai H., van de Vijver M.J., He Y.D., Hart A.A.M., Mao M., et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*, 2002, vol. 415, no. 6871, pp. 530–536. https://doi.org/10.1038/415530a
21. Zhao G., Wu Y. Feature subset selection for cancer classification using weight local modularity. *Scientific Reports*, 2016, vol. 6, pp. 34759. https://doi.org/10.1038/srep34759
22. Kolberg L., Raudvere U., Kuzmin I., Adler P., Vilo J., Peterson H. g: Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Research*, 2023, vol. 51, no. W1, pp. W207–W212. https://doi.org/10.1093/nar/gkad347
23. Herbst R.S., Morgensztern D., Boshoff C. The biology and management of non-small cell lung cancer. *Nature*, 2018, vol. 553, no. 7689, pp. 446–454. https://doi.org/10.1038/nature25183
24. Skoulidis F., Heymach J.V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. *Nature Reviews Cancer*, 2019, vol. 19, no. 9, pp. 495–509. https://doi.org/10.1038/s41568-019-0179-8
25. Nath S., Mukherjee P. Muc1: a multifaceted oncoprotein with a key role in cancer progression. *Trends in Molecular Medicine*, 2014, vol. 20, no. 6, pp. 332–342. https://doi.org/10.1016/j.molmed.2014.02.007

microarray data // Processes. 2023. V. 11. N 2. P. 562. https://doi.org/10.3390/pr11020562
7. Ahadzadeh B., Abdar M., Safara F., Khosravi A., Menhaj M.B., Suganthan P.N. SFE: A simple, fast, and efficient feature selection algorithm for high-dimensional data // IEEE Transactions on Evolutionary Computation. 2023. V. 27. N 6. P. 1896–1911. https://doi.org/10.1109/tevc.2023.3238420
8. Prajapati S., Das H., Gourisaria M.K. Feature selection using differential evolution for microarray data classification // Discover Internet of Things. 2023. V. 3. N 1. P. 12. https://doi.org/10.1007/s43926-023-00042-5
9. Hamla H., Ghanem K. A hybrid feature selection based on fisher score and svm-rfe for microarray data // Informatica. 2024. V. 48. N 1. P. 57–68. https://doi.org/10.31449/inf.v48i1.4759
10. Paul D., Jain A., Saha S., Mathew J. Multi-objective pso based online feature selection for multi-label classification // Knowledge-Based Systems. 2021. V. 222. P. 106966, https://doi.org/10.1016/j.knosys.2021.106966
11. Han F., Chen W.-T., Ling Q.-H., Han H. Multi-objective particle swarm optimization with adaptive strategies for feature selection // Swarm and Evolutionary Computation. 2021. V. 62. P. 100847. https://doi.org/10.1016/j.swevo.2021.100847
12. Dashtban M., Balafar M., Suravajhala P. Gene selection for tumor classification using a novel bio-inspired multi-objective approach // Genomics. 2018. V. 110. N 1. P. 10–17. https://doi.org/10.1016/j.ygeno.2017.07.010
13. Ghosh M., Adhikary S., Ghosh K.K., Sardar A., Begum S., Sarkar R. Genetic algorithm based cancerous gene identification from microarray data using ensemble of filter methods // Medical and Biological Engineering and Computing. 2019. V. 57. N 1. P. 159–176. https://doi.org/10.1007/s11517-018-1874-4
14. Storn R., Price K. Differential evolution: a simple and efficient adaptive scheme for global optimization over continuous spaces // International Computer Science Institute. 1995. V. 95. N 12. P. 1–12.
15. Charfaoui Y., Houari A., Boufera F. AMoDeBic: An adaptive multi-objective differential evolution biclustering algorithm of microarray data using a biclustering binary mutation operator // Expert Systems with Applications. 2024. V. 238. Part B. P. 121863, https://doi.org/10.1016/j.eswa.2023.121863
16. Goldberg D.E. Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley, 1989. 412 p.
17. Deb K., Pratap A., Agarwal S., Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II // IEEE Transactions on Evolutionary Computation. 2002. V. 6. N 2. P. 182–197. https://doi.org/10.1109/4235.996017
18. Noronha M.D., Henriques R., Madeira S.C., Zárate L.E. Impact of metrics on biclustering solution and quality: a review // Pattern Recognition. 2022. V. 127. P. 108612. https://doi.org/10.1016/j.patcog.2022.108612
19. Pomeroy S.L., Tamayo P., Gaasenbeek M., Sturla L.M., Angelo M., McLaughlin M.E., et al. Prediction of central nervous system embryonal tumour outcome based on gene expression // Nature. 2002. V. 415. N 6870. P. 436–442. https://doi.org/10.1038/415436a
20. van't Veer L.J., Dai H., van de Vijver M.J., He Y.D., Hart A.A.M., Mao M., et al. Gene expression profiling predicts clinical outcome of breast cancer // Nature. 2002. V. 415. N 6871. P. 530–536. https://doi.org/10.1038/415530a
21. Zhao G., Wu Y. Feature subset selection for cancer classification using weight local modularity // Scientific Reports. 2016. V. 6. P. 34759. https://doi.org/10.1038/srep34759
22. Kolberg L., Raudvere U., Kuzmin I., Adler P., Vilo J., Peterson H. g: Profiler—interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update) // Nucleic Acids Research. 2023. V. 51. N W1. P. W207–W212. https://doi.org/10.1093/nar/gkad347
23. Herbst R.S., Morgensztern D., Boshoff C. The biology and management of non-small cell lung cancer // Nature. 2018. V. 553. N 7689. P. 446–454. https://doi.org/10.1038/nature25183
24. Skoulidis F., Heymach J.V. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy // Nature Reviews Cancer. 2019. V. 19. N 9. P. 495–509. https://doi.org/10.1038/s41568-019-0179-8
25. Nath S., Mukherjee P. Muc1: a multifaceted oncoprotein with a key role in cancer progression // Trends in Molecular Medicine. 2014. V. 20. N 6. P. 332–342. https://doi.org/10.1016/j.molmed.2014.02.007

Authors

Mohamed Djellal Serandi — PhD Student, University Mustapha Stambouli, LISYS laboratory, Mascara, 29000, Algeria, https://orcid.org/0009-0009-5775-7956, mohamed.djellalserandi@univ-mascara.dz
Fatma Boufera – PhD, Full Professor, University Mustapha Stambouli, LISYS laboratory, Mascara, 29000, Algeria, https://orcid.org/0000-0002-5733-586X, fboufera@univ-mascara.dz

Amina Houari — PhD, Associate Professor, University Mustapha Stambouli, LISYS laboratory, Mascara, 29000, Algeria, sc 57021480300, https://orcid.org/0000-0002-3628-7483, amina.houari@univ-mascara.dz
Farid Flitti — PhD, Associate Professor, Higher Colleges of Technology in Dubai, Dubai, 500001, United Arab Emirates, sc 24821958500, https://orcid.org/0000-0002-2480-2580, fflitti@hct.ac.ae

Авторы

Джеллаль Серанди Мохамед — аспирант, Университет Мустафы Стамбули, лаборатория LISYS, Маскара, 29000, Алжир, https://orcid.org/0009-0009-5775-7956, djellalserandi@univ-mascara.dz
Буфера Фатма — кандидат технических наук, профессор, профессор, Университет Мустафы Стамбули, лаборатория LISYS, Маскара, 29000, Алжир, https://orcid.org/0000-0002-5733-586X, fboufera@univ-mascara.dz
Хуари Амина — PhD, доцент, Университет Мустафы Стамбули, лаборатория LISYS, Маскара, 29000, Алжир, sc 57021480300, https://orcid.org/0000-0002-3628-7483, amina.houari@univ-mascara.dz
Флитти Фарид — PhD, доцент, Высший колледж технологии, колледж в Дубае, Дубай, 500001, Объединенные Арабские Эмираты, sc 24821958500, https://orcid.org/0000-0002-2480-2580, fflitti@hct.ac.ae

Received 16.07.2025
Approved after reviewing 30.09.2025
Accepted 21.11.2025

Статья поступила в редакцию 16.07.2025
Одобрена после рецензирования 30.09.2025
Принята к печати 21.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1197-1207
УДК 620.178.152.341.4, 004.94, 303.724.32, 004.032.26

Определение оптимального метода машинного обучения для построения прогнозных моделей микротвердости по Виккерсу керамического композитного материала гидроксиапатит-многостенные углеродные нанотрубки

Анастасия Евгеньевна Резванова[1]✉, Борис Сергеевич Кудряшов[2], Владимир Юрьевич Погудин[3]

[1,2,3] Институт физики прочности и материаловедения Сибирского отделения Российской академии наук, Томск, 634055, Российская Федерация
[3] Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация

[1] ranast@ispms.ru✉, https://orcid.org/0000-0002-7067-7979
[2] bsk3@ispms.ru, https://orcid.org/0009-0000-5133-4893
[3] pogudin.vova@bk.ru, https://orcid.org/0009-0002-7191-163X

Аннотация
Введение. Долговечность и износостойкость керамических имплантатов, используемых в условиях высоких эксплуатационных нагрузок, в значительной степени зависят от их механических характеристик. Керамический композит на основе гидроксиапатита рассматривается как перспективный биоматериал для реконструкции поврежденных костных тканей и замещения костных дефектов благодаря своей биосовместимости и способности обеспечивать остеоинтеграцию с костной тканью. Для повышения механической прочности предлагается армировать керамику на основе гидроксиапатита многостенными углеродными нанотрубками, обладающими высокими физико-механическими характеристиками. Данный подход направлен на применение материалов в зонах имплантации, испытывающих значительные механические нагрузки. Эффективность армирования нанотрубками во многом зависит от состава композита, технологии синтеза и условий испытаний, что приводит к высокой вариативности итоговых характеристик. Прямое экспериментальное определение свойств каждого образца требует значительных временных затрат. Для оптимизации процесса исследования механических характеристик композитных материалов представляется актуальным использование математических моделей, основанных на методах машинного обучения, что позволяет прогнозировать микротвердость по Виккерсу в зависимости от нагрузки индентирования. **Метод.** Экспериментальные испытания на микротвердость по Виккерсу для шести серий керамических образцов проводились методом индентирования при нагрузках в диапазоне от 0,98 Н до 9,8 Н. Для прогнозирования полученных данных были применены три метода машинного обучения: нейронная сеть, метод случайного леса и градиентный бустинг. **Основные результаты.** В результате исследования после усреднения значений по всем нагрузкам индентирования для каждого образца определено, что с увеличением концентрации многостенных углеродных нанотрубок до 0,5 масс.% микротвердость композита по сравнению с гидроксиапатитом без добавок возрастает от 3,83 ± 0,39 ГПа до 4,71 ± 0,40 ГПа, а армирование становится эффективным на 19 %. Таким образом, наибольший вклад в повышение микротвердости композита внесли добавки с концентрацией 0,5 масс.%, при этом добавление 1 и 2 масс.% привело к значительному снижению микротвердости, что связано с возникновением агломерации нанотрубок в керамической матрице. **Обсуждение.** Результаты моделирования позволили на основе данных экспериментального исследования определить оптимальный метод машинного обучения для построения прогнозной модели микротвердости композитной керамики гидроксиапатит-многостенные углеродные нанотрубки в широком диапазоне нагрузок, а также установить взаимосвязь между составом композита и его механическими характеристиками, что открывает новые возможности для проектирования прочных и долговечных керамических имплантатов.

Ключевые слова
гидроксиапатит, многостенные углеродные нанотрубки, микротвердость по Виккерсу, прогнозирование, методы регрессии, машинное обучение

Благодарности

© Резванова А.Е., Кудряшов Б.С., Погудин В.Ю., 2025

Исследование выполнено за счет гранта Российского научного фонда № 25-23-20233 (https://rscf.ru/project/25-23-20233/) и гранта в форме субсидии, выделяемого Департаментом по научно-технологическому развитию и инновационной деятельности Томской области (Соглашение № 02/5/2025).
Исследования выполнены с использованием оборудования Центра коллективного пользования «Нанотех» Института физики прочности и материаловедения Сибирского отделения РАН.

Ссылка для цитирования: Резванова А.Е., Кудряшов Б.С., Погудин В.Ю. Определение оптимального метода машинного обучения для построения прогнозных моделей микротвердости по Виккерсу керамического композитного материала гидроксиапатит-многостенные углеродные нанотрубки // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1197–1207. doi: 10.17586/2226-1494-2025-25-6-1197-1207

Machine-learning method for the development of a Vickers microhardness predictive model of a hydroxyapatite-multi-walled carbon nanotube ceramic composite material

Anastasiya E. Rezvanova[1]✉, Boris S. Kudryashov[2], Vladimir Yu. Pogudin[3]

[1,2,3] Institute of Strength Physics and Materials Science of the Siberian Branch of the Russian Academy of Sciences, Tomsk, 634055, Russian Federation
[3] Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation

[1] ranast@ispms.ru✉, https://orcid.org/0000-0002-7067-7979
[2] bsk3@ispms.ru, https://orcid.org/0009-0000-5133-4893
[3] pogudin.vova@bk.ru, https://orcid.org/0009-0002-7191-163X

Abstract
The durability and wear resistance of ceramic implants, used under high operating loads largely, depend on their mechanical characteristics. Ceramic composite based on hydroxyapatite is considered as a promising biomaterial for reconstruction of damaged bone tissues and replacement bone defects due to its biocompatibility and ability to provide osseointegration with bone tissue. In this work, to increase mechanical strength, the hydroxyapatite ceramics was reinforced by multi-walled carbon nanotubes additives which have high physical and mechanical properties. The potential of such research lies in the use of these composites in implantation areas that experience significant mechanical loads. The effectiveness of nanotubes reinforcement depends on the ceramics composition, synthesis technology, and testing conditions, resulting in high variability in the final characteristics. At the same time, direct experimental study of the properties of each sample requires significant time-cost. The use of mathematical models based on machine learning methods to optimize the process of analysis the mechanical characteristics of composite materials is relevant study. This will allow us to predict Vickers microhardness depending on the indentation load. In this study, experimental microhardness tests were carried out by using Vickers method for six sets of ceramic materials which were exposed to indentation loads ranging from 0.98 N to 9.8 N. Three machine learning methods were used to predict the data obtained: neural network, random forest and gradient boosting. After averaging the values for all loads, it was determined that with the increasing the concentration of multi-walled carbon nanotubes to 0.5 wt. % the Vickers microhardness of the composite increases from 3.83 ± 0.39 GPa to 4.71 ± 0.40 GPa compared to hydroxyapatite without additives, and reinforcement becomes effective by 19 %. Thus, the greatest contribution to the increase in the microhardness of the composite was made by the additives of multi-walled carbon nanotubes with a concentration of 0.5 wt.%, while the addition of 1 and 2 wt.% led to a significant decrease in microhardness, which is associated with the appearance of multi-walled carbon nanotubes agglomeration in the ceramic hydroxyapatite matrix. The simulation results based on experimental data allowed us to determine the optimal machine learning method for constructing a predictive model of the microhardness of hydroxyapatite - multi-walled carbon nanotubes ceramic composite in a wide range of loads. In addition, it was possible to establish the relationship between the composition of the composite and its mechanical characteristics, which opens up new possibilities for designing strong and durable ceramic implants.

Keywords
hydroxyapatite, multi-walled carbon nanotubes, Vickers microhardness, prediction, regression methods, machine learning

Acknowledgements
This research was carried out with support from the Russian Science Foundation grant No. 25-23-20233 (https://rscf.ru/project/25-23-20233/) and a subsidy grant allocated by the Department of Scientific, Technological Development and Innovative Activities of the Tomsk Region (Agreement No. 02/5/2025).
The authors like to express their gratitude towards the management of Core Facility Centre “Nanotech” of the Institute of Strength Physics and Materials Science of the SB RAS for equipment employed in these studies.

For citation: Rezvanova A.E., Kudryashov B.S., Pogudin V.Yu. Machine-learning method for the development of a Vickers microhardness predictive model of a hydroxyapatite-multi-walled carbon nanotube ceramic composite material. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1197–1207 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1197-1207

Введение

В настоящее время ведется активная разработка композитных материалов, применимых в качестве медицинских имплантатов. Их устойчивость к внешним нагрузкам определяется множеством факторов, включая состав, структуру и физико-механические свойства материалов, из которых они изготовлены. Создание

современных композитов для восстановления или замены поврежденной костной ткани требует сочетания биосовместимости и достаточной механической прочности, что традиционно оценивается преимущественно экспериментально с использованием сложного оборудования и значительных ресурсов. Для улучшения эксплуатационных характеристик композитов в их состав вводят различные функциональные добавки. Свойства этих добавок — от химического состава и степени дисперсности до формы, твердости и концентрации — определяют итоговые характеристики материала.

Среди биодеградируемых керамических материалов особый интерес представляет гидроксиапатит (ГА) — фосфат кальция, структура которого близка к строению естественной костной ткани. Благодаря высокой биосовместимости и биоактивности ГА широко применяется в регенерации костной ткани и остеоинтеграции имплантатов [1–4]. Однако низкие, по сравнению с костными тканями, механические свойства ГА ограничивают его применение в зонах, подвергающихся значительным механическим нагрузкам [5–7]. Для повышения прочностных характеристик ГА все чаще комбинируют с различными нанодобавками, среди которых особое внимание привлекают многостенные углеродные нанотрубки (МУНТ). Добавление МУНТ способствует улучшению механической прочности, износостойкости и трещиностойкости композитной керамики, сохраняя ее высокую биоактивность и биосовместимость [8–14].

Одним из ключевых свойств композитов является твердость, которая отражает способность материалов противостоять упругим и пластическим деформациям под внешним воздействием. Для измерения микротвердости малых или тонких образцов широко применяется метод индентирования пирамидой Виккерса (микротвердость по Виккерсу), который, помимо измерения микротвердости, позволяет оценивать равномерность фазового распределения, локальные химические неоднородности и протяженность диффузионных слоев [15–18]. Создание композитных материалов на основе ГА с добавками МУНТ требует всестороннего изучения их физико-механических и химических характеристик. Экспериментальные измерения формируют обширные массивы числовых данных, которые обычно систематизируются в виде таблиц и графиков для дальнейшего анализа. При обработке таких данных традиционные методы аппроксимации часто сталкиваются с рядом ограничений: необходимость заранее выбирать тип подходящей функции, риск недостаточной гибкости при описании сложных нелинейных зависимостей, чувствительность к шуму и пробелам в исходной информации. Эти факторы затрудняют получение надежных зависимостей при большом объеме экспериментальных данных. В результате целесообразно использовать машинное обучение (МО), которое позволяет интерполировать недостающие данные и прогнозировать свойства материалов с учетом сложных взаимосвязей.

Методы машинного обучения в прогнозировании свойств материалов

Методы МО, в частности, искусственные нейронные сети, зарекомендовали себя как универсальные аппроксиматоры, способные точно выявлять нелинейные зависимости и прогнозировать физические и механические свойства материалов на основе ограниченных выборок [19–21]. Применение этих методов позволяет прогнозировать такие параметры, как микротвердость, прочность на разрыв и другие ключевые характеристики композитов [22–24], демонстрируя высокую точность даже при работе с небольшими наборами экспериментальных данных [25, 26].

Эффективность в прогнозировании также была доказана методом градиентного бустинга, с помощью которого предсказан коэффициент интенсивности напряжений пултрузионного композитного материала с погрешностью менее 10 %, эквивалентной экспериментальной ошибке [27]. В работе [28] рассмотрено прогнозирование изгибных свойств клееных деревянных элементов, усиленных полимерами, армированными углеродным волокном. Наилучший прогноз по оценке метрики R^2 был сделан с помощью метода случайного леса, что вносит вклад в предсказание свойств балок разных размеров. В [29] использован метод случайного леса с наилучшей точностью прогнозирования для оценки и ранжирования влияния входных переменных (состав смеси, параметры обработки и условия испытаний) на модуль Юнга, прочность на изгиб и вязкость разрушения сверхвысокотемпературных термопластов, что является подходящим методом для получения надежных прогнозов свойств до разработки теоретических моделей. Другие результаты применения МО показывают, что рекуррентные модели нейронных сетей могут точно предсказывать модуль упругости композитных материалов, армированных углеродными волокнами, достигая значений R^2 до 0,98 [30]. В работе [31] рассмотрено применение набора методов, основанных на анализе данных, для прогнозирования наиболее часто используемых механических свойств стальных труб на сталелитейных заводах, а именно предела текучести, предела прочности при растяжении и твердости. Модель искусственной нейронной сети удовлетворительно справилась с прогнозированием прочности на разрыв со средней ошибкой 3,7 %. В [32] применялся метод ансамблевого обучения для прямого прогнозирования твердости по Виккерсу в зависимости от нагрузки. Полученные модели показали, что твердость керамики или металлов можно точно спрогнозировать при любой заданной нагрузке с помощью тщательно настроенной модели МО. В работе [33] была создана успешная модель МО, в которой применен регрессор градиентного бустинга для прогнозирования твердости и использованы механические свойства твердого тела (модуль объемной упругости, модуль сдвига, модуль Юнга и коэффициент Пуассона) в качестве входных переменных. Модель была обучена на экспериментальной базе данных по твердости по Виккерсу для 143 материалов, что позволило охватить различные типы соединений. В [34] представлена система, основанная

на реальных данных, с применением экстремального градиентного бустинга для прогнозирования твердости керамики из нитрида кремния (Si_3N_4), армированной графеновыми волокнами. Определены параметры, которые особенно влияют на твердость Si_3N_4. Наиболее значимыми из них являются: давление при спекании, время спекания и плотность. Параметры спекания, в частности давление при спекании, температура, время выдержки и технология, максимально влияют на плотность, конечный размер зерен, состав Si_3N_4 и, как следствие, на твердость. Таким образом, работы [19–34] показали, что методы МО при обработке данных керамических и композитных материалов имеют высокую эффективность при прогнозировании механических свойств, которые представляются набором экспериментальных данных со сложными нелинейными зависимостями.

Цель работы — построение компьютерных моделей с помощью методов МО, таких как метод градиентного бустинга, метод случайного леса решений, а также используя нейросетевое моделирование на основе полносвязной нейронной сети, для прогнозирования значений микротвердости композитов на основе ГА с добавками МУНТ в широком диапазоне нагрузок индентирования. Такие модели позволят предсказывать твердость материала и нагрузку индентора без проведения дополнительных экспериментов, что существенно сокращает затраты ресурсов и ускоряет процесс разработки новых биоматериалов.

Объект исследований

Объектом исследований в настоящей работе стали керамические композитные материалы на основе ГА без добавок (серия 1), а также с различным содержанием МУНТ: 0,1; 0,2; 0,5; 1 и 2 масс.% (серии 2–6). Испытания на микротвердость по Виккерсу проводились методом индентирования на микротвердомере Affri-DM8 в диапазоне нагрузок (P) от 0,98 Н до 9,8 Н по 20 индентирований для каждого образца. На рис. 1 показаны зависимости микротвердости по Виккерсу от нагрузки индентирования керамических ГА материалов с различными концентрациями МУНТ. Усредненные значения микротвердости с указанием доверительных интервалов отражены в табл. 1.

По результатам рис. 1 и табл. 1 видно, что наибольшей микротвердостью по Виккерсу по сравнению с другими исследуемыми материалами обладает серия 4

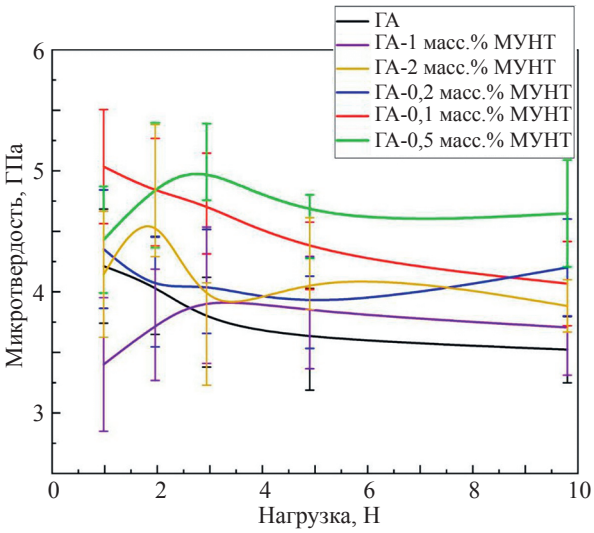

Рис. 1. Зависимость микротвердости по Виккерсу от нагрузки индентирования

Fig. 1. Dependence of Vickers microhardness on indentation load

образцов, имеющая в своем составе 0,5 масс.% МУНТ, что сопоставимо с результатами экспериментальных исследований в работах [35, 36]. При усреднении значений по всему диапазону нагрузок индентирования для каждой серии образцов установлено, что с увеличением нагрузки средняя твердость для данных материалов возрастает от 3,83 ± 0,39 ГПа до 4,71 ± 0,40 ГПа, увеличивая эффективность армирования на 18,68 %. Дальнейшее увеличение концентрации МУНТ до 1 и 2 масс.% приводит к значительному снижению микротвердости, что обуславливается возможной агломерацией МУНТ, возникающей в результате встраивания МУНТ в пористую структуру керамической матрицы ГА [37, 38].

Методы прогнозирования

Используем методы МО для аппроксимации полученных экспериментальных значений микротвердости по Виккерсу. Рассмотрим применение четырех методов.

Полиномиальная регрессия. Метод регрессионного анализа, который моделирует нелинейную зависимость между независимой переменной x и зависимой переменной y с помощью многочлена n-й степени. В основе полиномиальной регрессии лежит расширение

Таблица 1. Усредненные значения микротвердости керамических образцов, ГПа

Table 1. Average values of microhardness of ceramic samples, GPa

P, Н	Номер серии					
	1	2	3	4	5	6
0,98	4,21 ± 0,47	5,03 ± 0,47	4,35 ± 0,49	4,43 ± 0,44	3,40 ± 0,55	4,15 ± 0,52
1,96	4,05 ± 0,40	4,82 ± 0,44	3,99 ± 0,45	4,88 ± 0,52	3,73 ± 0,46	4,84 ± 0,55
2,94	3,75 ± 0,64	4,73 ± 0,42	4,09 ± 0,43	5,07 ± 0,32	3,97 ± 0,56	3,65 ± 0,42
4,90	3,61 ± 0,42	4,29 ± 0,28	3,83 ± 0,29	4,54 ± 0,26	3,83 ± 0,46	4,23 ± 0,38
9,80	3,52 ± 0,27	4,07 ± 0,35	4,20 ± 0,40	4,65 ± 0,44	3,71 ± 0,39	3,88 ± 0,22

линейной модели путем добавления степеней исходных признаков.

Случайный лес решений. Представляет собой ансамбль решающих деревьев, каждое из которых обучается на случайной подвыборке данных. Итоговое предсказание формируется как среднее значение по всем деревьям решений. Этот метод хорошо подходит для задач с зашумленными и нелинейными данными и устойчив к переобучению.

Метод градиентного бустинга. Построит последовательные деревья на ошибках предыдущих решений. Модель, получаемая с помощью данного метода, характеризуется высокой точностью и стабильностью за счет регуляризации и оптимизации структуры модели.

Методы случайного леса и градиентного бустинга часто используются в задачах, связанных с прогнозированием различных механических характеристик материалов, где показывают, особенно метод градиентного бустинга, высокую точность прогнозирования с коэффициентом детерминации $R^2 > 0{,}9$.

Нейросетевое моделирование. Реализуется в виде полносвязной архитектуры, где в качестве функции активации на слоях используется нелинейная функция, способствующая устойчивому обучению. Такие модели также показывают низкие показатели ошибок, в особенности на данных большой размерности и имеющих сложные закономерности. Для решения поставленной в настоящей работе задачи модели были построены с использованием библиотек Numpy, Pandas, Scikit-learn, XGBoost, а также Keras. У каждой модели присутствует большое количество различных параметров, также называемые гиперпараметрами, которые влияют на точность модели. Оптимальные параметры моделей были подобраны с помощью метода сеточного поиска GridSeacrhCV библиотеки Scikit-learn. Однако подбор архитектуры нейронной сети (количество нейронов на слоях, функции активации) осуществлялся методом байесовской оптимизации библиотеки Keras. В табл. 2 представлены результаты оптимизации гиперпараметров.

Качество всех моделей оценивалось с использованием метрик: R^2 — коэффициент детерминации; Mean Squared Error (MSE) — среднеквадратичная ошибка, Mean Squared Error (MAE) — средняя абсолютная ошибка; Mean Absolute Percentage Error (MAPE) — средняя абсолютная процентная ошибка. На вход моделям на основе методов (табл. 2) подавались значения концентрации МУНТ и нагрузки индентирования, а в качестве выходных параметров рассматривались предсказанные значения твердости и соответствующий доверительный интервал (для вероятности 0,95), рассчитанный по экспериментальным данным.

Моделирование микротвердости по Виккерсу керамики на основе гидроксиапатита

Рассмотрим результаты применения исследуемых методов МО для прогнозирования микротвердости композитов в зависимости от нагрузки и содержания МУНТ. Итоговые значения метрик приведены в табл. 3.

Из полученных результатов видно, что метод полиномиальной регрессии продемонстрировал наихудшие показатели точности в сравнении с другими методами: самое низкое значение $R^2 = 0{,}5308$ и высокие значения отклонений MSE = 0,0534. Для наглядной оценки был построен график отклонений предсказанных значений от экспериментальных (рис. 2, *a*), а также график зависимости твердости от нагрузки для одной из серии образцов (рис. 2, *b*), который включает экспериментальные точки, средние значения и предсказанные значения с доверительными интервалами на примере образца серии 2.

Таблица 2. Оптимальные параметры моделей

Table 2. Optimal parameters of the models

Метод обучения модели	Гиперпараметры модели	Величина
Метод случайного леса	Глубина дерева	10
	Минимальное число объектов в листе	3
	Минимальное число объектов для разбиения	2
	Число деревьев	150
Метод градиентного бустинга	Глубина дерева	4
	Коэффициент скорости обучения	0,1
	Минимальный вес узлов	5
	Доля признаков	1,0
	Доля выборки	1,0
	Число деревьев	1500
Метод нейросетевого моделирования	Функция активации	SeLU
	Оптимизатор	RMSprop
	Количество нейронов во входном слое	70
	Количество нейронов в скрытом слое	55

Таблица 3. Метрики оценки качества регрессионных моделей

Table 3. Performance metrics of prediction models

Метод обучения модели	R^2	MSE	MAE	MAPE, %
Полиномиальная регрессия	0,5308	0,0534	0,1538	0,09
Метод случайного леса	0,5403	0,0416	0,1510	0,10
Метод градиентного бустинга	0,8730	0,0069	0,0646	0,05
Метод нейросетевого моделирования	0,9546	0,0043	0,0353	0,02

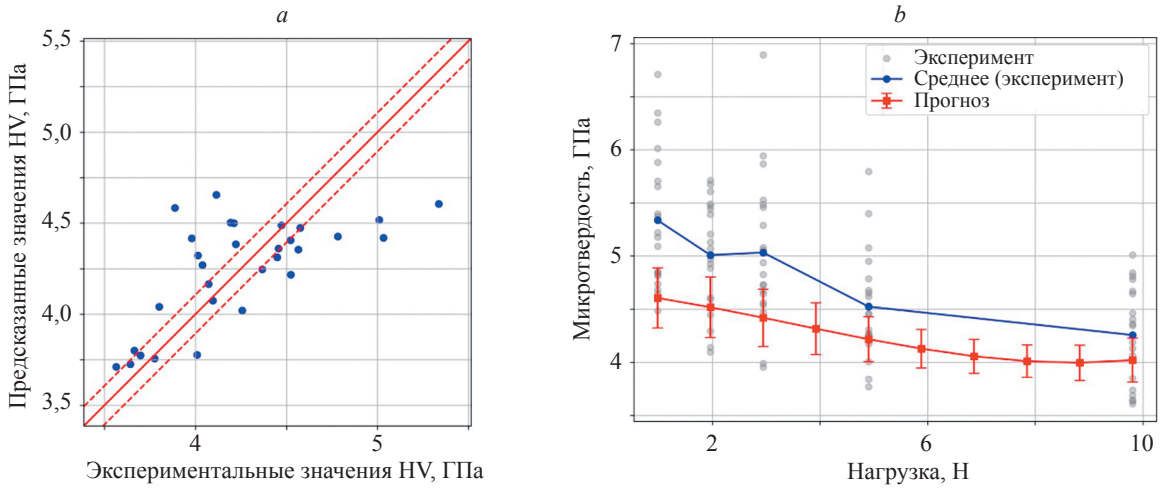

Рис. 2. Результаты моделирования методом полиномиальной регрессии: отклонение предсказанных от экспериментальных значений микротвердости по Виккерсу (HV) (*a*); зависимость экспериментальных и предсказанных значений микротвердости от нагрузки для образца серии 2 (*b*)

Fig. 2. Predicted Vickers microhardness using a polynomial regression model vs. experimental values (*a*) and indentation load (*b*)

Из рис. 2, *a* видно, что экспериментальные данные характеризуются значительным разбросом, особенно при средних и высоких нагрузках индентирования. В отдельных случаях доверительные интервалы, рассчитанные на основе предсказанных значений, не перекрывают средние экспериментальные значения, что указывает на ограниченную точность модели в некоторых областях. Тем не менее, общая тенденция изменения твердости с ростом нагрузки и концентрации МУНТ частично улавливается моделью. Метод полиномиальной регрессии позволил аппроксимировать зависимость твердости от концентрации МУНТ

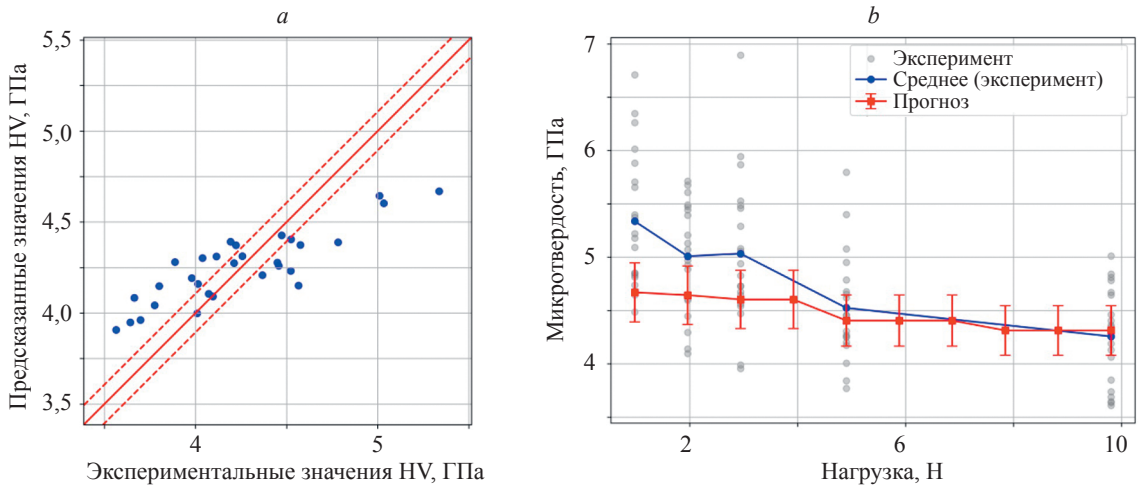

Рис. 3. Результаты моделирования методом случайного леса: отклонение предсказанных от экспериментальных значений микротвердости по Виккерсу (HV) (*a*); зависимость экспериментальных и предсказанных значений микротвердости от нагрузки (*b*) для образца серии 2

Fig. 3. Predicted Vickers microhardness using a random forest model vs. experimental values (*a*) and indentation load (*b*)

и нагрузки с учетом их нелинейной взаимосвязи (рис. 2, *b*). Метод случайного леса (рис. 3) демонстрирует небольшой прирост в сравнении с результатами полинома.

Оценка качества построенной модели методом случайного леса (табл. 3) показала следующие метрики: R^2 = 0,5403, *MSE* = 0,0416. Величины *MAE* = 0,1510 и *MAPE* = 0,10 %, в свою очередь, стали несколько выше, чем у модели полиномиальной регрессии. Анализ графика отклонений предсказанных значений от экспериментальных (рис. 3, *a*) показывает, что многие значения выходят за пределы отклонения ±5 %, что свидетельствует о достаточно высоком уровне ошибок.

Отметим, что на графике зависимости твердости от нагрузки наблюдается эффект «ступенчатости» (рис. 3, *b*): модель часто предсказывает одинаковые значения твердости для разных промежуточных значений нагрузки. Это характерно для метода случайного леса, с помощью которого формируются кусочно-постоянные приближения. Однако это приводит к недостоверной интерполяции, поскольку модель не учитывает, что твердость могла бы плавно возрастать или убывать, что характерно для таких свойств материалов.

Модель, построенная с помощью метода градиентного бустинга, показала значительно более высокое качество прогнозирования по сравнению с применением методов полиномиальной регрессии и случайного леса (рис. 4). На графике отклонения предсказанных значений от экспериментальных видно, что некоторые значения выходят за пределы допустимого отклонения в ±5 %, что свидетельствует о высокой точности полученной модели (рис. 4, *a*). Кроме того, значения метрик, таких как R^2 = 0,8730 и *MAPE* = 0,05 %, подтверждают способность метода хорошо подстраиваться под данные.

Несмотря на полученные высокие значения метрик, при анализе графика зависимости твердости от нагрузки можно заметить эффект «ступенчатости», аналогичный тому, что наблюдается в модели случайного леса (рис. 4, *b*). Это означает, что модель предсказывает одинаковые значения твердости для широких диапазонов нагрузок, особенно между значениями, на которых она была обучена. Такой эффект ограничивает способность модели к качественной интерполяции и снижает ее физическую интерпретируемость, особенно в задачах, где важно предсказать плавные зависимости механических свойств от нагрузки. Модель нейронной сети продемонстрировала наилучшие результаты среди всех рассмотренных методов (рис. 5). На графике отклонения предсказанных значений от экспериментальных наблюдается минимальное количество точек, выходящих за пределы допустимого отклонения в ±5 %, меньшее, чем у модели градиентного бустинга (рис. 5, *a*). Это подтверждается высокими значениями метрик качества: R^2 = 0,9546, а *MAPE* = 0,02 %, что указывает на высокую точность и стабильность полученной модели.

На графике зависимости твердости от нагрузки (рис. 5, *b*) видно, что модель эффективно аппроксимирует экспериментальные данные: отсутствует эффект «ступенчатости», присущий деревообразным моделям, и предсказанные значения плавно следуют за усредненными экспериментальными точками. Благодаря способности нейронной сети выявлять сложные нелинейные зависимости, модель эффективно справляется с интерполяцией. Таким образом, нейросетевое моделирование может быть наиболее эффективным методом прогнозирования твердости композитов, обеспечивая высокую точность и хорошую аппроксимацию экспериментальных данных по всей области нагрузок индентирования и концентраций упрочняющих добавок МУНТ.

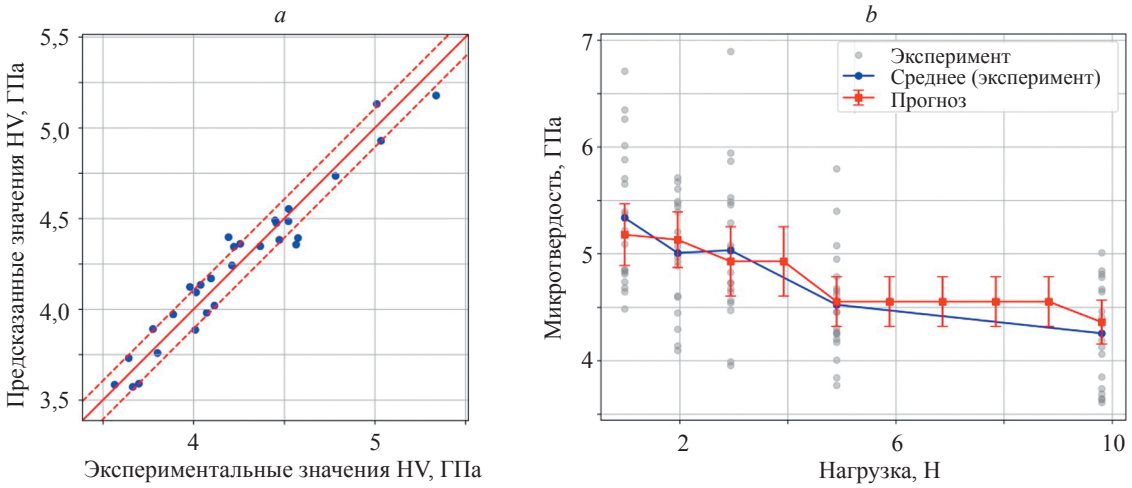

Рис. 4. Результаты моделирования методом градиентного бустинга: отклонение предсказанных от экспериментальных значений микротвердости по Виккерсу (HV) (*a*); зависимость экспериментальных и предсказанных значений микротвердости от нагрузки индентирования для образца серии 2 (*b*)

Fig. 4. Predicted Vickers microhardness using a gradient boosting model vs. experimental values (*a*) and indentation load (*b*)

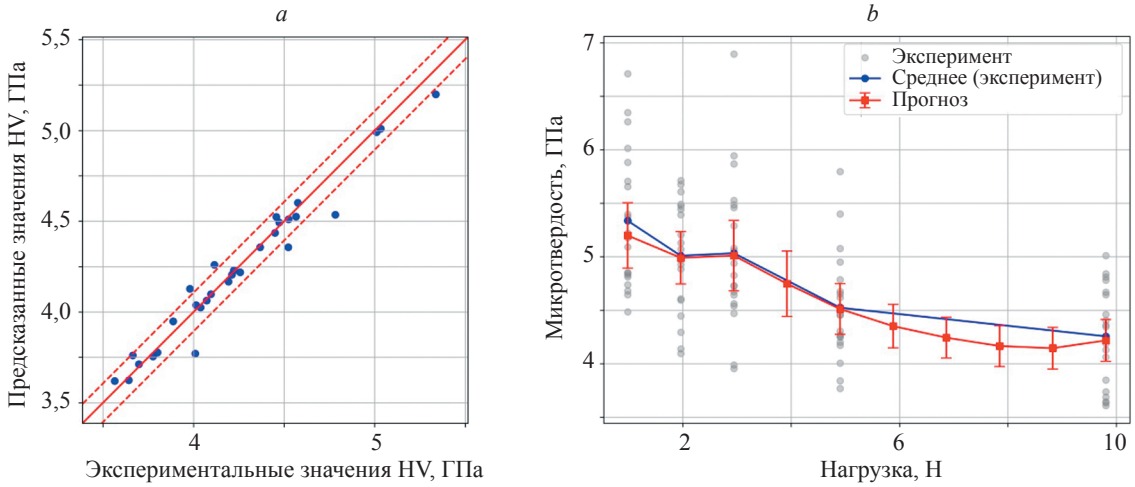

Рис. 5. Результаты нейросетевого моделирования: отклонение предсказанных от экспериментальных значений микротвердости по Виккерсу (HV) (*a*); зависимость экспериментальных и предсказанных значений микротвердости от нагрузки индентирования для образца серии 2 (*b*)

Fig. 5. Predicted Vickers microhardness using a neural network model vs. experimental values (*a*) and indentation load (*b*)

Заключение

Изучена возможность прогнозирования микротвердости композитного материала на основе гидроксиапатита с добавками многостенных углеродных нанотрубок с использованием методов машинного обучения. Экспериментально показано, что введение многостенных углеродных нанотрубок оказывает выраженное влияние на твердость материала: наблюдалось увеличение микротвердости по Виккерсу с 3,83 ± 0,39 ГПа (для образца серии 1) до 4,71 ± 0,40 ГПа (для образца серии 4 с добавкой 0,5 масс.% нанотрубок), что указывает на эффект упрочнения на 19 %. Дальнейшее увеличение концентрации нанотрубок до 1 и 2 масс.% привело к значительному снижению микротвердости, что может быть связано с агломерацией нанотрубок, негативно влияющей на механические свойства материала.

Для построения предиктивных моделей были применены четыре метода: полиномиальная регрессия, случайный лес решений, метод градиентного бустинга и нейросетевое моделирование на основе полносвязной нейронной сети. Обучение проводилось на экспериментальных данных, полученных при измерении микротвердости в диапазоне различных нагрузок индентирования от 0,98 Н до 9,8 Н. В качестве признаков использовались значения нагрузки и концентрации многостенных углеродных нанотрубок. Для оценки качества моделей применялись метрики R^2, MSE, MAE и MAPE. Все модели были обучены с использованием перекрестной проверки и подбора гиперпараметров.

По результатам анализа установлено, что наилучшая точность обучения была получена с помощью нейронной сети: предсказанные значения практически совпадали с экспериментальными, модель не демонстрировала ступенчатого характера предсказаний, а количество значений, выходящих за границы допустимого отклонения, было минимальным. Модель на основе метода градиентного бустинга также представила высокие значения R^2, подтверждающие ее способность хорошо адаптироваться к данным, описываемой в других исследованиях. Однако при использовании данной модели возник эффект «ступенчатости», аналогичный модели случайного леса, у которой, кроме того, наблюдалось большее число грубых отклонений. Полученная модель с применением полиномиальной регрессии продемонстрировала наихудшие результаты как по численным метрикам, так и по визуальному совпадению с экспериментальными значениями.

Таким образом, нейросетевое моделирование показало наивысшую эффективность в задаче прогнозирования микротвердости композитов на основе гидроксиапатита с добавками многостенных углеродных нанотрубок. Полученные модели могут быть использованы для интерполяции значений твердости при новых значениях нагрузки и концентрации нанотрубок и служить инструментом для планирования дальнейших экспериментальных исследований.

Литература

1. Fiume E., Magnaterra G., Rahdar A., Verné E., Baino F. Hydroxyapatite for biomedical applications: A short overview // Ceramics. 2021. V. 4. N 4. P. 542–563. https://doi.org/10.3390/ceramics4040039
2. Rahman M., Li Y., Wen C. HA coating on Mg alloys for biomedical applications: A review // Journal of Magnesium and Alloys. 2020. V. 8. N 3. P. 929–943. https://doi.org/10.1016/j.jma.2020.05.003

References

1. Fiume E., Magnaterra G., Rahdar A., Verné E., Baino F. Hydroxyapatite for biomedical applications: A short overview. *Ceramics*, 2021, vol. 4, no. 4, pp. 542–563. https://doi.org/10.3390/ceramics4040039
2. Rahman M., Li Y., Wen C. HA coating on Mg alloys for biomedical applications: A review. *Journal of Magnesium and Alloys*, 2020, vol. 8, no. 3, pp. 929–943. https://doi.org/10.1016/j.jma.2020.05.003

3. Gu Y.W., Loh N.H., Khor K.A., Tor S.B., Cheang P. Spark plasma sintering of hydroxyapatite powders // Biomaterials. 2002. V. 23. N 1. P. 37–43. https://doi.org/10.1016/S0142-9612(01)00076-X
4. White A.A., Best S.M., Kinloch I.A. Hydroxyapatite–carbon nanotube composites for biomedical applications: a review // International Journal of Applied Ceramic Technology. 2007. V. 4. N 1. P. 1–13. https://doi.org/10.1111/j.1744-7402.2007.02113.x
5. Zhao X., Zheng J., Zhang W., Chen X., Gui Z. Preparation of silicon coated-carbon fiber reinforced HA bio-ceramics for application of load-bearing bone // Ceramics International. 2020. V. 46. N 6. P. 7903–7911. https://doi.org/10.1016/j.ceramint.2019.12.010
6. Khalid P. Suman V.B. Carbon nanotube-hydroxyapatite composite for bone tissue engineering and their interaction with mouse fibroblast L929 In Vitro // Journal of Bionanoscience. 2017. V. 11. N 3. C. 233–240. https://doi.org/10.1166/jbns.2017.1431
7. Ferreira C.R.D., Santiago A.A.G., Vasconcelos R.C., Paiva D.F.F., Pirih F.Q., Araújo A.A., Motta F.V., Bomio M.R.D. Study of microstructural, mechanical, and biomedical properties of zirconia/hydroxyapatite ceramic composites // Ceramics International. 2022. V. 48. N 9. P. 12376–12386. https://doi.org/10.1016/j.ceramint.2022.01.102
8. Lahiri D., Singh V., Keshri A.K., Seal S., Agarwal A. Carbon nanotube toughened hydroxyapatite by spark plasma sintering: microstructural evolution and multiscale tribological properties // Carbon. 2010. V. 48. N 11. P. 3103–3120. https://doi.org/10.1016/j.carbon.2010.04.047
9. Mukherjee S., Kundu B., Chanda A., Sen S. Effect of functionalisation of CNT in the preparation of HAp–CNT biocomposites // Ceramics international. 2015. V. 41. N 3. Part A. P. 3766–3774. https://doi.org/10.1016/j.ceramint.2014.11.052
10. Henriques B., Fabris D., Lopes E., Moreira A.C., Mantovani I.F., Fernandes C.P, Fredel M.C. Influence of the addition of Ni-coated carbon nanotubes on the mechanical properties of highly porous zirconia cellular structures // Advanced Engineering Materials. 2022. V. 24. N 1. P. 2100624. https://doi.org/10.1002/adem.202100624
11. Yu L., Jia P., Song Y., Zhao B., Pan Y., Wang J., Cui H., Feng R., Li H., Cui X., Gao Z., Fang X., Zhang L. Effect of carbon nanotubes on the microstructure and properties of plasma electrolytic oxidized ceramic coatings on high silicon aluminum alloy // Journal of Materials Research and Technology. 2022. V. 18. P. 3541–3552. https://doi.org/10.1016/j.jmrt.2022.04.035
12. Thirugnanasambantham K.G., Sankaramoorthy T., Karthikeyan R., Kumar K.S. A comprehensive review: Influence of the concentration of carbon nanotubes (CNT) on mechanical characteristics of aluminium metal matrix composites: Part 1 // Materials Today: Proceedings. 2021. V. 45. P. 2561–2566. https://doi.org/10.1016/j.matpr.2020.11.267
13. Kumar S.P., Selvamani S.T., Vigneshwar M., Hariharan S.J. Tensile, microhardness, and microstructural analysis on Mg-CNT nano composites // Materials Today: Proceedings. 2018. V. 5. N 2. Part 2. P. 7882–7888. https://doi.org/10.1016/j.matpr.2017.11.469
14. Veljović Đ., Vuković G.D., Steins I., Palcevskis E., Uskoković P., Petrović R., Janaćković Đ. Improvement of the mechanical properties of spark plasma sintered hap bioceramics by decreasing the grain size and by adding multi-walled carbon nanotubes // Science of Sintering. 2013. V. 45. N 2. P. 233–243. https://doi.org/10.2298/sos1302233v
15. Currey J.D. Mechanical properties of bone tissues with greatly differing functions // Journal of Biomechanics. 1979. V. 12. N 4. P. 313–319. https://doi.org/10.1016/0021-9290(79)90073-3
16. Okamoto M., Dohi Y., Ohgushi H., Shimaoka H., Ikeuchi M., Matsushima A., Yonemasu K., Hosoi H. Influence of the porosity of hydroxyapatite ceramics on in vitro and in vivo bone formation by cultured rat bone marrow stromal cells // Journal of Materials Science: Materials in Medicine. 2006. V. 17. N 4. P. 327–336. https://doi.org/10.1007/s10856-006-8232-z
17. Imbeni V., Kruzic J.J., Marshall G.W., Marshall S.J., Ritchie R.O. The dentin–enamel junction and the fracture of human teeth // Nature Materials. 2005. V. 4. N 3. P. 229–232. https://doi.org/10.1038/nmat1323
18. Nastic A., Merati A., Bielawski M., Bolduc M., Fakolujo O., Nganbe M. Instrumented and Vickers indentation for the characterization of stiffness, hardness and toughness of zirconia toughened Al_2O_3 and SiC armor // Journal of Materials Science and Technology. 2015. V. 31. N 8. P. 773–783. https://doi.org/10.1016/j.jmst.2015.06.005

3. Gu Y.W., Loh N.H., Khor K.A., Tor S.B., Cheang P. Spark plasma sintering of hydroxyapatite powders. *Biomaterials*, 2002, vol. 23, no. 1, pp. 37–43. https://doi.org/10.1016/S0142-9612(01)00076-X
4. White A.A., Best S.M., Kinloch I.A. Hydroxyapatite–carbon nanotube composites for biomedical applications: a review. *International Journal of Applied Ceramic Technology*, 2007, vol. 4, no. 1, pp. 1–13. https://doi.org/10.1111/j.1744-7402.2007.02113.x
5. Zhao X., Zheng J., Zhang W., Chen X., Gui Z. Preparation of silicon coated-carbon fiber reinforced HA bio-ceramics for application of load-bearing bone. *Ceramics International*, 2020, vol. 46, no. 6, pp. 7903–7911. https://doi.org/10.1016/j.ceramint.2019.12.010
6. Khalid P. Suman V.B. Carbon nanotube-hydroxyapatite composite for bone tissue engineering and their interaction with mouse fibroblast L929 In Vitro. *Journal of Bionanoscience*, 2017, vol. 11, no. 3, pp. 233–240. https://doi.org/10.1166/jbns.2017.1431
7. Ferreira C.R.D., Santiago A.A.G., Vasconcelos R.C., Paiva D.F.F., Pirih F.Q., Araújo A.A., Motta F.V., Bomio M.R.D. Study of microstructural, mechanical, and biomedical properties of zirconia/hydroxyapatite ceramic composites. *Ceramics International*, 2022, vol. 48, no. 9, pp. 12376–12386. https://doi.org/10.1016/j.ceramint.2022.01.102
8. Lahiri D., Singh V., Keshri A.K., Seal S., Agarwal A. Carbon nanotube toughened hydroxyapatite by spark plasma sintering: microstructural evolution and multiscale tribological properties. *Carbon*, 2010, vol. 48, no. 11, pp. 3103–3120. https://doi.org/10.1016/j.carbon.2010.04.047
9. Mukherjee S., Kundu B., Chanda A., Sen S. Effect of functionalisation of CNT in the preparation of HAp–CNT biocomposites. *Ceramics international*, 2015, vol. 41, no. 3, part A, pp. 3766–3774. https://doi.org/10.1016/j.ceramint.2014.11.052
10. Henriques B., Fabris D., Lopes E., Moreira A.C., Mantovani I.F., Fernandes C.P, Fredel M.C. Influence of the addition of Ni-coated carbon nanotubes on the mechanical properties of highly porous zirconia cellular structures. *Advanced Engineering Materials*, 2022, vol. 24, no. 1, pp. 2100624. https://doi.org/10.1002/adem.202100624
11. Yu L., Jia P., Song Y., Zhao B., Pan Y., Wang J., Cui H., Feng R., Li H., Cui X., Gao Z., Fang X., Zhang L. Effect of carbon nanotubes on the microstructure and properties of plasma electrolytic oxidized ceramic coatings on high silicon aluminum alloy. *Journal of Materials Research and Technology*, 2022, vol. 18, pp. 3541–3552. https://doi.org/10.1016/j.jmrt.2022.04.035
12. Thirugnanasambantham K.G., Sankaramoorthy T., Karthikeyan R., Kumar K.S. A comprehensive review: Influence of the concentration of carbon nanotubes (CNT) on mechanical characteristics of aluminium metal matrix composites: Part 1. *Materials Today: Proceedings*, 2021, vol. 45, pp. 2561–2566. https://doi.org/10.1016/j.matpr.2020.11.267
13. Kumar S.P., Selvamani S.T., Vigneshwar M., Hariharan S.J. Tensile, microhardness, and microstructural analysis on Mg-CNT nano composites. *Materials Today: Proceedings*, 2018, vol. 5, no. 2, part 2. pp. 7882–7888. https://doi.org/10.1016/j.matpr.2017.11.469
14. Veljović Đ., Vuković G.D., Steins I., Palcevskis E., Uskoković P., Petrović R., Janaćković Đ. Improvement of the mechanical properties of spark plasma sintered hap bioceramics by decreasing the grain size and by adding multi-walled carbon nanotubes. *Science of Sintering*, 2013, vol. 45, no. 2, pp. 233–243. https://doi.org/10.2298/sos1302233v
15. Currey J.D. Mechanical properties of bone tissues with greatly differing functions. *Journal of Biomechanics*, 1979, vol. 12, no. 4, pp. 313–319. https://doi.org/10.1016/0021-9290(79)90073-3
16. Okamoto M., Dohi Y., Ohgushi H., Shimaoka H., Ikeuchi M., Matsushima A., Yonemasu K., Hosoi H. Influence of the porosity of hydroxyapatite ceramics on in vitro and in vivo bone formation by cultured rat bone marrow stromal cells. *Journal of Materials Science: Materials in Medicine*, 2006, vol. 17, no. 4, pp. 327–336. https://doi.org/10.1007/s10856-006-8232-z
17. Imbeni V., Kruzic J.J., Marshall G.W., Marshall S.J., Ritchie R.O. The dentin–enamel junction and the fracture of human teeth. *Nature Materials*, 2005, vol. 4, no. 3, pp. 229–232. https://doi.org/10.1038/nmat1323
18. Nastic A., Merati A., Bielawski M., Bolduc M., Fakolujo O., Nganbe M. Instrumented and Vickers indentation for the characterization of stiffness, hardness and toughness of zirconia toughened Al_2O_3 and SiC armor. *Journal of Materials Science and Technology*, 2015, vol. 31, no. 8, pp. 773–783. https://doi.org/10.1016/j.jmst.2015.06.005

19. Пашков Д.М., Беляк О.А., Гуда А.А., Колесников В.И. Обратное проектирование механических и трибологических свойств покрытий: результаты алгоритмов машинного обучения // Физическая мезомеханика. 2022. Т. 25. № 2. С. 24–34. https://doi.org/10.55652/1683-805X_2022_25_2_24
20. Abueidda D.W., Almasri M., Ammourah R., Ravaioli U., Jasiuk I.M., Sobh N.A. Prediction and optimization of mechanical properties of composites using convolutional neural networks // Composite Structures. 2019. V. 227. P. 111264. https://doi.org/10.1016/j.compstruct.2019.111264
21. DeVore R., Hanin B., Petrova G. Neural network approximation // Acta Numerica. 2021. V. 30. P. 327–444. https://doi.org/10.1017/s0962492921000052
22. Li Y., Li H., Jin C., Shen J. The study of effect of carbon nanotubes on the compressive strength of cement-based materials based on machine learning // Construction and Building Materials. 2022. V. 358. P. 129435. https://doi.org/10.1016/j.conbuildmat.2022.129435
23. Akbari P., Zamani M., Mostafaei A. Machine learning prediction of mechanical properties in metal additive manufacturing // Additive Manufacturing. 2024. V. 91. P. 104320. https://doi.org/10.1016/j.addma.2024.104320
24. Golkarnarenji G., Naebe M., Badii K., Milani A.S., Jazar R.N., Khayyam H. A machine learning case study with limited data for prediction of carbon fiber mechanical properties // Computers in Industry. 2019. V. 105. P. 123–132. https://doi.org/10.1016/j.compind.2018.11.004
25. Краснов Ф.В. Выявление ошибок разметки данных с помощью моделей классификации для небольших наборов данных // International Journal of Open Information Technologies. 2023. Т. 11. № 5. С. 54–62.
26. Xu P., Ji X., Li M., Lu W. Small data machine learning in materials science // npj Computational Materials. 2023. V. 9. N 1. P. 42. https://doi.org/10.1038/s41524-023-01000-z
27. Karamov R., Akhatov I., Sergeichev I.V. Prediction of fracture toughness of pultruded composites based on supervised machine learning // Polymers. 2022. V. 14. N 17. P. 3619. https://doi.org/10.3390/polym14173619
28. Şimşek Türker Y., Kilinçarslan S., Yilmaz Ince E. Performance of ANN, Random Forest and XGBoost methods in predicting the flexural properties of wood beams reinforced with carbon-FRP // Wood Material Science and Engineering. 2025. V. 20. N 3. P. 657–668. https://doi.org/10.1080/17480272.2024.2370942
29. Han T., Huang J., Sant G., Neithalath N., Kumar A. Predicting mechanical properties of ultrahigh temperature ceramics using machine learning // Journal of the American Ceramic Society. 2022. V. 105. N 11. P. 6851–6863. https://doi.org/10.1111/jace.18636
30. Shah V., Zadourian S., Yang C., Zhang Z., Gu, G.X. Data-driven approach for the prediction of mechanical properties of carbon fiber reinforced composites // Materials Advances. 2022. V. 3. N 19. P. 7319–7327. https://doi.org/10.1039/d2ma00698g
31. Carneiro M.V., Salis T.T., Almeida G.M., Braga A.P. Prediction of mechanical properties of steel tubes using a machine learning approach // Journal of Materials Engineering and Performance. 2021. V. 30. N 1. P. 434–443. https://doi.org/10.1007/s11665-020-05345-0
32. Zhang Z., Mansouri Tehrani A., Oliynyk A.O., Day B., Brgoch J. Finding the next superhard material through ensemble learning // Advanced Materials. 2021. V. 33. N 5. P. 2005112. https://doi.org/10.1002/adma.202005112
33. Dovale-Farelo V., Tavadze P., Lang L., Bautista-Hernandez A., Romero A.H. Vickers hardness prediction from machine learning methods // Scientific Reports. 2022. V. 12. N 1. P. 22475. https://doi.org/10.1038/s41598-022-26729-3
34. Qadir A., Ali S., Dusza J., Rafaja D. Predicting hardness of graphene-added Si_3N_4 using machine learning: A data-driven approach // Open Ceramics. 2024. V. 19. P. 100634. https://doi.org/10.1016/j.oceram.2024.100634
35. Barabashko M.S., Tkachenko M.V., Neiman A.A., Ponomarev A.N., Rezvanova A.E. Variation of Vickers microhardness and compression strength of the bioceramics based on hydroxyapatite by adding the multi-walled carbon nanotubes // Applied Nanoscience. 2020. V. 10. N 8. P. 2601–2608. https://doi.org/10.1007/s13204-019-01019-z
36. Barabashko M., Ponomarev A., Rezvanova A., Kuznetsov V., Moseenkov S. Young's modulus and vickers hardness of the hydroxyapatite bioceramics with a small amount of the multi-walled carbon nanotubes // Materials. 2022. V. 15. N 15. P. 5304. https://doi.org/10.3390/ma15155304

19. Pashkov D.M., Belyak O.A., Guda A.A., Kolesnikov V.I. Reverse engineering of mechanical and tribological properties of coatings: results of machine learning algorithms. *Physical Mesomechanics*, 2022, vol. 25, no. 4, pp. 296–305. https://doi.org/10.1134/s1029959922040038
20. Abueidda D.W., Almasri M., Ammourah R., Ravaioli U., Jasiuk I.M., Sobh N.A. Prediction and optimization of mechanical properties of composites using convolutional neural networks. *Composite Structures*, 2019, vol. 227, pp. 111264. https://doi.org/10.1016/j.compstruct.2019.111264
21. DeVore R., Hanin B., Petrova G. Neural network approximation. Acta *Numerica*, 2021, vol. 30, pp. 327–444. https://doi.org/10.1017/s0962492921000052
22. Li Y., Li H., Jin C., Shen J. The study of effect of carbon nanotubes on the compressive strength of cement-based materials based on machine learning. *Construction and Building Materials*, 2022, vol. 358, pp. 129435. https://doi.org/10.1016/j.conbuildmat.2022.129435
23. Akbari P., Zamani M., Mostafaei A. Machine learning prediction of mechanical properties in metal additive manufacturing. *Additive Manufacturing*, 2024, vol. 91, pp. 104320. https://doi.org/10.1016/j.addma.2024.104320
24. Golkarnarenji G., Naebe M., Badii K., Milani A.S., Jazar R.N., Khayyam H. A machine learning case study with limited data for prediction of carbon fiber mechanical properties. *Computers in Industry*, 2019, vol. 105, pp. 123–132. https://doi.org/10.1016/j.compind.2018.11.004
25. Krasnov F.V. Identifying data labeling errors using classification models for small datasets. *International Journal of Open Information Technologies*, 2023, vol. 11, no. 5, pp. 54–62. (in Russian)
26. Xu P., Ji X., Li M., Lu W. Small data machine learning in materials science. *npj Computational Materials*, 2023, vol. 9, no. 1, pp. 42. https://doi.org/10.1038/s41524-023-01000-z
27. Karamov R., Akhatov I., Sergeichev I.V. Prediction of fracture toughness of pultruded composites based on supervised machine learning. *Polymers*, 2022, vol. 14, no. 17, pp. 3619. https://doi.org/10.3390/polym14173619
28. Şimşek Türker Y., Kilinçarslan S., Yilmaz Ince E. Performance of ANN, Random Forest and XGBoost methods in predicting the flexural properties of wood beams reinforced with carbon-FRP. *Wood Material Science and Engineering*, 2025, vol. 20, no. 3, pp. 657–668. https://doi.org/10.1080/17480272.2024.2370942
29. Han T., Huang J., Sant G., Neithalath N., Kumar A. Predicting mechanical properties of ultrahigh temperature ceramics using machine learning. *Journal of the American Ceramic Society*, 2022, vol. 105, no. 11, pp. 6851–6863. https://doi.org/10.1111/jace.18636
30. Shah V., Zadourian S., Yang C., Zhang Z., Gu, G.X. Data-driven approach for the prediction of mechanical properties of carbon fiber reinforced composites. *Materials Advances*, 2022, vol. 3, no. 19, pp. 7319–7327. https://doi.org/10.1039/d2ma00698g
31. Carneiro M.V., Salis T.T., Almeida G.M., Braga A.P. Prediction of mechanical properties of steel tubes using a machine learning approach. *Journal of Materials Engineering and Performance*, 2021, vol. 30, no. 1, pp. 434–443. https://doi.org/10.1007/s11665-020-05345-0
32. Zhang Z., Mansouri Tehrani A., Oliynyk A.O., Day B., Brgoch J. Finding the next superhard material through ensemble learning. *Advanced Materials*, 2021, vol. 33, no. 5, pp. 2005112. https://doi.org/10.1002/adma.202005112
33. Dovale-Farelo V., Tavadze P., Lang L., Bautista-Hernandez A., Romero A.H. Vickers hardness prediction from machine learning methods. *Scientific Reports*, 2022, vol. 12, no. 1, pp. 22475. https://doi.org/10.1038/s41598-022-26729-3
34. Qadir A., Ali S., Dusza J., Rafaja D. Predicting hardness of graphene-added Si_3N_4 using machine learning: A data-driven approach. *Open Ceramics*, 2024, vol. 19, pp. 100634. https://doi.org/10.1016/j.oceram.2024.100634
35. Barabashko M.S., Tkachenko M.V., Neiman A.A., Ponomarev A.N., Rezvanova A.E. Variation of Vickers microhardness and compression strength of the bioceramics based on hydroxyapatite by adding the multi-walled carbon nanotubes. *Applied Nanoscience*, 2020, vol. 10, no. 8, pp. 2601–2608. https://doi.org/10.1007/s13204-019-01019-z
36. Barabashko M., Ponomarev A., Rezvanova A., Kuznetsov V., Moseenkov S. Young's modulus and vickers hardness of the hydroxyapatite bioceramics with a small amount of the multi-walled

37. Шутилов Р.А., Мызь А.Л., Кузнецов В.Л., Карагедов Г.Р. Токопроводящие керамические композиты на основе Al_2O_3, модифицированные многослойными углеродными нанотрубками // Перспективные материалы. 2016. № 8. С. 64–73.
38. Сидоренко Д.А., Зайцев А.А., Курбаткина В.В., Левашов Е.А., Андреев В.А., Рупасов С.И., Севастьянов П.И. Влияние добавок углеродных нанотрубок на структуру и свойства металлических связок для алмазного инструмента // Известия высших учебных заведений. Порошковая металлургия и функциональные покрытия. 2012. № 1. С. 38–43.

carbon nanotubes. *Materials*, 2022, vol. 15, no. 15, pp. 5304. https://doi.org/10.3390/ma15155304
37. Shutilov R., Myz A., Kuznetsov V., Karagedov G. Current conductive AL_2O_3 ceramic composites modified by multiwall carbon nanotubes. *Perspektivnye Materialy*, 2016, no. 8, pp. 64–73. (in Russian)
38. Sidorenko D.A., Zaitsev A.A., Kurbatkina V.V., Levashov E.A., Anreyev V.A., Rupasov S.I., Sevastianov P.I. The effect of carbon nanotube additives on structure and propertiesof metal binders for diamond cutting tools. *Powder Metallurgy and Functional Coatings*, 2012, no. 1, pp. 38–43. (in Russian)

Авторы

Резванова Анастасия Евгеньевна — младший научный сотрудник, Институт физики прочности и материаловедения Сибирского отделения Российской академии наук, Томск, 634055, Российская Федерация, sc 57199302281, https://orcid.org/0000-0002-7067-7979, ranast@ispms.ru

Кудряшов Борис Сергеевич — аспирант, инженер-исследователь, Институт физики прочности и материаловедения Сибирского отделения Российской академии наук, Томск, 634055, Российская Федерация, sc 57656690100, https://orcid.org/0009-0000-5133-4893, bsk3@ispms.ru

Погудин Владимир Юрьевич — инженер, Институт физики прочности и материаловедения Сибирского отделения Российской академии наук, Томск, 634055, Российская Федерация; студент, Томский государственный университет систем управления и радиоэлектроники, Томск, 634050, Российская Федерация, https://orcid.org/0009-0002-7191-163X, pogudin.vova@bk.ru

Authors

Anastasiya E. Rezvanova — Junior Researcher, Institute of Strength Physics and Materials Science of the Siberian Branch of the Russian Academy of Sciences, Tomsk, 634055, Russian Federation, sc 57199302281, https://orcid.org/0000-0002-7067-7979, ranast@ispms.ru

Boris S. Kudryashov — PhD Student, Engineer-Researcher, Institute of Strength Physics and Materials Science of the Siberian Branch of the Russian Academy of Sciences, Tomsk, 634055, Russian Federation, sc 57656690100, https://orcid.org/0009-0000-5133-4893, bsk3@ispms.ru

Vladimir Yu. Pogudin — Engineer, Institute of Strength Physics and Materials Science of the Siberian Branch of the Russian Academy of Sciences, Tomsk, 634055, Russian Federation; Student, Tomsk State University of Control Systems and Radioelectronics, Tomsk, 634050, Russian Federation, https://orcid.org/0009-0002-7191-163X, pogudin.vova@bk.ru

Статья поступила в редакцию 28.07.2025
Одобрена после рецензирования 03.10.2025
Принята к печати 23.11.2025

Received 28.07.2025
Approved after reviewing 03.10.2025
Accepted 23.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1208-1219

Aerodynamic analysis of SD7037 airfoil with flex-skin flap at low angles of attack and low Reynolds numbers

Ahmed Tawfeeq Mustafa Ali Abed[1]✉, Azad A. Hussein[2], Saba Hafedh Mahdi[3]

[1,2,3] Aeronautical Engineering Department, College of Engineering, University of Baghdad, Baghdad, 10071, Iraq

[1] a.t.abed@coeng.uobaghdad.edu.iq✉, https://orcid.org/0000-0002-8025-9367
[2] azad.a@coeng.uobaghdad.edu.iq, https://orcid.org/0000-0001-9792-7774
[3] saba.hafedh@coeng.uobaghdad.edu.iq, https://orcid.org/0009-0009-9479-0536

Abstract
This work concentrates on the effect of a flex-skin trailing edge flap on the aerodynamic characteristics of SD7037 airfoil at low Reynolds numbers, in the range of $2\cdot10^5$ to $5\cdot10^5$ using computational methods. The study used a range of angle of attack (AOA) associated with the take-off phase and different flap angles. The numerical model was set up in Siemens STAR-CCM+ package using the κ-ω shear stress transport turbulence model and the (γ-Reθ) transition model which ensured the approximate solution of Navier-Stokes equations. The verification of the computational solution was done by the comparison with the available experimental data of the plain flap, and it was discovered that the results matched pretty well at lower AoAs. Results indicated that certain sets of AoAs and flap angles can notably achieve the lift over the drag ratio above the baseline conditions, thus improved the performance especially during take-off stage. Besides, some combinations were found to be inefficient, and these were recommended to be discarded. Additionally, the results showed that the flex-skin flap generated higher lift coefficient but also higher drag coefficient at the same range of AoAs as compared to that of the plain flap.

Keywords
UAV, aerodynamics, lift, drag, performance

For citation: Abed A.T.M.A., Hussein A.A., Mahdi S.H. Aerodynamic analysis of SD7037 airfoil with flex-skin flap at low angles of attack and low Reynolds numbers. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1208–1219. doi: 10.17586/2226-1494-2025-25-6-1208-1219

УДК 533.6

Аэродинамический анализ профиля SD7037 с гибким закрылком при малых углах атаки и малых числах Рейнольдса

Ахмед Тауфик Мустафа Али Абед[1]✉, Азад А. Хусейн[2], Саба Хафед Махди[3]

[1,2,3] Отделение авиационной техники Инженерного колледжа, Багдадский университет, Багдад, 10071, Ирак

[1] a.t.abed@coeng.uobaghdad.edu.iq✉, https://orcid.org/0000-0002-8025-9367
[2] azad.a@coeng.uobaghdad.edu.iq, https://orcid.org/0000-0001-9792-7774
[3] saba.hafedh@coeng.uobaghdad.edu.iq, https://orcid.org/0009-0009-9479-0536

Аннотация
С использованием вычислительных методов изучено влияние закрылка беспилотного летательного аппарата с гибкой обшивкой на аэродинамические характеристики профиля SD7037 при низких числах Рейнольдса в диапазоне от $2\cdot10^5$ до $5\cdot10^5$. Исследование проводилось в диапазоне углов атаки (Angles of Attacks, AoAs), связанных с фазой взлета и различными углами закрылка беспилотного летательного аппарата. Численная модель реализована в пакете Siemens STAR-CCM+ с использованием модели турбулентности переноса касательных напряжений κ-ω и модели перехода γ-Reθ, которая обеспечивает приближенное решение уравнений Навье–Стокса. Проверка вычислительного решения выполнена путем сравнения с имеющимися экспериментальными данными, полученными для плоского закрылка. Обнаружено, что результаты довольно хорошо совпадают при более низких AoAs. Показано, что определенные наборы AoAs и углов закрылка могут заметно обеспечить

© Abed A.T.M.A., Hussein A.A., Mahdi S.H., 2025

подъемную силу с аэродинамическим качеством выше базовых условий, тем самым улучшая характеристики, особенно на этапе взлета. Установлено, что некоторые комбинации параметров оказались неэффективными, и от них было рекомендовано отказаться. Обнаружено, что гибкая обшивка закрылка обеспечивает более высокий коэффициент подъемной силы, но также и более высокий коэффициент лобового сопротивления при том же диапазоне углов атаки по сравнению с обычным закрылком беспилотного летательного аппарата.

Ключевые слова
БПЛА, аэродинамика, подъемная сила, сопротивление, производительность, беспилотный летательный аппарат

Ссылка для цитирования: Абед А.Т.М.А., Хусейн А.А., Махди С.Х. Аэродинамический анализ профиля SD7037 с гибким закрылком при малых углах атаки и малых числах Рейнольдса // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1208–1219 (на англ. яз.). doi: 10.17586/2226-1494-2025-25-6-1208-1219

Introduction

Aircraft lift can be boosted by the use of flaps which also affects stall speeds, take-off, and landing distances by lowering these characteristics. These effects are utilized in Unmanned Aerial Vehicles (UAVs) to improve the lift, but with increased drag. The UAV flight stability is also enhanced. At specific situations the efficiency of the airfoil is improved, which in return improves UAVs range. For these positive outcomes various flap types have been formed over the years to get definite aerodynamic characteristics [1, 2]. Flex-skin flaps are very simple and can be easily fabricated with very simple tools and techniques at truncated cost. Comparing to other types of flaps, this minimalism comes with limited reliability that solely depends on the type of the material used to form the skin of the airfoil and the flap, as presented by Ooda [3]. Therefore, studying airfoils with flaps, including flex-skin type, could provide valuable insights for a broad range of UAV engineers, hobbyist, and makers [4].

Literature Review

Due to its obvious importance many studies were carried for various types of flaps on different airfoil sections such as the work done by Frolov and Wenming [5] who analyzed the effect of plain flap on airfoil lift by using computational methods. The complex variable function theory was used to analyze the proposed model accompanied with conformal mapping and the discrete vortex method. The study showed linear relation between the increase of lift coefficient *cl* and the increase of the plain flap angle.

While Todorov [6] analyzed numerically a single slotted flap airfoil NACA-23012, the study included a numerical analysis of aerodynamic coefficient at Reynolds number of $3{\cdot}10^6$ and simulated the forces on the flap using computational fluid dynamics utilized by ANSYS Fluent code. The proposed technique uses a turbulent model with Spalart-Allmaras method. Also, the airfoil has 1 m chord with an additional 0.32 m chord for the slotted flap. The study found out that the slotted flap generated a greater lift coefficient *cl* than that of the (NACA-23012) profile, with less drag coefficient *cl* for the flap model.

Srivastava [7] modeled and analyzed airfoil NACA-2412 with flap to generate high-lift. The 3D model simulated using Reynolds number 1.4 million (velocity 20.42 m/s). The dimensions of the model were 1 m and 0.332 m for the chord and flap, respectively. The study showed that the maximum lift generated at 10° flap angle.

The flap deflection on NACA-2412 airfoil was simulated numerically by Dutta et al. [8]. The flow regime selected to be subsonic. Also, for each angle of attack (AoA) from the set (0°, 4°, 8°, and 12°) the model used a different flap deflection. The model simulated using ANSYS Fluent with structured mesh-grid and the model dimensions set to be 150 mm and 300 mm for chord length and span, respectively. The end results indicated that with the increasing the AoA, the lift coefficient *cl* enhanced up to a certain degree. As well as, the study simulate the relation between different flap deflection and the flight condition.

Also, the airfoil section NACA-2412 with plain flap was studied using computational techniques by Thejarju et al. [9] using ANSYS Fluent for various flap angles. The study showed that the maximum lift coefficient *cl* observed at 32° flap angle.

Because the calculations of aerodynamic forces are accurate for flows where boundary layer transition may occur, such as in wind turbine and in many other crucial aerospace applications, Michna et al. [10] evaluated the accuracy of the γ-Reθ boundary layer transition method in forecasting the aerodynamics of the DU-91-W2-250 profile at high Reynolds numbers. The study inspects transition onset, boundary layer development, and flow separation by comparing numerical results with experimental measurements. While the model effectively gets key transition phenomena, inconsistencies arise in separation-induced transition and turbulent reattachment regions. The findings indicate that γ-Reθ is suitable for preliminary aerodynamic analysis but requires refinements for high-fidelity simulations.

The aerodynamic effects of fixed wing UAV airfoil SD-7037 were numerically evaluated by Ramanan et al. [11]. The model simulated at low Reynolds number, with dimensions of the model (719 mm, 292 mm, and 120 mm for length, width and total height, respectively). The model simulated using ANSYS Fluent. The result showed that the simulation analysis is satisfying to meet the performance aspects.

The change of the flap angle influences on the performance of airfoil type NACA-4415 is studied in the research carried by Vinod and Joth [12]. The model simulated at Reynolds number $4{\cdot}10^6$ and AoA varied from –10° to 30°, using Xfoil software. Also, the flap angle was (–90°, –45°, +45° and +90°). The analysis showed that during the take-off situation the flap can be useful due to

the generation of sudden lift. The viscosity effects were not considered in this work as the Xfoil software provides limited supported for boundary layer simulation.

The lift and drag forces of airfoil (M21) with plain flapper at diverse AoAs was investigated numerically by Patel et al. [13]. The different flapped designs simulated using ANSYS Fluent. Also, the κ-ω turbulence model used with 2,000,006 Reynolds number as initial boundary conditions. The chord length selected to be 1.0148 m. The mesh/grid generation tested around 395,846 elements. The study showed that the drag and lift increase with the AoA increases. The angle of stall found out to be around 24°. The plain flapper angle that improved the lift found out to be around 10°.

Flaps are also used in wind turbines to reap its benefits, thus, Ma et al. [14], simulated numerically and analyzed the effect of adding flap to airfoil S 809 wind turbine. The flap added to the upper surface of the S 809 airfoil with a cord length of 1 meter and the Reynolds number value of $1 \cdot 10^6$. The 2D model was simulated using ANSYS Fluent with 2D incompressible Reynolds Averaged Navier-Stokes (RANS) equations. The total mesh selected to be around 160,000 cells. The study showed that the added flap improve the aerodynamic performance of the optimal performance found out to be at AoA = 14°, where the drag decreased by 3.98 % and the lift increased by 13.67 %.

Computational and experimentally methods was used to investigate the flap angles on the performance of a home-built wing NACA-23012 in the work of Salam et al. [15]. The tested angles of attack were 0°, 5°, 10°, 15° and 20°, and the flap angles were –15°, 0°, 15°, 30°, and 45°. The chord and the span of the model was taken as 1.6 m and 0.29 m, respectively. The study found out that the maximum drag was at 20° AoA and 45° of flap angle. Whereas, the maximum lift was at 0° AoA and at 15° of flap angle. The model optimal performance was at 15° AoA and 45° of flap angle.

Krishnan et al. [16] conducted a simulation to illustrate the consequences of adding flap at different angles on a wing NACA-6412. The AoA for the model varied from –4° to 16° and Mach number from 0.2 to 0.4. The flap angles used were 10°, 20°, and 30°. The model was simulated using ANSYS ICEM CFD. The aerodynamic analysis showed that the maximum lift occurs at 14° AoA. Also, the drag and the lift increased with increasing the flap angle.

Soman et al. [17] investigated the outcomes of changing the numerical settings on the behavior of the γ-Reθ model in simulating both two dimensional and three dimensional bluff body transitional flows using the Open-FOAM software. The influence of various solver configurations and turbulence models on the accuracy and convergence of the simulations was examined, and it was shown that numerical settings significantly affect the transition prediction, especially in complex flow scenarios involving separation and reattachment. The findings emphasize the need for careful optimization of computational parameters to improve the reliability of transition modeling in practical applications, such as aerodynamic design and fluid-structure interaction analysis.

The well-known NACA-4412 airfoil with a plain flap was evaluated numerically by Xiao et al. [18] as this type of airfoil is widely used in UAVs. The simulated model uses κ-ω Shear Stress Transport (SST) method with 1 m chord length and $6 \cdot 10^6$ Reynolds number and around 4° AoA. The study showed that at a certain AoA with increasing the flap angle, the suction shortfall increases on the lower wing surface.

Parluhutan et al. [19] evaluated numerically the effect of plain flap on the lift of symmetrical airfoils NACA-0015. The model simulated at $4 \cdot 10^6$ Reynolds number with κ-Epsilon turbulent model and different angle of deflection of 15° and 30°, the AoA selected to be from 0° to 25°. The study found out that the lift coefficient *cl* increases due to the presence of the flap.

It is worth to mention that less complex methods, other than those that uses control volume methods applied in Computational Fluid Dynamics (CFD), may be used to analyze aerodynamic characteristics of airfoils with flaps, such as the panel method, both linear and non-linear, utilized by Ali [20, 21], but it lacks the capabilities to calculate viscosity effects. These methods can't be used to calculate airfoil performance for practical applications although it requires much less computational power as compared to control volume methods used in this work.

From the presented literature review it can be seen that neither recent nor previous studies have investigated the use of the SD7037 airfoil with a flex-skin flap, in spite of its potential to improve efficiency and its simple manufacturing methods.

This work is a continuation of a previous study carried by Abed [4]. In this work, there exists the aerodynamic performance of the SD7037 airfoil section with trailing edge flex-skin flap, employing CFD at low AoA and Reynolds numbers within the range of $2 \cdot 10^5$ to $5 \cdot 10^5$. The analysis encompasses a variety of low AoA alongside differing flap angles at take-off phase by solving the flow equations of Navier-Stokes by the application of CFD with utilizing κ-ω SST turbulence method and γ-Reθ transition method available in Siemens STAR-CCM+ package to help in the prediction of the performance of UAV during take-off phase.

Numerical Modeling

For this research, the SD7037 airfoil has been selected, with a 0.23 m as the chord length (c) to be compared with the plain flap data from Abed [4]. A flex-skin flap is added to the trailing edge of the airfoil section at the ($0.79c$) position, thus the flap chord is ($0.21c$) of the total chord. The study is carried at sea level conditions. The value of Reynolds number for the flow is set to be ranging from $2 \cdot 10^5$ to $5 \cdot 10^5$.

These settings were preferred to assess the numerical calculations against the experimental measurements for verification, and also to be compared with CFD results from Abed [4]. It is to be noted that no experimental information is available for such case, thus the verification will be carried based on the general behavior of the numerical results for the flex-skin flap with the data from MS Selig et al. [22] for the plain flap of the same airfoil section. The airfoil profile information were based on those from [22] and shaped using utilities provided by the software package, which are presented in Fig. 1.

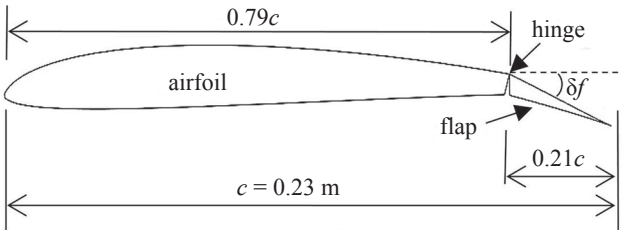

Fig. 1. Drawing of the airfoil and the flex-skin flap with deflection angle δ*f*

For the presented case, it is expected that the flow across the airfoil to be mixed of both laminar and turbulent nature, due to the local Reynolds number could reach and may surpasse the value of $5 \cdot 10^5$, thus become turbulent and encounter a transition phase. Therefore, RANS equations were applied to estimate the viscosity influences. The airfoil has been numerically simulated for steady-state viscous flow conditions at sea level at two dimensions with constant density.

In order to accurately account for the viscosity effects and include the calculations of the transition phenomena, κ-ω with shear stress transport turbulence model presented by Menter, F.R. [23] with γ-Reθ transition method has been selected. The correlations required to formulate this model is based on those typically provided within the model with no calibration, Menter [24].

The computational domain has been chosen to be a square of a large size, 20 m were defined as shown in Fig. 2, this is to eliminate and reduce the effects from the domain boundaries on the airfoil. The inlet has been established to be a velocity inlet, and the outlet at the domain exit was set to a pressure outlet.

Unstructured mesh is preferred to discretize the domain because of its flexibility as presented by Versteeg, H.K. and Malalasekera [25], while a structured mesh was used to model the surface of the airfoil and the flex-skin flap so as to capture the flow nature correctly and in details. The polygonal type of elements has been applied due to their higher quality, accuracy, and stability as compared to the other available types. This type of elements requires more time due its higher requirements of computational power [25].

The boundary layer over the model has been studied, using 60 layers of prism cells near the model solid boundaries. The height of the cell next to the wall calculated to be at $1.2 \cdot 10^{-5}$ m using the method introduced in [25]. The wake of the airfoil was modeled with a length of 20 times of the airfoil chord. Mesh refinement was applied and checked by running a mesh independence study to achieve the accuracy and stability of the solution.

From Fig. 3 below, it can be concluded that the numerical solution will stabilize around 1,563,143 elements, with acceptable accuracy. At higher number of elements the solution become more stable but requires much higher computational power, which will limit the number of cases that can be computed.

Thus, compromise between the level of the accuracy of the solution and the computational power required may be applied. It can be seen from Table 1 below that the difference of the lift coefficient *cl* value between the successive assessments is very small for the last three trials in the mesh independence study, thus the selected mesh yields results with acceptable accuracy and less computational time. Similar technique was used by Abed [4] to increase computational efficiency.

Based on this analysis, the mesh applied in the solution with 1,563,143 elements is presented in Fig. 4 below.

Results and Discussion

As it was mentioned before, the general behavior of the numerical results for the flex-skin flap is compared with that of the experimental measurements from MS Selig, et al. [22] for the plain flap of the same airfoil section in order to verify the correct application of the suggested CFD model.

The behavior of the numerical results show conformity with the experimental information from [22] for the studied cases, as shown in Fig. 5 below, where experimental measurements presented in (Exp) in the figure as solid lines for various flap deflection angles δ*f* are compared to the computational results (CFD) presented in bullets. Thus

Table 1. Change of lift coefficient *cl* among the last trails of the mesh independence study

Number of Elements	Percentage change of lift coefficient *cl* between successive trails, %
1,563,143	0.19
2,635,899	0.11
5,625,178	0.00

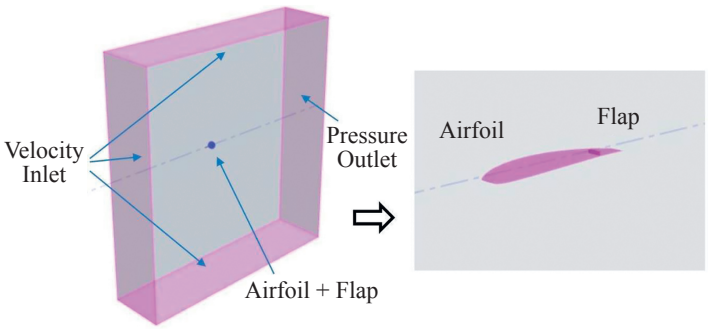

Fig. 2. The computational domain, airfoil with flap

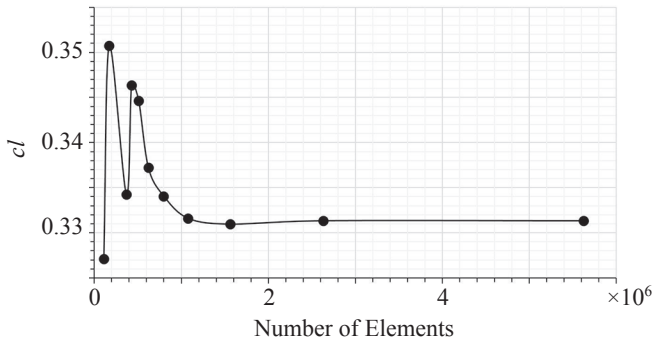

Fig. 3. Mesh Independence Study

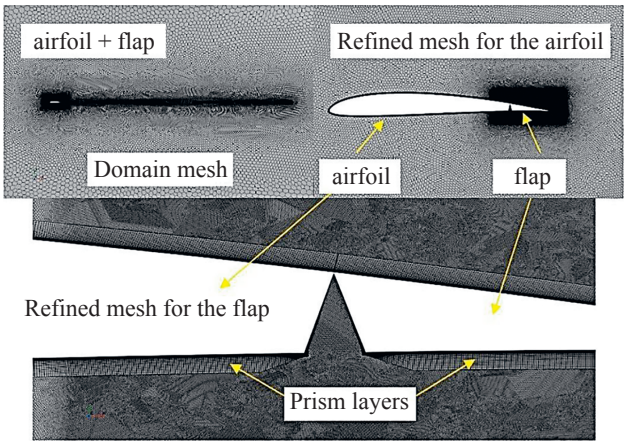

Fig. 4. The applied mesh of the model

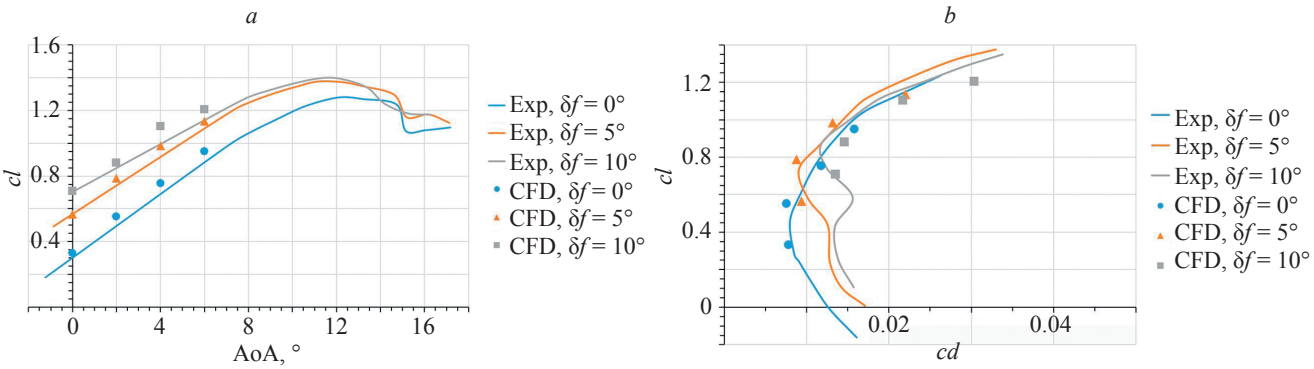

Fig. 5. Validation of CFD results (CFD, bullets) with experimental measurements (Exp, the solid lines) for lift coefficient *cl* (*a*) vs. AoA and drag polar (*b*) vs. drag coefficient *cd* for Reynolds number $3 \cdot 10^5$ for various flap deflection angles in degree

the suggested numerical model may be used to calculate the aerodynamic characteristics for the flex-skin flap on the SD7037 airfoil with acceptable accuracy.

In this work, calculations was carried over the range of AoA from 0° to 6° for a set of various flap deflection angles at 0°, 5°, 10°, 15°, and 20° in the range of Reynolds numbers from $2 \cdot 10^5$ to $5 \cdot 10^5$ with a $0.5 \cdot 10^5$ step, totaling 140 cases. The results are shown and discussed in the following paragraphs.

By the analysis of the computational results shown in Fig. 6, *a–d* below, it can be determined that the value of the lift to drag ratio coefficients (*cl*/*cd*), or the aerodynamic performance of the model, is the highest at AoA 2° at all values of Reynolds numbers, while at AoA 6° is the lowest.

At AoA 2°, δf 10°, Re $3 \cdot 10^5$ further examination of the flow uncovers that near the flap hinge and in the direction of the trailing edge the pressure gradient is unfavorable with many small regions of separated flow as shown in Fig. 7 below. The pressure and velocity contours indicate that drag is much higher in the case of δf 10° as compared to that of δf 5° due to the small regions of flow separations on the upper surface and due to the drop in pressure at the hinge region of the flap on the lower surface.

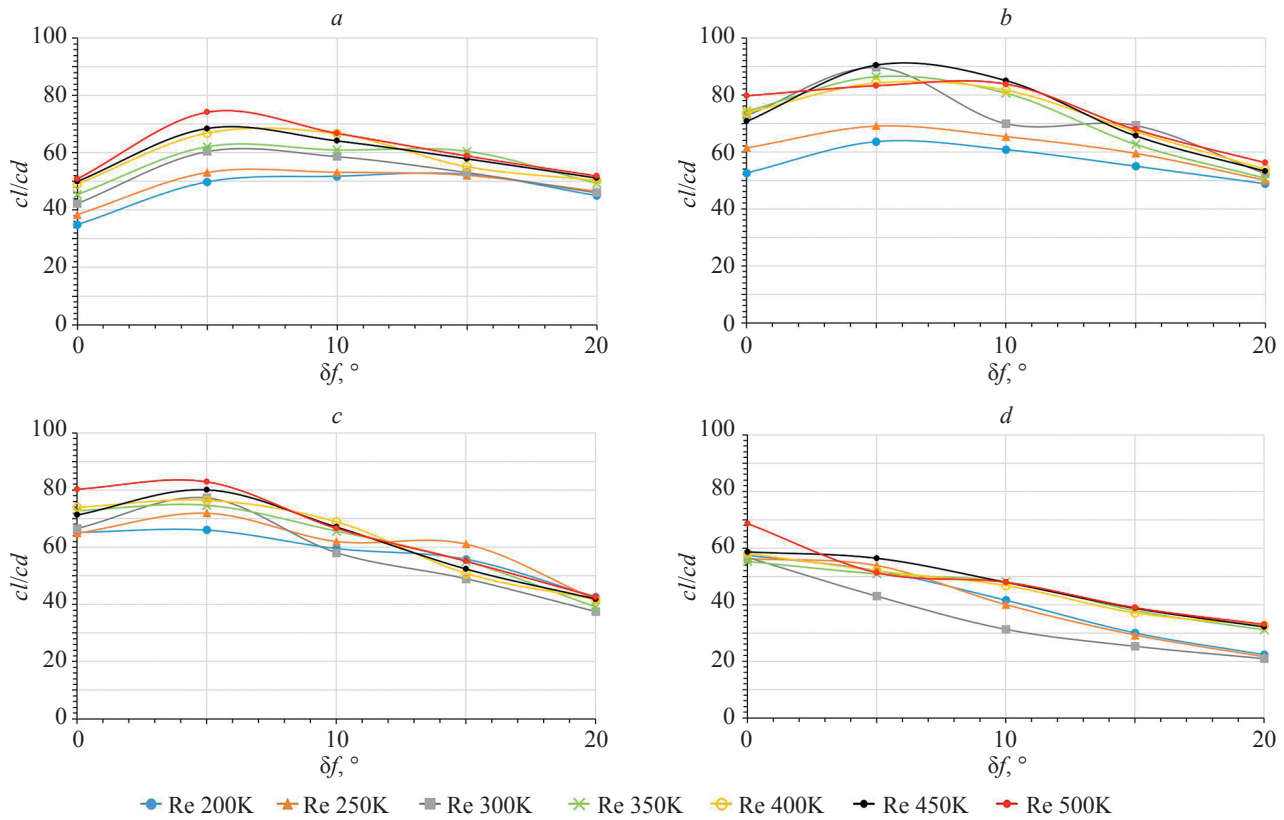

Fig. 6. Lift to drag coefficient ratio *cl*/*cd* of the model at various flap angles for different Reynolds numbers at AoA 0° (*a*), 2° (*b*), 4° (*c*), 6° (*d*)

Fig. 8 shows a closer view of the velocity contours at the flap hinge region which confirms the turbulent nature of the flow near and around the flap.

Fig. 9 demonstrates the moment coefficient (*cm*) as a function of δ*f* for different Reynolds numbers for various AoAs. The moment coefficient *cm* is calculated at the

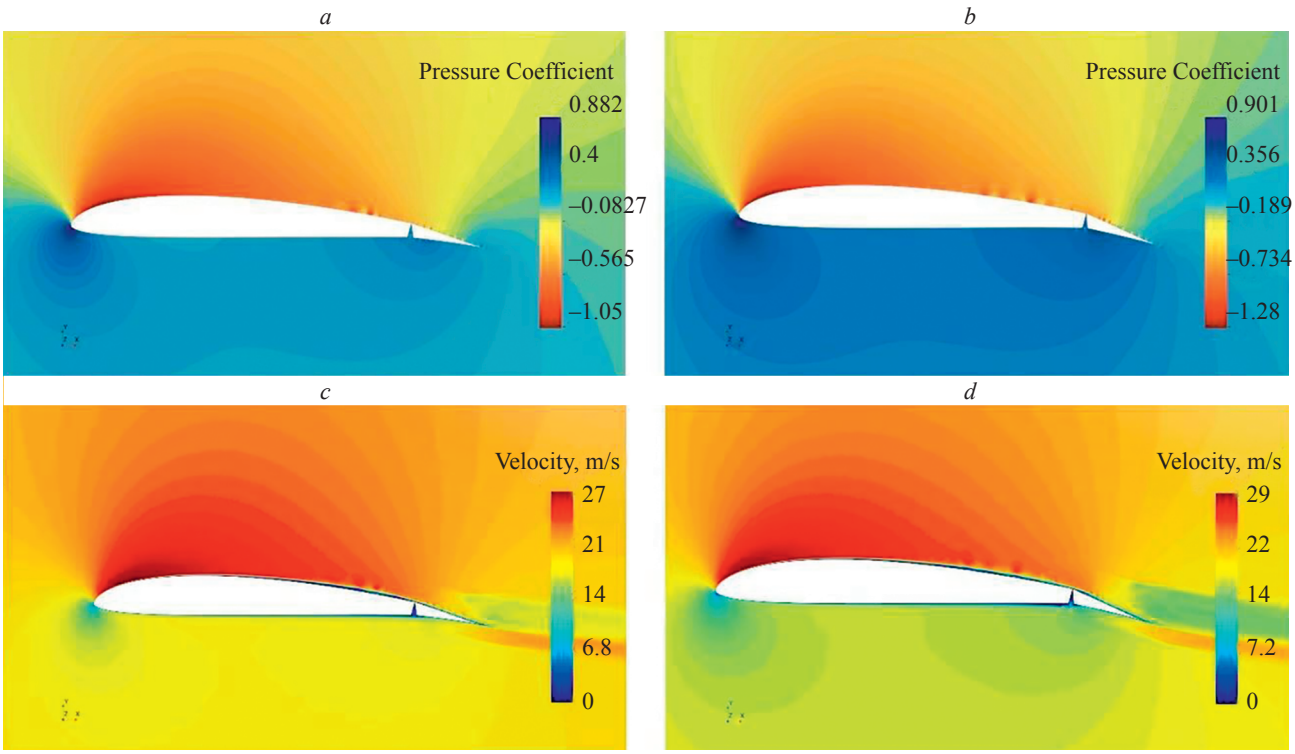

Fig. 7. Pressure and velocity contours at AoA 2°, Re 3·10^5, δ*f* 5° (*a*, *c*), δ*f* 10° (*b*, *d*)

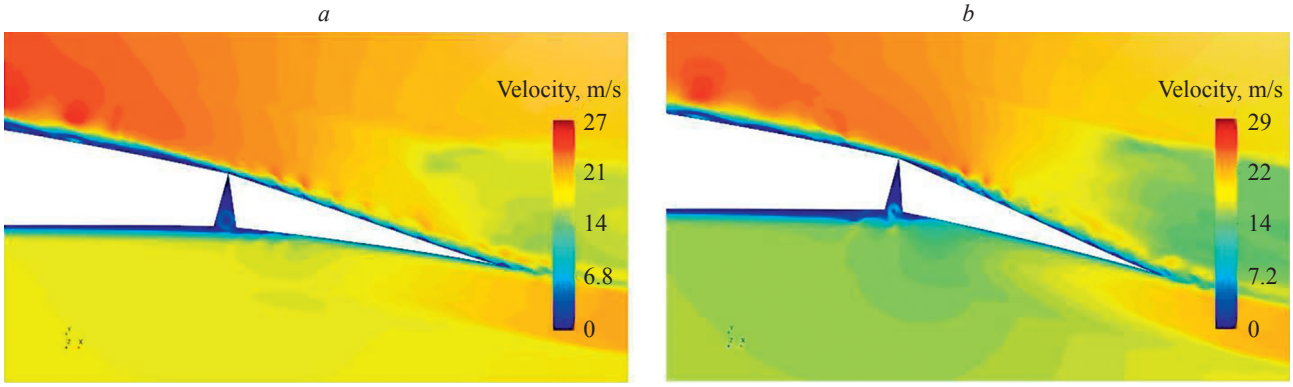

Fig. 8. Velocity contours near flap hinge at AoA 2°, Re 3·10^5, δ*f* 5° (*a*), δ*f* 10° (*b*)

airfoil leading edge with its negative values indicates a counterclockwise direction, or the pitch down moment coefficient *cm*. It can be concluded that as the AoA increases the pitching moment increases as well. This is due to the increase in the forces over the top surface of the model, which means that during take-off a higher δ*f* causes the airfoil to pitch down to counterbalance the moment. It is to be taken into considerations that Reynolds number has a limited effect on the moment coefficient *cm*, except at AoA 6°, δ*f* 20°, as shown in Fig. 9, *d*.

A comparison of the aerodynamic characteristics has been carried between the model with the flex-skin flap and that with the plain flap presented by Abed [4] showed that at Reynolds number of Re 3·10^5 for the lift coefficient *cl* for different AoA at various flap deflection angles is practically the same as can be seen in Fig. 10. This is true except for the situation when the AoA is at 4° and 6° at flap deflection angle of 10°. In those cases the lift for the plain flap model drops noticeably as compared to that of the flex-skin flap. The flex-skin flap retained a more steady

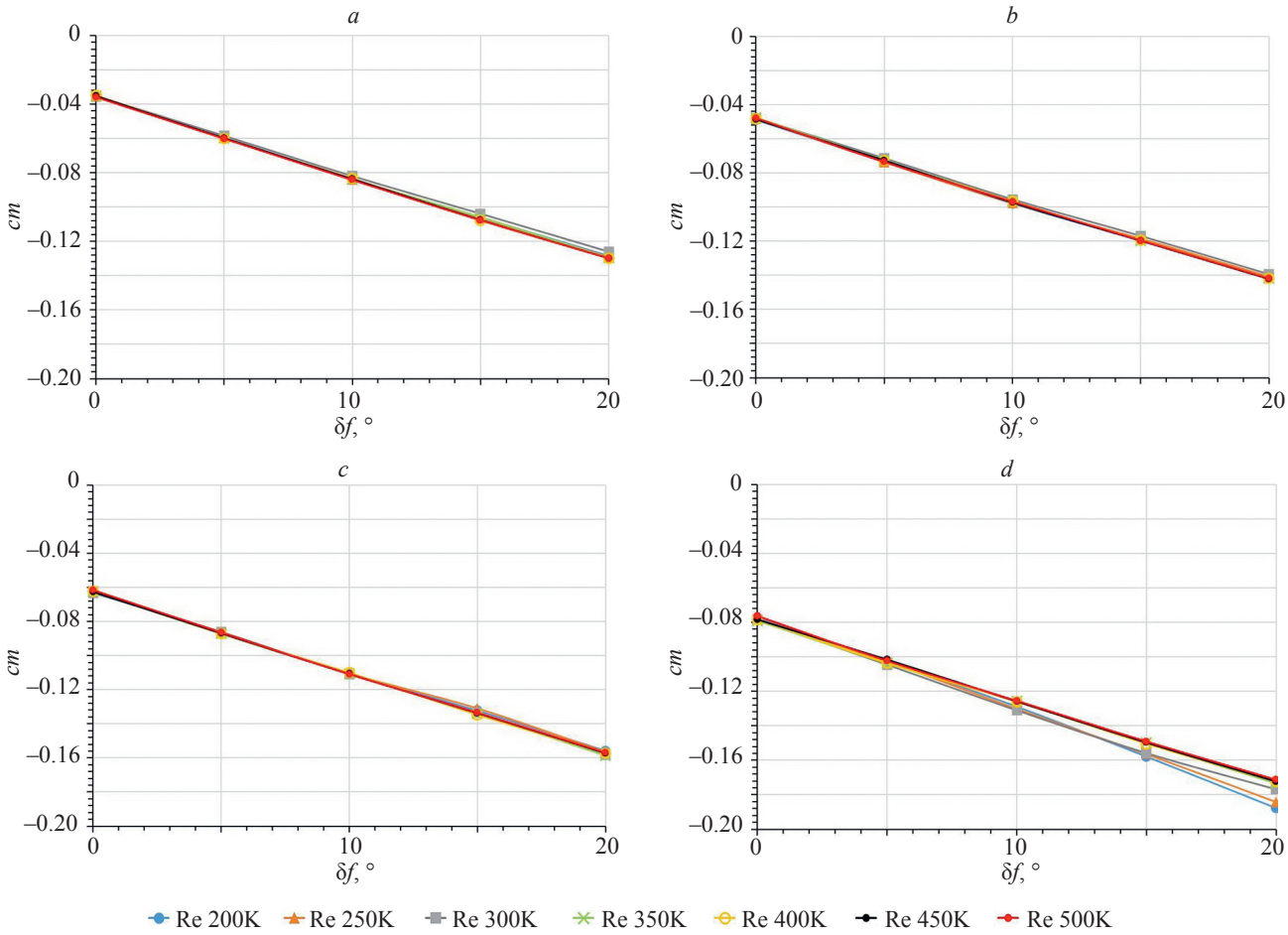

Fig. 9. Moment coefficient *cm* of the model vs. various flap angles for different Reynolds numbers at AoA: 0° (*a*), 2° (*b*), 4° (*c*), 6° (*d*)

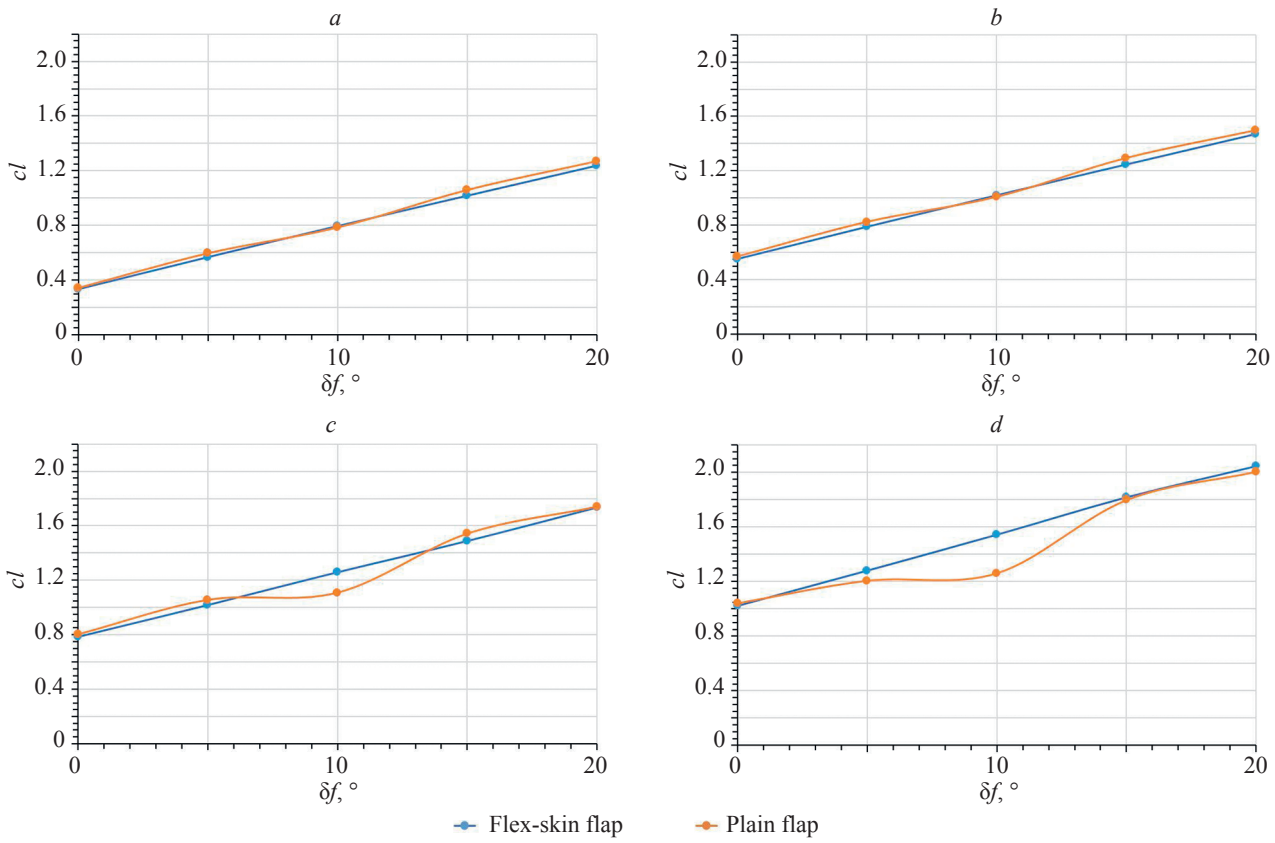

Fig. 10. Lift coefficient *cl* comparison between the flex-skin and the plain flap at Re $3 \cdot 10^5$ for various flap deflection angles at AoA: 0° (*a*), 2° (*b*), 4° (*c*), 6° (*d*)

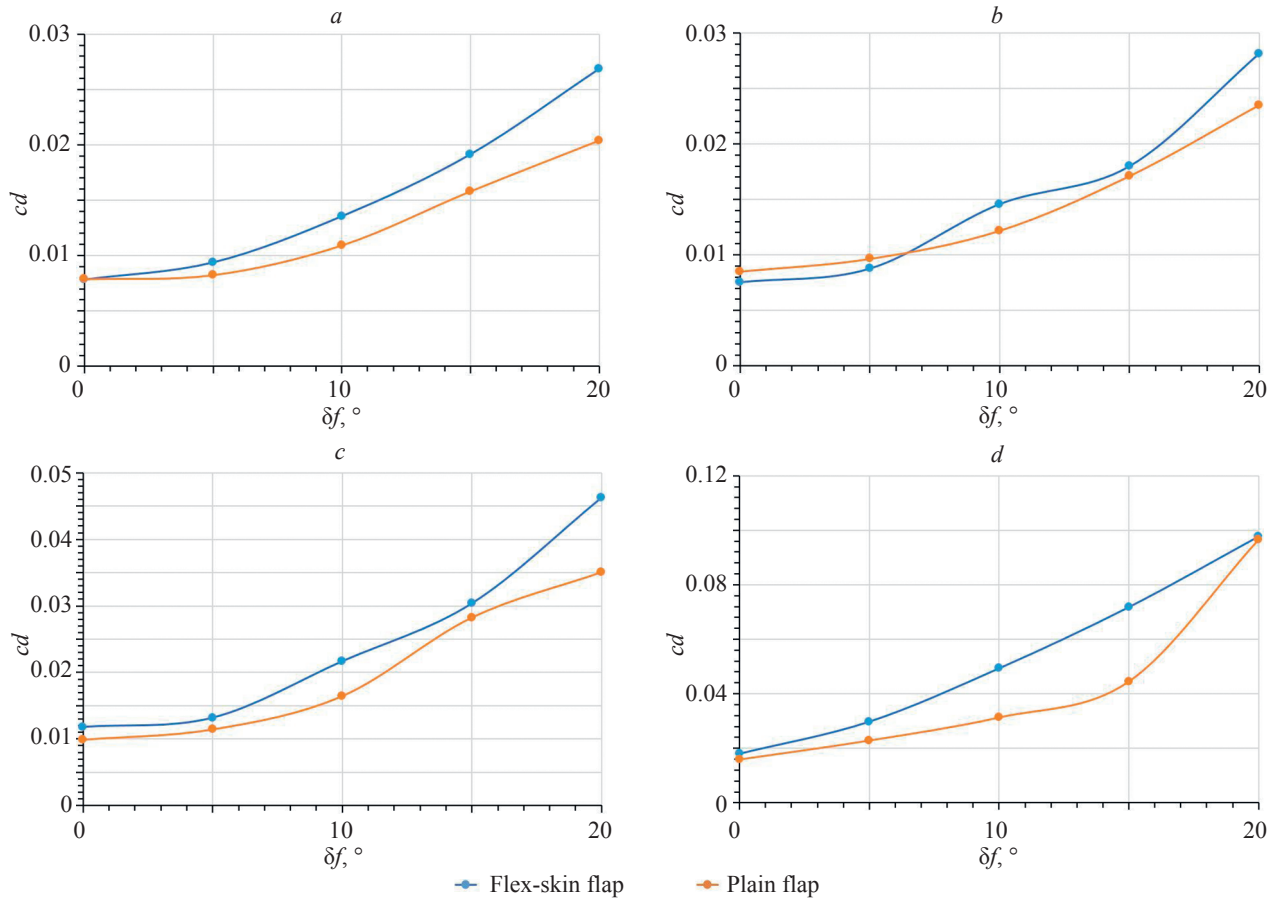

Fig. 11. Drag coefficient *cd* comparison between the flex-skin and the plain flap at Re $3 \cdot 10^5$ for various flap deflection angles at AoA: 0° (*a*), 2° (*b*), 4° (*c*), 6° (*d*)

flow and enhanced pressure gradient over the upper surface due to the continuous surface with no abrupt changes in the geometry.

The flex-skin flap model showed higher drag coefficient *cd* than that of the plain flap as can be concluded from Fig. 11 for the same conditions and cases. This is because of the unsmooth change in the flow route on the bottom surface of the model with flex-skin flap at the hinge point between the wing and the flap. This sudden change of the direction of the flow as it passes the hinge area produced losses of energy resulting in the increase of the drag.

Analysis of the value of the lift to drag of the model presented in the current work as compared to that of the model with the plain flap from Abed [4] is shown in Fig. 12 in similar conditions. It can be deuced that the flex-skin flap offers higher performance at AoAs 0° and 2° while it provides less performance at AoA 4° and 6° with the highest performance recorded at AoA 2° for the model with the plain flap. This is because of the higher drag produced by the application of the flex-skin flap.

The moment coefficient *cm* of the model with the flex-skin flap at Reynolds number of $3 \cdot 10^5$ is compared to that

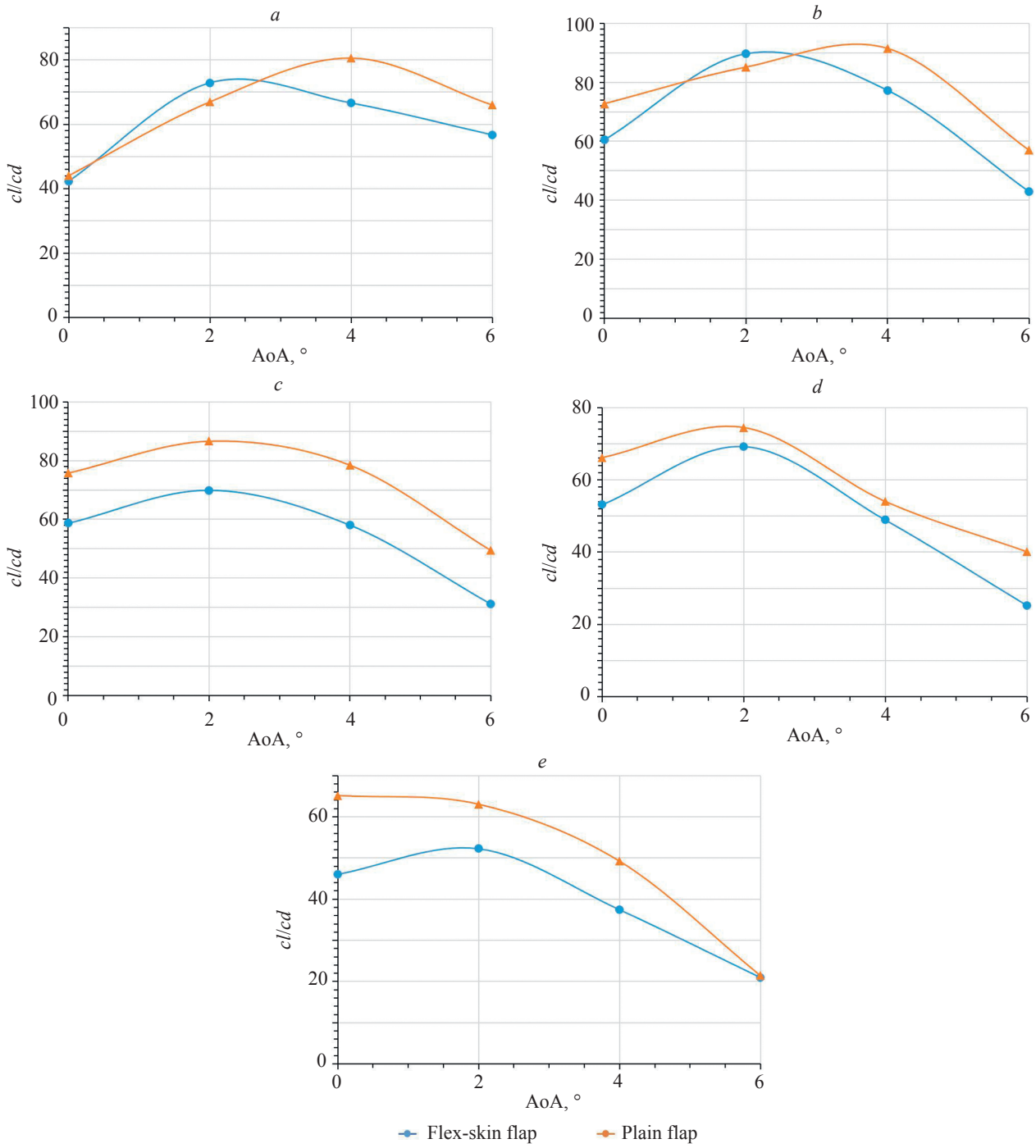

Fig. 12. Lift to drag ratio *cl*/*cd* comparison between the flex-skin and the plain flap at Re $3 \cdot 10^5$ for various angles of attack at δf: 0° (*a*), 5° (*b*), 10° (*c*), 15° (*d*), 20° (*e*)

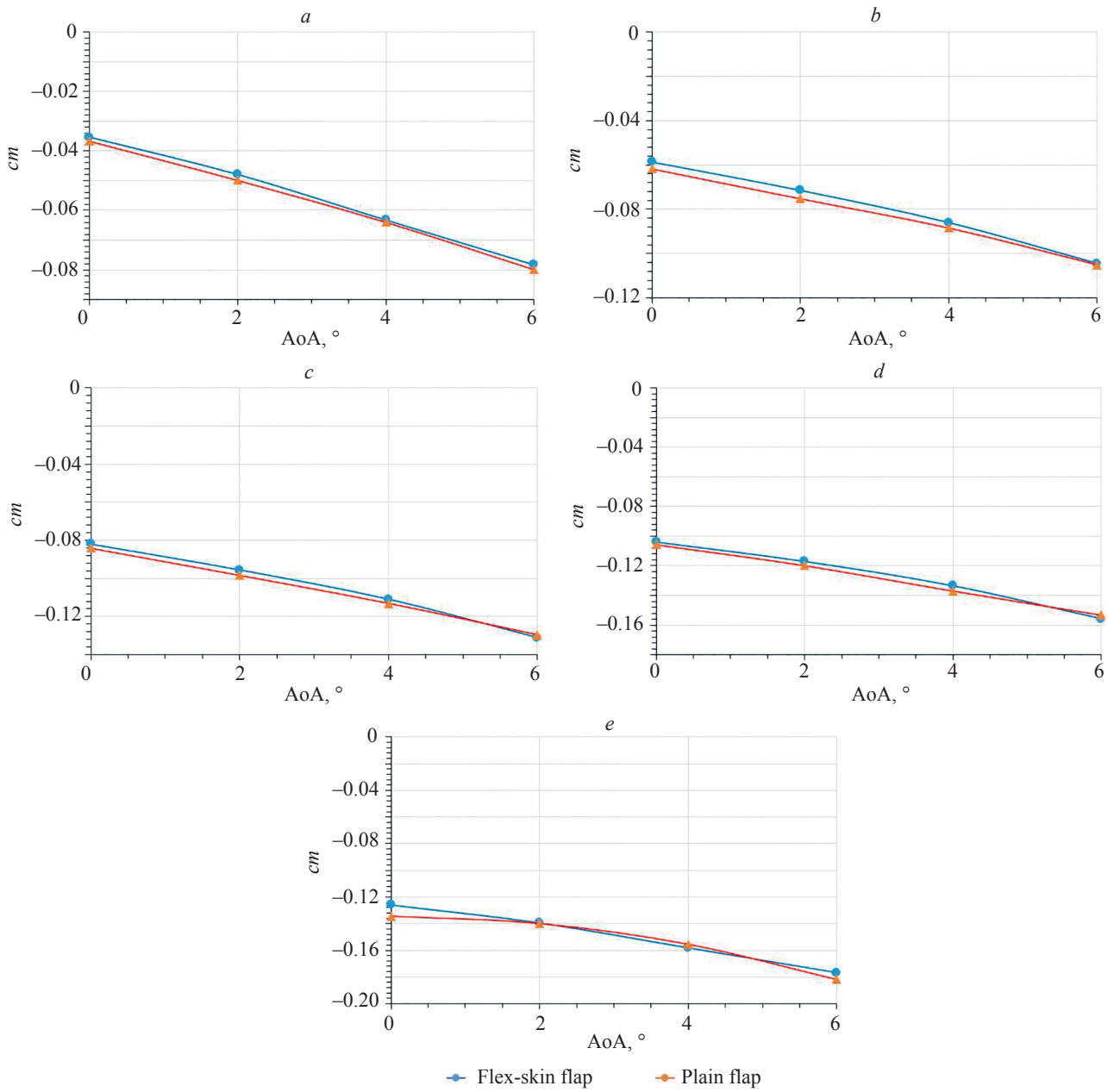

Fig. 13. Moment coefficient *cm* comparison between the flex-skin and the plain flap at Re $3 \cdot 10^5$ for various angles of attack at AoA 0° (*a*), 5° (*b*), 10° (*c*), 15° (*d*), 20° (*e*)

of the plain flap results from Abed [4] which showed that there are insignificant differences, as can be seen in Fig. 13 below.

Conclusion

This research investigates the aerodynamic behavior of an airfoil with SD7037 profile and flex-skin flap at the trailing edge during take-off phase with small Angles of Attack (AoAs) and Reynolds numbers within the range of $2 \cdot 10^5$ to $5 \cdot 10^5$ by the application of Computational Fluid Dynamics (CFD) simulation through the use of Siemens STAR-CCM+ software package.

The research showed that using the κ-ω SST model together with the γ-Reθ transition method is suitable for capturing the complex flow behavior around the SD7037 airfoil with a flex-skin flap at small angles of attack, since it can represent the mixed laminar-turbulent characteristics with acceptable accuracy. The numerical results showed that the best aerodynamic performance, in terms of lift-to-drag ratio, was obtained at a Reynolds number of $4.5 \cdot 10^5$, AoA 2°, and flap deflection angle of 5°. In fact, for the different Reynolds numbers considered, the maximum lift-to-drag ratio consistently appeared at AoA = 2° and δf = 5°, whereas the lowest lift-to-drag ratio occurred at AoA = 6° for the various flap deflection angles examined. The analysis also indicated that the flex-skin flap can provide a higher lift coefficient *cl* than the plain flap, but this improvement is accompanied by an increase in drag, which ultimately reduces the overall aerodynamic efficiency, even though the pitching-moment coefficient remains comparable.

To further enhance the reliability of the numerical predictions, additional work on the correlations and parameter settings within the γ-Reθ transition model is recommended, particularly with respect to identifying the minimum number of mesh elements needed to maintain solution stability and accuracy. Finally, an experimental investigation is advised to validate and complement the CFD findings presented in this research.

References

1. Anderson J.D. *Introduction to Flight.* McGraw Hill, 2015, 928 p.
2. Sadraey M.H. *Design of Unmanned Aerial Systems*. John Wiley & Sons Ltd., 2020, 637 p.
3. Ooda I.J. Numerical investigation of aerodynamic characteristics of supercritical RAE2822 airfoil with gurney flap. *Journal of Engineering*, 2022, vol. 28, no. 6, pp. 1–14. https://doi.org/10.31026/j.eng.2022.06.01
4. Abed A.T.M.A. Computational analysis of SD7037 airfoil with plain flap. *Archive of Mechanical Engineering*, 2024, vol. 71, no. 4, pp. 561–580. https://doi.org/10.24425/ame.2024.151337
5. Frolov V.A., Dong W.M. Lift investigation of airfoil with flaps. *Applied Mechanics and Materials*, 2014, vol. 627, pp. 89–92. https://doi.org/10.4028/www.scientific.net/AMM.627.89
6. Todorov M.D. Aerodynamic characteristics of airfoil with single slotted flap for light airplane wing. *Proc. of the International Conference of Scientific Paper AFASES*, 2015, pp. 509–515.
7. Srivastava M.S., Aditya C.V.N. Analysis on NACA 2412 airfoil for UAV based on high-lift devices. *International Journal of Engineering Applied Sciences and Technology*, 2016, vol. 1, no. 6, pp. 13–16.
8. Pracheta D., Anup A., Shahriar A., Saifur R.B. Computational study on effect of flap deflection on NACA 2412 airfoil in subsonic flow. *Applied Mechanics and Materials*, 2016, vol. 829, pp. 9–14. https://doi.org/10.4028/www.scientific.net/AMM.829.9
9. Thejaraju R., Dana J., Niveditha V.R., Vishal S., Sumanthran A. A study of airfoil flap deflection angle using CFD simulation techniques. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, 2019, vol. 9, special issue, pp. 125–132.
10. Michna J., Rogowski K., Bangga G., Hansen M.O.L. Accuracy of the Gamma Re-theta transition model for simulating the DU-91-W2-250 airfoil at high reynolds numbers. *Energies*, 2021, vol. 14, no. 24, pp. 8224. https://doi.org/10.3390/en14248224
11. Ramanan G., Radha Krishnan P., Ranjan H.M. An aerodynamic performance study and analysis of SD7037 fixed wing UAV airfoil. *Materials Today: Proceedings*, 2021, vol. 47, pp. 2547–2552. https://doi.org/10.1016/j.matpr.2021.05.051
12. Abhishek Vinod G., Jothi T.J.S. Effect of flap deflection angle on flow characteristics of aerofoil. *IOP Conference Series: Materials Science and Engineering*, 2021, vol. 1189, no. 1, pp. 012039. https://doi.org/10.1088/1757-899X/1189/1/012039
13. Patel V., Rathod V., Patel C. Numerical investigation of M21 aerofoil and effect of plain flapper at various angle of attack. *Journal of Physics: Conference Series*, 2021, vol. 2070, pp. 012153. https://doi.org/10.1088/1742-6596/2070/1/012153
14. Ma Q., Wang J., Zhang Y., Liu X. Parametric research and aerodynamic performance analysis of wind turbine airfoil with added flap. *AIP Advances*, 2022, vol. 12, no. 10, pp. 105010. https://doi.org/10.1063/5.0109713
15. Salam N., Tarakka R., Jalaluddin, Iriansyah D., Ihsan M. The effects of flap angles on the aerodynamic performances of a homebuilt aircraft wing model. *International Journal of Mechanical Engineering and Robotics Research*, 2022, vol. 11, no. 12, pp. 908–914. https://doi.org/10.18178/ijmerr.11.12.908-914
16. Radha Krishnan P., Mukesh R., Hasan I., Booma Devi P., Alebachew W.A. Comparative study and aerodynamic analysis of rectangular wing using high-lift systems. *International Journal of Aerospace Engineering*, 2023, vol. 2023, pp. 5813557. https://doi.org/10.1155/2023/5813557
17. Soman S., Melvin A., Variya D., Murallidharan J.S. Impact of numerical settings on simulation performance of the Gamma-Re-Theta model for 2D and 3D bluff body transitional flow cases using OpenFOAM. *Lecture Notes in Mechanical Engineering*, 2023, pp. 31–36. https://doi.org/10.1007/978-981-19-6970-6_6
18. Xiao Y., Wang F., Sun P. Numerical simulation of two-dimensional plain flap ground effect based on NACA4412 airfoil. *Journal of Physics: Conference Series*, 2024, vol. 2756, no. 1, pp. 012049. https://doi.org/10.1088/1742-6596/2756/1/012049

Литература

1. Anderson J.D. Introduction to Flight. McGraw Hill, 2015. 928 p.
2. Sadraey M.H. Design of Unmanned Aerial Systems. John Wiley & Sons Ltd., 2020. 637 p.
3. Ooda I.J. Numerical investigation of aerodynamic characteristics of supercritical RAE2822 airfoil with gurney flap // Journal of Engineering. 2022. V. 28. N 6. P. 1–14. https://doi.org/10.31026/j.eng.2022.06.01
4. Abed A.T.M.A. Computational analysis of SD7037 airfoil with plain flap // Archive of Mechanical Engineering. 2024. V. 71. N 4. P. 561–580. https://doi.org/10.24425/ame.2024.151337
5. Frolov V.A., Dong W.M. Lift investigation of airfoil with flaps // Applied Mechanics and Materials. 2014. V. 627. P. 89–92. https://doi.org/10.4028/www.scientific.net/AMM.627.89
6. Todorov M.D. Aerodynamic characteristics of airfoil with single slotted flap for light airplane wing // Proc. of the International Conference of Scientific Paper AFASES. 2015. P. 509–515.
7. Srivastava M.S., Aditya C.V.N. Analysis on NACA 2412 airfoil for UAV based on high-lift devices // International Journal of Engineering Applied Sciences and Technology. 2016. V. 1. N 6. P. 13–16.
8. Pracheta D., Anup A., Shahriar A., Saifur R.B. Computational study on effect of flap deflection on NACA 2412 airfoil in subsonic flow // Applied Mechanics and Materials. 2016. V. 829. P. 9–14. https://doi.org/10.4028/www.scientific.net/AMM.829.9
9. Thejaraju R., Dana J., Niveditha V.R., Vishal S., Sumanthran A. A study of airfoil flap deflection angle using CFD simulation techniques // International Journal of Mechanical and Production Engineering Research and Development (IJMPERD). 2019. V. 9. Special Issue. P. 125–132.
10. Michna J., Rogowski K., Bangga G., Hansen M.O.L. Accuracy of the Gamma Re-theta transition model for simulating the DU-91-W2-250 airfoil at high reynolds numbers // Energies. 2021. V. 14. N 24. P. 8224. https://doi.org/10.3390/en14248224
11. Ramanan G., Radha Krishnan P., Ranjan H.M. An aerodynamic performance study and analysis of SD7037 fixed wing UAV airfoil // Materials Today: Proceedings. 2021. V. 47. P. 2547–2552. https://doi.org/10.1016/j.matpr.2021.05.051
12. Abhishek Vinod G., Jothi T.J.S. Effect of flap deflection angle on flow characteristics of aerofoil // IOP Conference Series: Materials Science and Engineering. 2021. V. 1189. P. 012039. https://doi.org/10.1088/1757-899X/1189/1/012039
13. Patel V., Rathod V., Patel C. Numerical investigation of M21 aerofoil and effect of plain flapper at various angle of attack // Journal of Physics: Conference Series. 2021. V. 2070. P. 012153. https://doi.org/10.1088/1742-6596/2070/1/012153
14. Ma Q., Wang J., Zhang Y., Liu X. Parametric research and aerodynamic performance analysis of wind turbine airfoil with added flap // AIP Advances. 2022. V. 12. N 10. P. 105010. https://doi.org/10.1063/5.0109713
15. Salam N., Tarakka R., Jalaluddin, Iriansyah D., Ihsan M. The effects of flap angles on the aerodynamic performances of a homebuilt aircraft wing model // International Journal of Mechanical Engineering and Robotics Research. 2022. V. 11. N 12. P. 908–914. https://doi.org/10.18178/ijmerr.11.12.908-914
16. Radha Krishnan P., Mukesh R., Hasan I., Booma Devi P., Alebachew W.A. Comparative study and aerodynamic analysis of rectangular wing using high-lift systems // International Journal of Aerospace Engineering. 2023. V. 2023. P. 5813557. https://doi.org/10.1155/2023/5813557
17. Soman S., Melvin A., Variya D., Murallidharan J.S. Impact of numerical settings on simulation performance of the Gamma-Re-Theta model for 2D and 3D bluff body transitional flow cases using OpenFOAM // Lecture Notes in Mechanical Engineering. 2023. P. 31–36. https://doi.org/10.1007/978-981-19-6970-6_6
18. Xiao Y., Wang F., Sun P. Numerical simulation of two-dimensional plain flap ground effect based on NACA4412 airfoil // Journal of Physics: Conference Series. 2024. V. 2756. N 1. P. 012049. https://doi.org/10.1088/1742-6596/2756/1/012049

19. Parluhutan Y.M., Fahrudin F., Rhakasywi D. Investigating the impact of plain flap as lift enhancement on symmetrical airfoils. *International Journal of Marine Engineering Innovation and Research*, 2024, vol. 9, no. 1, pp. 97–104. https://doi.org/10.12962/j25481479.v9i1.19848
20. Ali A.H. Computational method for unsteady motion of two-dimensional airfoil. *Journal of Engineering*, 2008, vol. 14, no. 4, pp. 1–17. https://doi.org/10.31026/j.eng.2008.04.21
21. Ali A.H. Aerodynamic characteristics of a rectangular wing using non-linear vortex ring method. *Journal of Engineering*, 2017, vol. 23, no. 4, pp. 1–17. https://doi.org/10.31026/j.eng.2017.04.08
22. Selig M.S., Lyon C.A., Giguère P., Ninham C.N., Guglielmo J.J. *Summary of Low-Speed Airfoil Data*, V. 2. Virginia Beach, Virginia, SoarTech Publications, 1996.
23. Menter F.R. Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA Journal*, 1994, vol. 32, no. 8, pp. 1598–1605. https://doi.org/10.2514/3.12149
24. Menter F.R., Langtry R.B., Likki S.R., Suzen Y.B., Huang P.G., Völker S. A correlation-based transition model using local variables. *Proc. of the ASME Turbo Expo*, 2004, pp. 1–11.
25. Versteeg H.K., Malalasekera W. *An Introduction to Computational Fluid Dynamics: The Finite Volume Method*. Pearson Education, 2007, 520 p.

19. Parluhutan Y.M., Fahrudin F., Rhakasywi D. Investigating the impact of plain flap as lift enhancement on symmetrical airfoils // International Journal of Marine Engineering Innovation and Research. 2024. V. 9. N 1. P. 97–104. https://doi.org/10.12962/j25481479.v9i1.19848
20. Ali A.H. Computational method for unsteady motion of two-dimensional airfoil // Journal of Engineering. 2008. V. 14. N 4. P. 1–17. https://doi.org/10.31026/j.eng.2008.04.21
21. Ali A.H. Aerodynamic characteristics of a rectangular wing using non-linear vortex ring method // Journal of Engineering. 2017. V. 23. N 4. P. 1–17. https://doi.org/10.31026/j.eng.2017.04.08
22. Selig M.S., Lyon C.A., Giguère P., Ninham C.N., Guglielmo J.J. Summary of Low-Speed Airfoil Data, V. 2. Virginia Beach, Virginia: SoarTech Publications, 1996.
23. Menter F.R. Two-equation eddy-viscosity turbulence models for engineering applications // AIAA Journal. 1994. V. 32. N 8. P. 1598–1605. https://doi.org/10.2514/3.12149
24. Menter F.R., Langtry R.B., Likki S.R., Suzen Y.B., Huang P.G., Völker S. A correlation-based transition model using local variables // Proc. of the ASME Turbo Expo. 2004. P. 1–11.
25. Versteeg H.K., Malalasekera W. An Introduction to Computational Fluid Dynamics: The Finite Volume Method. Pearson Education, 2007. 520 p.

Authors

Ahmed Tawfeeq Mustafa Ali Abed — Magister, Lecturer Assistant, Aeronautical Engineering Department, College of Engineering, University of Baghdad, Baghdad, 10071, Iraq, sc 59392733600, https://orcid.org/0000-0002-8025-9367, a.t.abed@coeng.uobaghdad.edu.iq

Azad A. Hussein — PhD, Lecturer, Aeronautical Engineering Department, College of Engineering, University of Baghdad, Baghdad, 10071, Iraq, sc 57297985700, https://orcid.org/0000-0001-9792-7774, azad.a@coeng.uobaghdad.edu.iq

Saba Hafedh Mahdi — Magister, Lecturer Assistant, Aeronautical Engineering Department, College of Engineering, University of Baghdad, Baghdad, 10071, Iraq, https://orcid.org/0009-0009-9479-0536, saba.hafedh@coeng.uobaghdad.edu.iq

Авторы

Али Ахмед Тауфик Мустафа Абед — магистр, помощник преподавателя, Отделение авиационной техники Инженерного колледжа, Багдадский университет, Багдад, 10071, Ирак, sc 59392733600, https://orcid.org/0000-0002-8025-9367, a.t.abed@coeng.uobaghdad.edu.iq

А. Азад Хусейн — PhD, преподаватель, Отделение авиационной техники Инженерного колледжа, Багдадский университет, Багдад, 10071, Ирак, sc 57297985700, https://orcid.org/0000-0001-9792-7774, azad.a@coeng.uobaghdad.edu.iq

Хафед Саба Махди — магистр, помощник преподавателя, Отделение авиационной техники Инженерного колледжа, Багдадский университет, Багдад, 10071, Ирак, https://orcid.org/0009-0009-9479-0536, saba.hafedh@coeng.uobaghdad.edu.iq

Received 04.05.2025
Approved after reviewing 10.10.2025
Accepted 12.11.2025

Статья поступила в редакцию 04.05.2025
Одобрена после рецензирования 10.10.2025
Принята к печати 12.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2025-25-6-1220-1228
УДК 53.072; 53:004

Теплофизические аспекты фазовых превращений охлаждаемого расплава высокотемпературного кориума в ядерных энергетических установках

Николай Васильевич Пилипенко[1], Артём Олегович Коптюхов[2]✉

[1] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация
[2] НИЦ «Курчатовский институт» — ПИЯФ, Гатчина, 188300, Российская Федерация

[1] pilipenko38@mail.ru, https://orcid.org/0000-0001-9328-3166
[2] koptyukhov_ao@pnpi.nrcki.ru✉, https://orcid.org/0000-0002-9660-3601

Аннотация
Введение. Рассмотрены вопросы трехмерного моделирования теплового взаимодействия между расплавом активной зоны и устройством локализации расплава (ловушкой) в условиях тяжелой аварии на атомной электростанции. Предложена оптимизированная конфигурация заполнения устройства локализации жертвенным материалом. Выполнены интеграции усредненного по Рейнольдсу уравнения Навье–Стокса, численного решения уравнения теплопроводности и модели динамики уровня двух несмешивающихся жидкостей, что позволяет одновременно учитывать турбулентное течение в жидких компонентах расплава, изменение границы расплавляющегося жертвенного материала (задача Стефана) и стратификацию расплавов с инверсией. **Метод.** Расчеты произведены в три последовательных этапа. На этапе первом этапе моделируется плавление жертвенного материала, на втором — стратификация слоев, а на третьем — теплоперенос после стратификации. На основе полученных данных формируется оптимальная конфигурация заполнения ловушки. **Основные результаты.** Представлено детальное объемное распределение температур на протяжении всех трех этапов, расчет распределения теплового потока на поверхности стенок ловушки и определение максимальной толщины подплавления корпуса, обусловленной интенсивным тепловым взаимодействием. Сравнение трехмерных расчетов с аналогичными двумерными исследованиями показало, что трехмерное моделирование более точно фиксирует характерные временные интервалы застывания и последующего расплавления металла. **Обсуждение.** Продемонстрированы преимущества предложенного подхода в сравнении с существующими методами. Показана его применимость при проектировании и оптимизации устройств локализации расплава, представлены перспективы дальнейшего развития модели, включая учет химических реакций и адаптацию к другим типам реакторов. Полученные данные убедительно подтверждают, что принятая конфигурация обладает потенциалом для существенного продления периода эффективного противодействия последствиям тяжелой аварии на атомных электростанциях.

Ключевые слова
тигельная ловушка, кориум, локализации расплава активной зоны, аварии атомных станций, жертвенный материал

Ссылка для цитирования: Пилипенко Н.В., Коптюхов А.О. Теплофизические аспекты фазовых превращений охлаждаемого расплава высокотемпературного кориума в ядерных энергетических установках // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1220–1228. doi: 10.17586/2226-1494-2025-25-6-1220-1228

Thermophysical aspects of phase transformations of the cooled melt of high-temperature corium in nuclear power plants

Nikolay V. Pilipenko[1], Artem O. Koptyukhov[2]✉

[1] ITMO University, Saint Petersburg, 197101, Russian Federation
[2] NRC “Kurchatov Institute” — PNPI, Gatchina, 188300, Russian Federation

[1] pilipenko38@mail.ru, https://orcid.org/0000-0001-9328-3166
[2] koptyukhov_ao@pnpi.nrcki.ru✉, https://orcid.org/0000-0002-9660-3601

© Пилипенко Н.В., Коптюхов А.О., 2025

Abstract
This study addresses three-dimensional modeling of the thermal interaction between the core melt and the melt localization device (trap) during a severe accident at a nuclear power plant. An optimized configuration for filling the localization device with sacrificial material is proposed. The calculations incorporate the Reynolds-averaged Navier–Stokes equations, numerical solutions of the heat conduction equation, and a two-fluid interface dynamics model, enabling simultaneous consideration of turbulent flow within the liquid phases, the moving boundary of the melting sacrificial material (Stefan problem), and stratification with inversion. The analysis proceeds in three consecutive stages. The first stage models the melting of the sacrificial material; the second simulates the stratification of layers; the third evaluates heat transfer after stratification. Based on the results, an optimal filling configuration for the trap is developed. The study presents detailed volumetric temperature distributions throughout all three stages, the heat flux distribution on the trap walls, and the maximum thickness of the melted shell caused by intense thermal interaction. Comparison between three-dimensional simulations and similar two-dimensional studies demonstrates that 3D modeling more accurately captures the characteristic timing of solidification and subsequent melting processes. The advantages of the proposed approach over existing methods are highlighted. Its applicability for designing and optimizing melt localization devices is shown, and prospects for future development are discussed, including incorporating chemical reactions and adapting the model to other reactor types. The data convincingly suggest that the adopted configuration has significant potential to extend the period during which effective mitigation of severe accident consequences at nuclear power plants can be maintained.

Keywords
crucible trap, corium, core-melt localization, nuclear power-plant accidents, sacrificial material

For citation: Pilipenko N.V., Koptyukhov A.O. Thermophysical aspects of phase transformations of the cooled melt of high-temperature corium in nuclear power plants. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1220–1228 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1220-1228

Введение

Нарушение теплообмена при тяжелых авариях на атомных электростанциях (АЭС) является причиной расплавления активной зоны, в результате чего образуется высокотемпературный кориум. Основу кориума составляют неэвтектические смеси металлов и их оксидов, происходящие из тепловыделяющих элементов и кронштейнов [1]. Поскольку кориум способен проплавить корпус реактора и выйти в окружающую среду, это создает угрозу тяжелых экологических последствий.

Современные реакторы АЭС оснащены системами для предотвращения последствий аварий. Эти системы включают концепции удержания кориума (расплава топлива) либо в корпусе реактора, либо в специальном устройстве локализации расплава под активной зоной [2]. Для отвода остаточного радиационного тепла из ванны с кориумом используется охлаждение стенок корпуса. Чтобы снизить температурное воздействие на корпус и уменьшить выделение водорода при контакте пара с цирконием, применяют теплопоглощающий жертвенный материал на основе оксидов железа и алюминия [3]. Тепло поглощается за счет плавления этого материала и эндотермических химических реакций с кориумом.

При поступлении в систему локализации расплав кориума формирует бислойную оксидно-металлическую ванну. Согласно фазовой диаграмме [4], на начальном этапе плотность металлической фазы уступает плотности оксидной, что обуславливает ее верхнее положение. Тем не менее, по мере достижения критической плотности в результате взаимодействия расплавленного теплопоглощающего материала с оксидными компонентами, металлическая фаза мигрирует вверх.

Одним из факторов, лимитирующих корпусное удержание расплава, выступает формирование металлической корки. В пределах данной зоны происходит фокусировка максимального теплового потока [5]. Эффективность удержания расплава находится в прямой зависимости от интенсивности теплоотдачи на охлаждаемых водой поверхностях корпуса. Критерием успешного удержания является достижение критического теплового потока. Естественная конвекция тепла играет существенную роль в реализации данных стратегий, поскольку она напрямую определяет величину тепловой нагрузки.

Изучение кориума было предметом многочисленных международных экспериментальных исследований. Работы, проведенные UCLA (1996), ACOPO (1997), RASPLAV (1998), COPO (1999), BALI (1999), SIMECO (2002) и COPRA (2016), позволили получить ценные данные о теплопередаче в конструкциях, моделирующих бассейн расплава кориума [6]. Следует отметить, что из-за высокой стоимости и технических сложностей, многие из этих экспериментальных установок были мелкомасштабными и характеризовались более низкими числами Рэлея [7]. Основной целью этих экспериментов являлась верификация численных моделей. В настоящее время, благодаря заметному прогрессу в численных методах и росту вычислительных мощностей, методы моделирования турбулентных течений становятся все более применимыми. Однако существуют и ограничения. Дин и Нургалиев (1997) указали на неспособность двумерной модели турбулентности k-ε точно описывать тепловые потоки при высоких числах Рэлея [8]. Сравнение моделей турбулентности k-ε и больших вихрей (Large Eddy Simulation, LES), основанное на эксперименте BALI [9], показало, что k-ε модель демонстрирует низкую точность в описании термической стратификации, в то время как LES-модель завышает теплоотдачу на боковой стенке. Исследования, описанные в работе [10], подтвердили хорошее согласование высокоразрешающих методов Computational Fluid Dynamics, таких как моделирование LES и прямое численное моделирование, с эксперимен-

тальными данными. Тем не менее, эти модели пока не учитывают взаимодействия с жертвенным материалом.

Предложена модель, которая описывает три этапа физических процессов в кориуме внутри системы локализации расплава.

Гравитационная сегрегация. На начальном этапе расплав расслаивается по плотности: под действием силы тяжести более тяжелые оксидные компоненты опускаются вниз. Одновременно происходит теплофизическое взаимодействие с жертвенным материалом.

Плотностная инверсия. В процессе смешивания жертвенного материала с кориумом происходит перераспределение фаз: оксидные компоненты перемещаются вверх.

Кристаллизация. На финальном этапе расплав охлаждается и кристаллизуется благодаря естественной конвекции и внешнему охлаждению корпуса тигельной ловушки, в которой локализован расплав.

Модель базируется на решении уравнения Навье–Стокса с использованием k-ε модели и задачи Стефана. Задача Стефана с изменяющимися границами раздела фаз сводится к задаче тепло-массообмена в вещественной среде кориумной ловушки. В этой среде присутствуют подвижные источники теплоты: эндотермическое разложение и плавление жертвенного материала, а также тепло экзотермических реакций с цирконием, хромом и ураном в расплаве активной зоны [11].

В настоящей работе представлены результаты разработки и валидации трехмерной модели теплофизических процессов, происходящих при фазовых превращениях и тепловом взаимодействии в системе локализации расплава кориума при тяжелой аварии на АЭС, с целью выявления оптимальных конструктивных решений для повышения безопасности и продления периода противодействия аварийным последствиям.

Метод

Геометрически-конструктивная система. Геометрическая схема исследуемой ванны расплава (рис. 1) описывает цилиндрическую емкость радиусом R = 2900 мм и высотой H = 2400 мм, с толщиной стенки 100 мм.

В качестве жертвенного материала используется конфигурация из треугольных пирамид с ребром 210 мм и высотой 50 мм. Эти элементы формируют шестигранные сборки, в которых каждый ряд имеет два отсутствующих элемента. Последовательные ряды смещены друг относительно друга на 60°, что обеспечивает спиральное движение кориума по сборкам. Общее количество таких элементов в ловушке локализации достигает 26 600 единиц, а их суммарная масса составляет 82 т.

Ванна расплавов окружена водяной оболочкой, поддерживающей температуру около 303 К. Для обеспечения равномерного распределения воды в верхней части устройства локализации установлен кольцевой коллектор. Он включает один входной и 32 выходных патрубка, подключенных к кольцевому каналу. Нижняя часть устройства оборудована выходным патрубком для отвода воды в бак.

Моделирование тепловых процессов. В работе [12] предложена математическая модель, которая описывает динамику расплава как вязкую жидкость с переменной плотностью, зависящей от температуры и состава. При применении модели была использована классическая система уравнений для описания расплава — уравнение Навье–Стокса, уравнение энергии и модели турбулентности — для моделирования вязкой жидкости с переменной плотностью и тепловыми эффектами фазовых переходов. Однако для более точного и полного описания процессов плавления, кристаллизации и межфазных взаимодействий стало очевидно, что необходимо ввести более детализированный учет интерфейсов и фазовых границ.

Для повышения точности модели в настоящей работе дополнительно использован фазово-полевой подход, который позволяет динамически описывать интерфейсы между различными фазами через переменную φ. Этот подход является мощным инструментом для моделирования сложных фазовых превращений, особенно когда важны тонкие структуры и взаимодействия на микроскопическом уровне — что критично при моделировании кристаллизации, плавления и образования новых фаз.

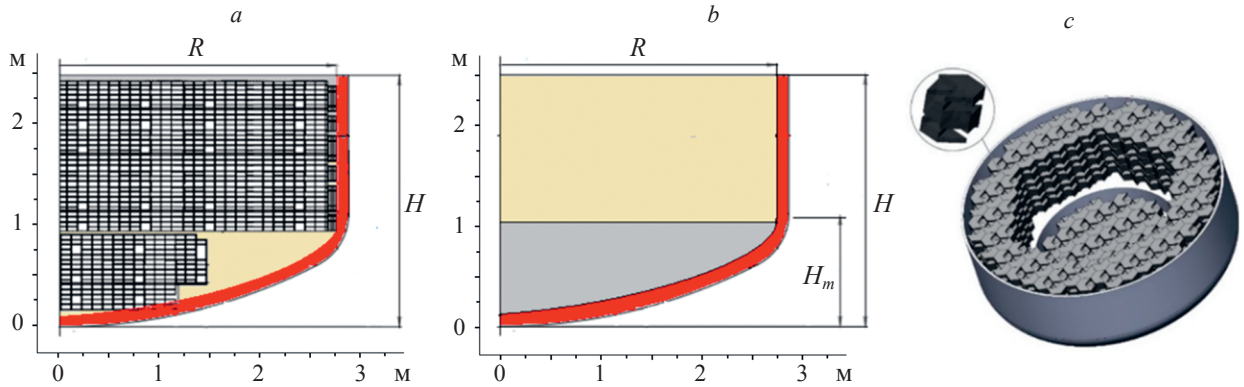

Рис. 1. Габаритный чертеж ванны расплава: до инверсии (*a*); после инверсии (*b*); общий трехмерный вид (*c*). H_m = 0,9 м — высота слоя металлической компоненты после инверсии

Fig. 1. Overall drawing of the melt bath: before inversion (*a*); after inversion (*b*); overall three-dimensional view (*c*). H_m = 0.9 m — height of the metallic component layer after inversion

Для учета плавления жертвенного материала и кристаллизации кориума применена фазовая переменная φ, которая принимает значения φ = 0 в одной фазе и φ = 1 в другой. Эволюция φ смоделирована с помощью фазово-полевой модели, например, модифицированного уравнения Алена–Кана, что позволяет описывать распределение интерфейса и толщину краев инверсии. Пример обобщенной формы фазового уравнения [13]:

$$\frac{\partial \varphi}{\partial t} + \mathbf{u}\nabla\varphi = \frac{\partial f}{\partial \varphi} - \varepsilon^2\nabla^2\varphi,$$

где φ — соотношение твердой фазы к жидкой, безразмерная (от 0 до 1); t — время; $\mathbf{u}$ — скорость потока, в котором переносятся фазы, задается отдельно решением уравнения Навье–Стокса; ∇φ — градиент φ; ε — параметр, задающий толщину границы; $\nabla^2\varphi$ — Лапласиан φ.

Связь φ с физическими величинами реализуется через зависимость плотности $\rho(\varphi, T) = \varphi\rho_2 + (1 - \varphi)\rho_1$, где ρ_1 и ρ_2 — плотности фаз. Латентное тепло вводится через член $L = d\varphi/dt$ в уравнение теплопереноса, отображая энергию, высвобождающуюся или поглощающуюся при фазовых превращениях.

Реакция между жертвенным материалом и кориумом сопровождается изменением состава и, как следствие, плотности смеси. В результате происходит инверсия стратификации, при которой менее плотная оксидная компонента расплава перемещается вверх под действием силы плавучести. Для описания динамики этого процесса используется фазово-полевая модель, что и при моделировании плавления-кристаллизации, т. е. уравнение переноса скалярной функции уровня φ.

Изменение температуры при естественной конвекции в расплаве учитывается конвективным членом ($\mathbf{v}\cdot\nabla T$) уравнения теплопереноса. Для моделирования естественной и вынужденной конвекций применялась усредненная по скорости система уравнений Навье–Стокса и сохранения массы.

$$\rho\left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v}\nabla\mathbf{v}\right) = -\nabla p + \nabla\mu(\nabla\mathbf{v} + \nabla\mathbf{v}^T) + \rho\mathbf{g} + F,\ \nabla\mathbf{u} = 0,$$

где ρ — плотность среды; p — давление; μ — вязкость; $\mathbf{g}$ — вектор ускорения свободного падения; F — составляющие силы поверхностного натяжения. Решение уравнения базируется на использовании замыкания полуэмпирической модели турбулентности согласно стандартной k-ε модели [14].

Если в некоторой зоне плотность снижается (при окислении элементов расплава), в уравнении Навье–Стокса появляется меньший член $\rho\mathbf{g}$, а значит, вес жидкости в этой зоне уменьшается. В результате слой получает дополнительный подъемный импульс, а нижний — опускается.

Вычислительный анализ тепловых процессов в устройстве локализации расплава был выполнен методом конечных элементов. Теплообмен с участием расплава активной зоны моделировался численно с использованием мультифизического пакета COMSOL Multiphysics. Для повышения точности расчетов в области пограничного слоя применялась анизотропная сетка, шаг которой уменьшался на границах расчетной области. Минимальный размер элемента сетки для наибольшей расчетной области бассейна расплава в корпусе охлаждаемой ловушки 2,74 × 2,85 м составлял 0,003 м, максимальный — 0,119 м при общем количестве 4 076 212 сеточных элементов.

Моделирование проводилось в три этапа (рис. 2).

Этап 1. Изучение тепловых процессов при плавлении жертвенного материала. Расчет остаточной мощности осуществлялся по экспоненциальному закону, исходя из номинальной мощности реактора 1000 МВт.

Этап 2. Стратификация оксидного и металлических слоев.

Этап 3. Моделирование теплопереноса в инверсном бассейне. На этом этапе учитывалось затвердевание корки кориума и дополнительное граничное условие охлаждения бассейна сверху — впрыск воды, сопровождающийся пузырьковым кипением.

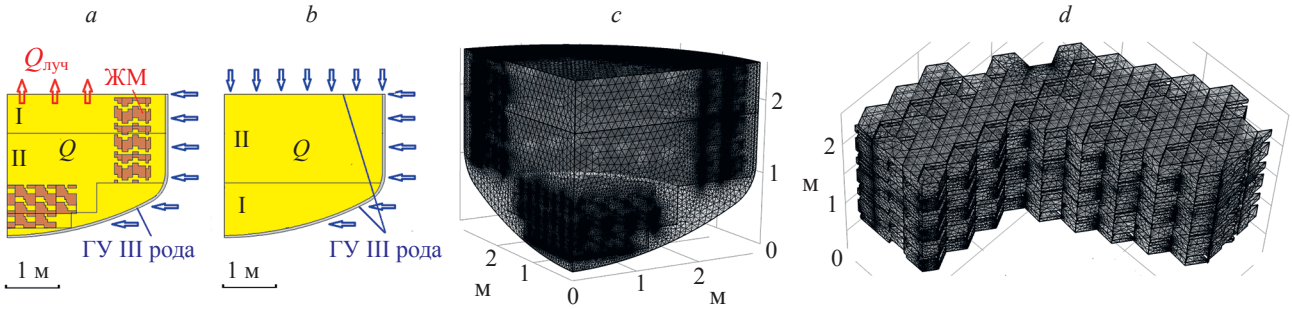

Рис. 2. Расчетная модель взаимодействия расплава кориума с корзиной жертвенного материала при плавлении жертвенного материала (*a*), кристаллизация расплавов после инверсии (*b*) и сетка конечных элементов (*c*, *d*).

ЖМ — жертвенный материал; ГУ — граничное условие; I — металлические компоненты; II — оксидные компоненты; III — конвективный теплообмен; Q — объемные тепловыделения; $Q_{луч}$ — тепловой поток излучением

Fig. 2. Computational model of the interaction between the core melt and the sacrificial material basket during melting (*a*), crystallization of melts after inversion (*b*), and the finite element mesh (*c*, *d*)

ЖМ — sacrificial material; ГУ — boundary condition; I — metallic components; II — oxide components; III — convective heat transfer; Q — volumetric heat sources; $Q_{луч}$ — radiative heat flux

Основные результаты

Результаты численного моделирования включают объемное распределение температур на всех трех этапах, а также идентификацию фазовых областей, связанных с плавлением жертвенного материала и последующим затвердеванием кориума. Динамика температурных полей и фазовых состояний в ловушке, охватывающая весь цикл от плавления жертвенного материала до затвердевания кориума, показана на рис. 3.

В течение 60 мин жертвенный материал полностью расплавляется. Мощность остаточных тепловыделений в кориуме заметно снижается, демонстрируя экспоненциальное затухание во времени. Одновременно с этим наблюдается затвердевание оксидных и металлических компонентов (рис. 4).

На начальном этапе (рис. 4, *a*) наблюдается охлаждение кориума несмотря на остаточное радиационное тепловыделение. Это обусловлено затратами тепловой энергии на нагрев и плавление жертвенного материала. При медленном прогреве кориумом поверхности холодного жертвенного материала образуется тонкий слой кристаллической корки из тугоплавких компонентов расплава (рис. 4, *d*). Это связано с высокой теплоемкостью жертвенного материала. Низкая теплопроводность образовавшейся корки ограничивает тепловой

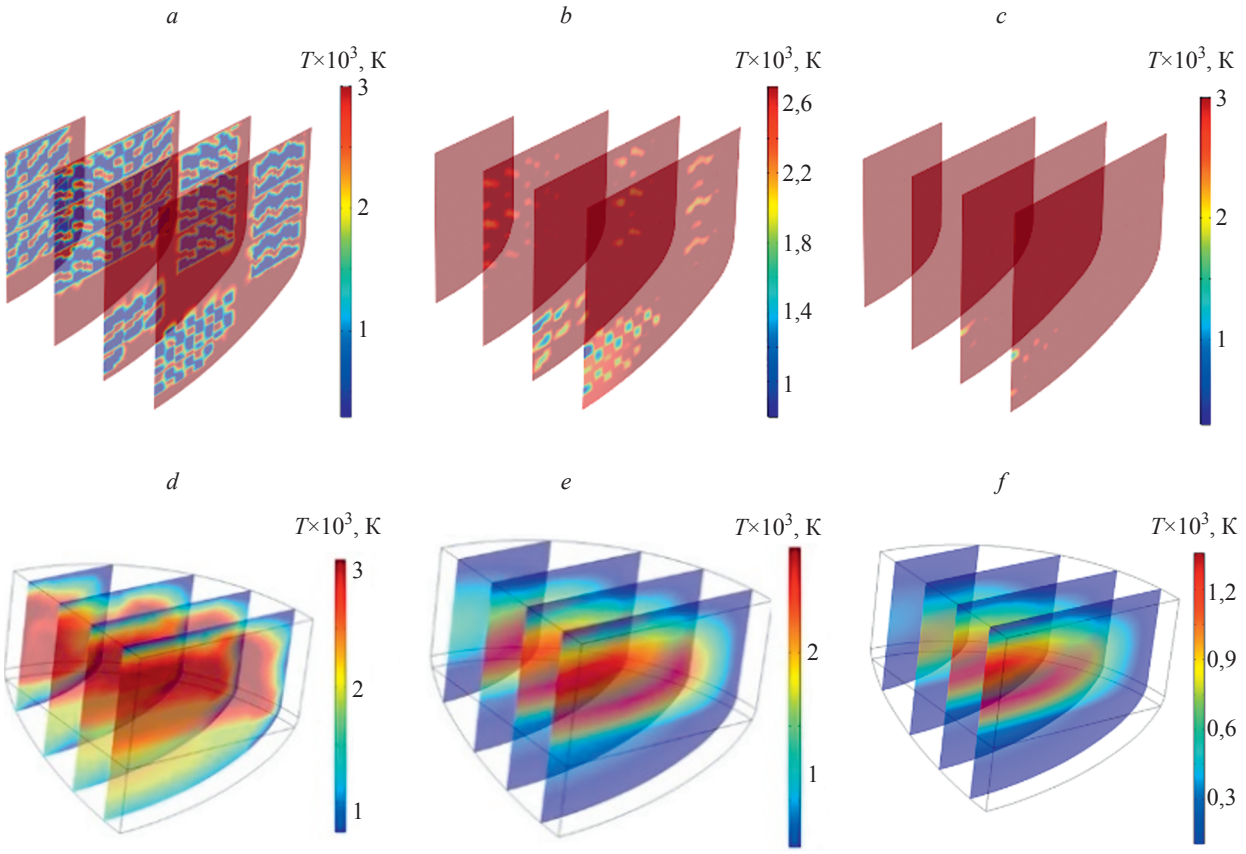

Рис. 3. Температурные поля в сечениях бассейна расплава в течении: 1 мин (*a*); 20 мин (*b*): 60 мин (*c*); 10 ч (*d*); 100 ч (*e*); 200 ч (*f*)

Fig. 3. Temperature fields in cross-sections of the melt pool at durations of: 1 minute (*a*); 20 minutes (*b*); 60 minutes (*c*); 10 hours (*d*); 100 hours (*e*); 200 hours (*f*)

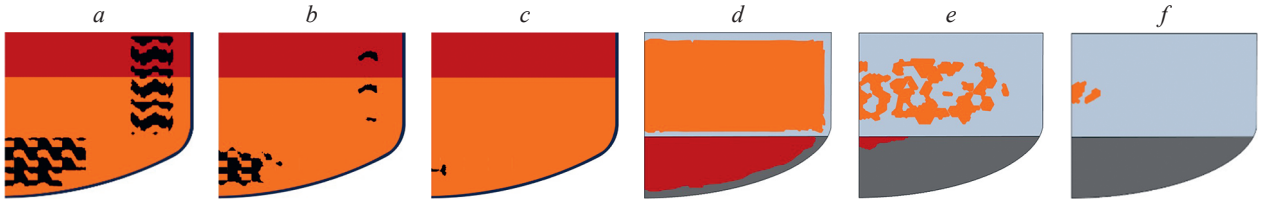

Рис. 4. Фазовый состав системы в течении: 1 мин (*a*); 20 мин (*b*); 60 мин (*c*); 10 ч (*d*); 30 ч (*e*); 91 ч (*f*).
Твердая фаза представлена черной областью, а жидкие фазы — металлической компонентой (красная область) и оксидной компонентой (оранжевая область)

Fig. 4. Phase composition of the system. The solid phase is shown as the black region, while the liquid phases are the metallic component (red region) and the oxide component (orange region)

эффект протекающих химических реакций и теплоотдачу. Параллельно наблюдается возрастание влияния химических реакций на процессы тепло-массообмена и фазовые переходы в системе. По мере дальнейшего прогрева происходит совместное плавление корки и жертвенного материала (рис. 4, *e*).

Примерно через 60 мин (рис. 4, *c*) достигается критическая масса расплавленного жертвенного материала в бассейне. Плотность металла становится ниже плотности смеси оксидных компонентов, что инициирует инверсию стратификации. Тепловая стратификация и поток тепла в слоях кориума влияют на их стабильность и турбулентацию. Этот теплоперенос обусловлен термогравитационным механизмом генерации турбулентности. Устойчивое расположение оксидного слоя позволяет безопасно подавать воду на поверхность расплава (рис. 4, *f*).

В условиях неизотермичности формируется гарнисажная корка, которая затрудняет теплообмен (рис. 5). Анализ графика показывает, что оксидная корка образуется практически мгновенно, и уже через 30 мин после инверсии достигается состояние, близкое к стационарному. Через 60 мин после инверсии температура оксидных компонентов начинает снижаться, что приводит к постепенному затвердеванию кориума. Температура затвердевания оксидной компоненты достигается через 23 ч, после чего интенсивность затвердевания возрастает. Аналогично, интенсивность затвердевания расплавленного металла активной зоны увеличивается через 17 ч. Затвердевание бассейна происходит достаточно быстро: через 44 ч более теплопроводная металлическая компонента полностью затвердевает. Полное затвердевание бассейна завершается через 91 ч (рис. 4, *f*).

В рамках проведенных исследований установлено, что тепловые потоки от кориума к стенкам ловушки находятся в диапазоне от 0,2 до 0,6 МВт/м^2. Эти значения существенно ниже критической тепловой нагрузки [15] на внешней охлаждаемой поверхности корпуса устройства, что гарантирует достаточный запас безопасности.

При достижении максимальной тепловой нагрузки оптимальным является расход воды в охлаждающей рубашке ловушки $G = 20$ кг/с. В стационарном режиме, при таком расходе, максимальная температура стенки составляет $t_ст = 2330$ К. Гидравлический расчет

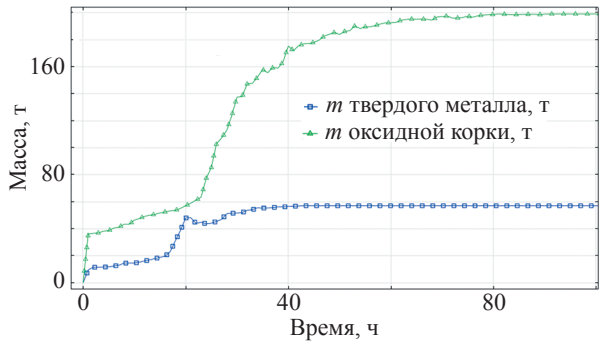

Рис. 5. Масса (*m*) оксидной корки и затвердевшего металла в бассейне расплава после инверсии

Fig. 5. Mass (*m*) of the oxide crust and solidified metal in the melt pool after inversion

подтверждает, что для обеспечения данного режима необходимо давление воды на входе 1,25 атм. Важно отметить, что дальнейшее увеличение расхода воды не приводит к существенному снижению температуры стенки.

Результаты анализа демонстрируют, что при расходе воды более 20 кг/с достигается заметное улучшение равномерности температурного поля по сравнению с режимами, использующими меньший расход (рис. 6).

В результате воздействия тепловых потоков наблюдалось подплавление корпуса устройства локализации (рис. 7). Процесс плавления начался через 60 мин после аварии. Благодаря образованию корки затвердевшего кориума на относительно холодной поверхности корпуса, мгновенного нагрева не произошло. Нагрев корпуса происходил неравномерно, что привело к тому, что плавление изначально затронуло днище и верхнюю часть стенки.

По мере снижения тепловыделений, количество расплавляющегося металла уменьшалось. Через 49 мин после начала нагрева стенки максимальная температура корпуса не превысила температуры плавления ($T = 1688$ К), что привело к полной остановке процесса плавления.

Максимальный тепловой поток при нагреве стенки составлял 640 кВт/м^2, а после нагрева снизился до 400 кВт/м^2. В течение 30 ч тепловой поток упал до 80 кВт/м^2.

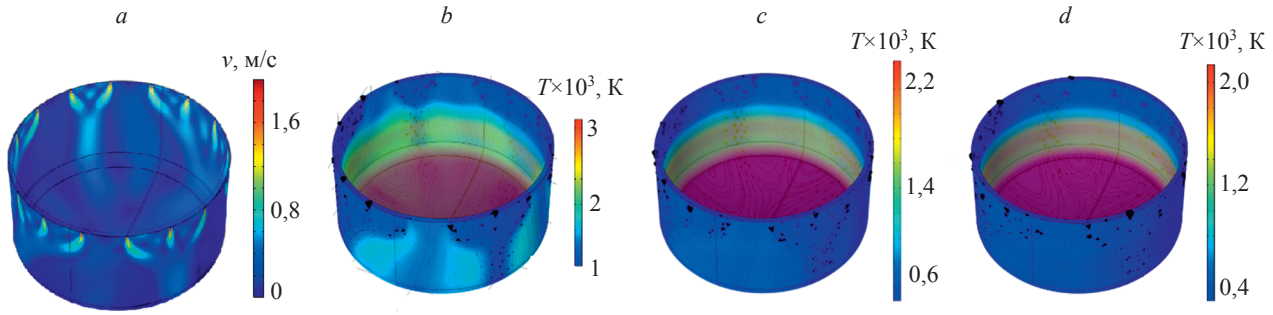

Рис. 6. Поле скоростей при массовом расходе воды 20 кг/с (*a*), температурные поля и линии тока при массовом расходе воды: 5 кг/с (*b*); 20 кг/с (*c*); 60 кг/с (*d*)

Fig. 6. Velocity field at a water mass flow rate of 20 kg/s (*a*); temperature fields and streamlines at water mass flow rates: 5 kg/s (*b*); 20 kg/s (*c*); 60 kg/s (*d*)

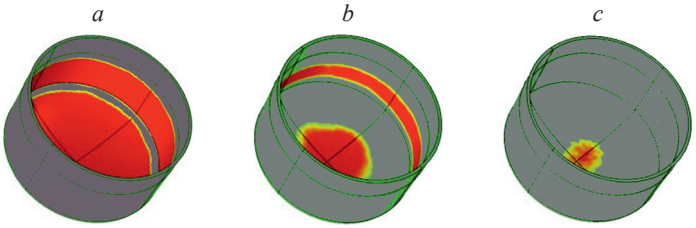

Рис. 7. Процесс проплавления корпуса устройства локализации: 60 мин (*a*); 1 ч 20 мин (*b*); 1 ч 40 мин (*c*). Процесс плавления (красная область)

Fig. 7. Melt-through of the core-catcher vessel: 60 min (*a*); 1 h 20 min (*b*); 1 h 40 min (*c*). Melting process (red area)

Вывод

Для моделирования тепловых процессов в устройстве локализации расплава разработана модель, основанная на обобщенной формулировке задачи Стефана. Модель включает уравнения теплопроводности и движений, учитывая тепловыделение кориума, реакции с жертвенным материалом, теплоперенос конвекцией и охлаждение водой. Использован пакет COMSOL Multiphysics для дискретизации, автоматизации условий и расчетов. Это снижает общую неопределенность моделирования и повышает детализацию термокинетических характеристик процессов при разработке устройств локализации расплава.

Результаты трехмерного моделирования, учитывающего пространственные особенности естественной конвекции компонентов расплава кориума, показали сокращение продолжительности до-инверсной и пост-инверсной стадий по сравнению с двумерной моделью. Анализ трехмерных расчетов тепловых потоков, температурных полей, термических и гидростатических напряжений корпуса ловушки выявил возможность частичного подплавления стенок с внутренней стороны в районе днища и обечайки корпуса устройства локализации на границе заполнения расплавом. Глубина подплавления составляет 1,5–2 мм и прекращается в течение первого часа после поступления расплава в ловушку, когда температура расплава снижается на этапе взаимодействия с жертвенным материалом до инверсии оксидной и металлической составляющих.

Более быстрый темп остывания металлического слоя и более легкоплавкие составляющие приводят к его опережающему затвердеванию по сравнению с тугоплавкими компонентами оксидного слоя. В оксидном слое в течение длительного времени сохраняется расплавленное ядро, и даже через 200 ч в закристаллизовавшихся зонах остаются отдельные очаги жидкой фазы, где возможно скопление оксидов повышенной радиоактивности.

Численный эксперимент продемонстрировал, что удержание кориума в материальной ловушке способно минимизировать последствия аварии с расплавлением активной зоны на АЭС. Использование предложенного решения по размещению жертвенного материала может продлить период борьбы с тяжелой аварией.

Обсуждение

Результаты численного моделирования показывают, что трехмерный анализ турбулентной конвекции и фазовых превращений значительно повышает точность предсказаний тепловых режимов системы. Такой подход позволяет выявить критические зоны подплавления, понять механизмы образования и динамики корки, что невозможно при двухмерных моделях.

Особенностью трехмерной модели является учет локальных эффектов турбулентных вихрей и асимметрий распределения тепла, особенно вблизи днища и верхней обечайки корпуса устройства локализации. Это важно для оценки надежности защитных конструкций и выявления слабых мест системы.

Анализ показал, что конфигурация из треугольных пирамид жертвенного материала обеспечивает равномерное распределение теплового потока, способствуя своевременному формированию стабильной корки и снижая риск превышения температуры корпуса. Модель также позволила определить оптимальные параметры охлаждения, включая расход воды и временные интервалы.

При использовании предложенной конструкции и расхода воды около 20 кг/с температуры стенок и тепловые потоки остаются в пределах нормативных требований. Однако в реальных условиях возможны значительные вариации параметров из-за изменений тепловыделения, тепловых потерь или форм и материалов конструкции. Потому необходимо расширять исследования диапазоном аварийных сценариев.

Важно также учитывать образование пористых структур и трещин в корке, что может привести к непредсказуемому тепловому режиму и разрушению защитных слоев. Необходимо развивать модели миграции радиоактивных веществ и их влияние на теплообмен, так как радиоактивные компоненты могут существенно менять тепловые свойства системы.

Перспективы включают развитие моделей с учетом ядерных реакций, миграции радиоактивных веществ, а также внедрение искусственного интеллекта для автоматической интерпретации данных и принятия решений. Кроме того, целесообразно расширять моделирование для различных типов реакторов и использовать мультифизические подходы, объединяющие теплофизику, гидродинамику, химию и радиационную безопасность.

Заключение

В работе создана и успешно реализована трехмерная математическая модель теплофизических процессов, протекающих в системе локализации расплава кориума на примере атомной электростанции. Модель с учетом фазовых превращений, турбулентных течений и взаимодействия с жертвенным материалом позволила получить детальное представление о динамике распределения температур, формировании и развитии корки, а также о временных характеристиках плавления и затвердевания различных компонентов расплава.

Полученные данные подтверждают, что предложенная конструкция системы и оптимальные параметры охлаждения обеспечивают стабильное удержание кориума без выхода за пределы критических значений тепловых нагрузок на корпус. Использование трехмерного моделирования выявило ключевые зоны риска, связанные с локальным подплавлением и образованием трещин, что важно учитывать при создании резервных и защитных систем.

Полученные результаты подтвердили возможность прогнозирования поведения системы в условиях аварийных сценариев, что существенно повышает уровень безопасности при эксплуатации атомной электростанции и снижает вероятность экологических последствий.

Литература

1. Xi Z., Hu Y., Gong H., Zhang L., Ma Z., Sun W., et al. Numerical study on the corium pool heat transfer with OpenFOAM // Frontiers in Energy Research. 2020. V. 8. P. 80. https://doi.org/10.3389/fenrg.2020.00080
2. Кухтевич И.В., Безлепкин В.В., Грановский В.С. [и др.]. Концепция локализации расплава кориума на внекорпусной стадии запроектной аварии АЭС с ВВЭР-1000 // Теплоэнергетика. 2001. № 9. С. 2–7.
3. Гусаров В.В., Альмяшев В.И., Хабенский В.Б., Бешта С.В., Грановский В.С. Взаимодействие материала на основе оксидов алюминия и железа с расплавом металлов // Журнал прикладной химии. 2007. Т. 80. № 4. С. 541–548.
4. Asmolov V.G., Zagryazkin V.N., Tsurikov D.F. The thermodynamics of U-Zr-Fe-O melts // High Temperature. 2007. V. 45. N 3. P. 305–312. https://doi.org/10.1134/s0018151x07030042
5. Viot L., Le Tellier R., Peybernes M. Modeling of the corium crust of a stratified corium pool during severe accidents in light water reactors // Nuclear Engineering and Design. 2020. V. 368. P. 110816. https://doi.org/10.1016/j.nucengdes.2020.110816
6. Gaus-Liu X., Miassoedov A., Gabriel S. Review of experimental studies on the heat transfer behaviour of volumetrically-heated pool with different boundary conditions and the influence of crust formation // Proc. of the 24th International Conference on Nuclear Engineering. 2016. V. 2. P. UNSP V002T07A007. https://doi.org/10.1115/icone24-60268
7. Zhang L., Luo S., Zhang Y., Tian W., Su G., Qiu S. Large eddy simulation on turbulent heat transfer in reactor vessel lower head corium pools // Annals of Nuclear Energy. 2018. V. 111. P. 293–302. https://doi.org/10.1016/j.anucene.2017.08.055
8. Динь Ч.Н., Нигматулин Б.И., Нургалиев Р.Р., Рассохин Н.Г. Оценка тепловой нагрузки на стадии взаимодействия расплавленного кориума с корпусом ВВЭР при тяжелой аварии // Теплоэнергетика. 1996. № 3. С. 9–17.
9. Fukasawa M., Hayakawa S., Saito M. Thermal-hydraulic analysis for inversely stratified molten corium in lower vessel // Journal of Nuclear Science and Technology. 2008. V. 45. N 9. P. 873–888. https://doi.org/10.1080/18811248.2008.9711489
10. Tran C.T., Kudinov P. The effective convectivity model for simulation of molten metal layer heat transfer in a boiling water reactor lower head // Science and Technology of Nuclear Installations. 2013. V. 2013. P. 231501. https://doi.org/10.1155/2013/231501
11. Коваленко А.Н., Коптюхов А.О., Мещеряков Д.К., Щуклинов А.П. Моделирование термокинетических процессов в тигельных ловушках расплава кориума при авариях реакторов на АЭС // Научно-технические ведомости Санкт-Петербургского государственного политехнического университета. Физико-математические науки. 2020. Т. 13. № 4. С. 61–76. https://doi.org/10.18721/JPM.13405
12. Коваленко А.Н., Коптюхов А.О., Мещеряков Д.К., Щуклинов А.П. Расширенное трехстадийное 3D-моделирование термокинетических процессов в тигельных ловушках расплава кориума при авариях реакторов на АЭС // Научно-технические ведомости Санкт-Петербургского государственного политехнического уни-

References

1. Xi Z., Hu Y., Gong H., Zhang L., Ma Z., Sun W., et al. Numerical study on the corium pool heat transfer with OpenFOAM. *Frontiers in Energy Research*, 2020, vol. 8, pp. 80. https://doi.org/10.3389/fenrg.2020.00080
2. Kukhtevich I.V., Bezlepkin V.V., Granovskii V.S., et al. The concept of localization of the corium melt in the ex-vessel stage of a severe accident at a nuclear power station with a VVER-1000 reactor. *Thermal Engineering*, 2001, vol. 48, no. 9, pp. 699–706.
3. Gusarov V.V., Al'myashev V.I., Khabenskii V.B., Beshta S.V., Granovskii V.S. Interaction of a material based on aluminum and iron oxides with a metal melt. *Russian Journal of Applied Chemistry*, 2007, vol. 80, no. 4, pp. 528–535. https://doi.org/10.1134/S1070427207040027
4. Asmolov V.G., Zagryazkin V.N., Tsurikov D.F. The thermodynamics of U-Zr-Fe-O melts. *High Temperature*, 2007, vol. 45, no. 3, pp. 305–312. https://doi.org/10.1134/s0018151x07030042
5. Viot L., Le Tellier R., Peybernes M. Modeling of the corium crust of a stratified corium pool during severe accidents in light water reactors. *Nuclear Engineering and Design*, 2020, vol. 368, pp. 110816. https://doi.org/10.1016/j.nucengdes.2020.110816
6. Gaus-Liu X., Miassoedov A., Gabriel S. Review of experimental studies on the heat transfer behaviour of volumetrically-heated pool with different boundary conditions and the influence of crust formation. *Proc. of the 24th International Conference on Nuclear Engineering*, 2016, vol. 2, pp. V002T07A007. https://doi.org/10.1115/icone24-60268
7. Zhang L., Luo S., Zhang Y., Tian W., Su G., Qiu S. Large eddy simulation on turbulent heat transfer in reactor vessel lower head corium pools. *Annals of Nuclear Energy*, 2018, vol. 111, pp. 293–302. https://doi.org/10.1016/j.anucene.2017.08.055
8. Dinh T.N., Nigmatulin B.I., Nurgaliev R.R., Rassokhin N.G. Assessment of the thermal load when molten debris interact with a VVER reactor vessel during a severe accident. *Thermal Engineering*, 1996, vol. 43, no. 3, pp. 189–197.
9. Fukasawa M., Hayakawa S., Saito M. Thermal-hydraulic analysis for inversely stratified molten corium in lower vessel. *Journal of Nuclear Science and Technology*, 2008, vol. 45, no. 9, pp. 873–888. https://doi.org/10.1080/18811248.2008.9711489
10. Tran C.T., Kudinov P. The effective convectivity model for simulation of molten metal layer heat transfer in a boiling water reactor lower head. *Science and Technology of Nuclear Installations*, 2013, vol. 2013, pp. 231501. https://doi.org/10.1155/2013/231501
11. Kovalenko A.N., Koptyukhov A.O., Meshcheryakov D.K., Schuklinov A.P. Thermokinetic processes in the corium traps of high-temperature melt during the reactor accidents at an a-plant: simulation. *St. Petersburg Polytechnical State University Journal. Physics and Mathematics*, 2020, vol. 13, no. 4, pp. 61–76. (in Russian). https://doi.org/10.18721/JPM.13405
12. Kovalenko A. N., Koptyukhov A. O., Meshcheryakov D. K., Schuklinov A. P., Thermokinetic processes in the corium traps of high-temperature melt during the reactor accidents at an A-plant: advanced three-stage 3D simulation. *St. Petersburg Polytechnical State University Journal. Physics and Mathematics*, 2021,

верситета. Физико-математические науки. 2021. Т. 14. № 4. С. 69–83. https://doi.org/10.18721/JPM.14405
13. Boettinger W.J., Warren J.A., Beckermann C., Karma A. Phase-field simulation of solidification // Annual Review of Materials Research. 2002. V. 32. P. 163–194. https://doi.org/10.1146/annurev.matsci.32.101901.155803
14. Wang Q., Yu Q., Du W., Fang Z., Li K., Wang Q. Fluid distribution in a two-phase space accumulator predicted by a coupled multi-scale model based on single-domain approach // International Communications in Heat and Mass Transfer. 2025. V. 162. P. 108567. https://doi.org/10.1016/j.icheatmasstransfer.2024.108567
15. Кириллов П.Л., Богословская Г.П. Тепломассообмен в ядерных энергетических установках. М.: Энергоатомиздат, 2000. 456 с.

vol. 14, no. 3, pp. 69–83. (in Russian). https://doi.org/10.18721/JPM.14405
13. Boettinger W.J., Warren J.A., Beckermann C., Karma A. Phase-field simulation of solidification. *Annual Review of Materials Research*, 2002, vol. 32, pp. 163–194. https://doi.org/10.1146/annurev.matsci.32.101901.155803
14. Wang Q., Yu Q., Du W., Fang Z., Li K., Wang Q. Fluid distribution in a two-phase space accumulator predicted by a coupled multi-scale model based on single-domain approach. *International Communications in Heat and Mass Transfer*, 2025, vol. 162, pp. 108567. https://doi.org/10.1016/j.icheatmasstransfer.2024.108567
15. Kirillov P.L., Bogoslovskaya G.P. *Heat and Mass Transfer in Nuclear Power Plants*. Moscow, Energoatomizdat Publ., 2000, 456 p. (in Russian)

Авторы

Пилипенко Николай Васильевич — доктор технических наук, профессор, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 7006938207, https://orcid.org/0000-0001-9328-3166, pilipenko38@mail.ru

Коптюхов Артём Олегович — кандидат технических наук, научный сотрудник, НИЦ «Курчатовский институт» — ПИЯФ, Гатчина, 188300, Российская Федерация, sc 57213518403, https://orcid.org/0000-0002-9660-3601, koptyukhov_ao@pnpi.nrcki.ru

Authors

Nikolay V. Pilipenko — D.Sc., Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 7006938207, https://orcid.org/0000-0001-9328-3166, pilipenko38@mail.ru

Artem O. Koptyukhov — PhD, Scientific Researcher, NRC “Kurchatov Institute” — PNPI, Gatchina, 188300, Russian Federation, sc 57213518403, https://orcid.org/0000-0002-9660-3601, koptyukhov_ao@pnpi.nrcki.ru

Статья поступила в редакцию 07.07.2025
Одобрена после рецензирования 15.10.2025
Принята к печати 17.11.2025

Received 07.07.2025
Approved after reviewing 15.10.2025
Accepted 17.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

КРАТКИЕ СООБЩЕНИЯ

BRIEF PAPERS

doi: 10.17586/2226-1494-2025-25-6-1229-1233
УДК 004.2

Проектные механизмы микроархитектурного уровня для встраиваемых систем

Максим Вячеславович Кольчурин[1]✉, Алексей Евгеньевич Платунов[2]

[1] ООО «ЛМТ», Санкт-Петербург, 199034, Российская Федерация
[2] Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[1] maxim.kolchurin@gmail.com✉, https://orcid.org/0000-0002-7061-9357
[2] aeplatunov@itmo.ru, https://orcid.org/0000-0003-3003-3949

Аннотация
Стремительно растущая потребность в проектировании специализированных вычислительных систем (особенно встраиваемых систем и систем на кристалле) сдерживается ограниченными возможностями разработчиков из-за высокой сложности задачи. Предлагается «проектный механизм» как абстракция (абстрактное понятие) с явным разделением на архитектурную логику и уровень реализации по комплексному критерию. Его применение наиболее эффективно при проектировании подсистем с высокой внутренней вариативностью и множеством технологических альтернатив. Разработанная абстракция на всех этапах маршрута проектирования позволяет эффективнее формировать и анализировать пространство проектных решений, радикально сокращая концептуальные ошибки и расширяя для повторного использования число результатов проектной деятельности.

Ключевые слова
встраиваемые системы, архитектурные решения, метрики

Ссылка для цитирования: Кольчурин М.В., Платунов А.Е. Проектные механизмы микроархитектурного уровня для встраиваемых систем // Научно-технический вестник информационных технологий, механики и оптики. 2025. Т. 25, № 6. С. 1229–1233. doi: 10.17586/2226-1494-2025-25-6-1229-1233

Design mechanisms at the microarchitectural level for embedded systems

Maxim V. Kolchurin[1]✉, Alexey E. Platunov[2]

[1] LMT Ltd., Saint Petersburg, 199034, Russian Federation
[2] ITMO University, Saint Petersburg, 197101, Russian Federation

[1] maxim.kolchurin@gmail.com✉, https://orcid.org/0000-0002-7061-9357
[2] aeplatunov@itmo.ru, https://orcid.org/0000-0003-3003-3949

Abstract
The rapidly growing demand for the design of specialized computing systems, primarily embedded systems and systems-on-chip, constrained by the limited capabilities of developers, due to the high complexity of the task. A Design Mechanism is proposed as an abstraction (a conceptual construct) that explicitly separates architectural logic from the implementation level according to a comprehensive criterion. Its use is most effective in the design of subsystems characterized by high internal variability and numerous technological alternatives. Such an abstraction, applied at all stages of the design process, enables more efficient formation and analysis of the design solution space, radically reducing conceptual errors and expanding the number of design outcomes available for reuse.

Keywords
embedded systems, design patterns, architectural solutions, metrics

For citation: Kolchurin M.V., Platunov A.E. Design mechanisms at the microarchitectural level for embedded systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2025, vol. 25, no. 6, pp. 1229–1233 (in Russian). doi: 10.17586/2226-1494-2025-25-6-1229-1233

© Кольчурин М.В., Платунов А.Е., 2025

В области проектирования встраиваемых систем проблема заполнения семантического разрыва между представлениями концептуального (архитектурного) уровня и уровня реализации по-прежнему стоит очень остро [1]. Она заключается в существовании «серой зоны» между архитектурным этапом проектирования и этапом реализации, недостаточно обеспеченной методологически и инструментально. Сокращение семантического разрыва возможно за счет выделения в архитектурном проектировании микроархитектурного этапа [2], который включает в себя определение внутренней организации компонентов, выбор алгоритмов, адаптацию решений под платформу, обоснование компромиссов. Другими словами, основная задача микроархитектурного этапа проектирования — формирование, визуализация и документирование пространства детализированных проектных решений [3].

В настоящей работе рассматривается проблема представления инженерных решений на микроархитектурном уровне для встраиваемых систем и предлагается подход, направленный на повышение степени формализации и переносимости таких решений между технологическими контекстами. Ключевое отличие состоит в принудительном разделении каждого инженерного решения на архитектурную логику и множество вариантов реализации, адаптированных под конкретные платформы и ограничения, а также в способе такого разделения, основанном на явной фиксации инварианта решения и критериев корректности.

Современные методологии проектирования систем на кристалле, встраиваемых и киберфизических систем развивались в ответ на растущее многообразие, внутреннюю сложность и многовариантность технических решений. Существующие методологии, такие как модель-ориентированные подходы (MBSE, SysML/UML) [4, 5], формальные методы (TLA+, UPPAAL) [6, 7], отраслевые фреймворки (DDD, TOGAF) [8] — частично преодолевают семантический разрыв между замыслом и реализацией, но сосредоточены на отдельных аспектах проектирования, таких как согласованность представлений, верификация свойств, структурирование процессов.

Работы, связанные с развитием платформно-ориентированного проектирования (Platform Based Design, PBD) [9, 10], заложили важную концептуальную основу — ортогонализацию (разделение) между функцией (что система делает) и архитектурой (как система это делает). Вместе с тем PBD сосредоточено на выборе платформы и отображении функциональности на ее компоненты через симуляцию и оценку производительности, документируя результаты этого процесса, но не рассматривает (как самостоятельную) задачу извлечения и формализации архитектурного знания, направляющего эти решения, в форме, пригодной для повторного использования. Практики записи архитектурных решений (Architecture Decision Records, ADR) [11] документируют принятые решения, но фиксируют их постфактум, не раскрывая компромиссы между альтернативами.

Таким образом, существующие подходы либо описывают процесс принятия решений (PBD), либо фиксируют их результат (ADR), но не предлагают артефакта, который обеспечивал бы явную связь между сутью принятого проектного решения и свойствами его технологических воплощений.

Для заполнения этого пробела предлагается концепция, основанная на абстракции (абстрактном понятии) — «проектный механизм» (ПМ) — подход к формализации инженерных решений, направленный на структурирование архитектурного знания, который обеспечивает явное разделение логики проектного решения и конкретных технических воплощений. ПМ включает в себя:

— архитектурную логику — технологически нейтральное описание решения на микроархитектурном уровне [1, 2], фиксирующее обоснование выбора, ключевые компромиссы, метрики для сравнения альтернатив и границы применимости;
— реализации — конкретные воплощения архитектурной логики с учетом особенностей и ограничений конкретных технологических платформ.

Практическая применимость ПМ опирается на четыре критерия корректности:

1. Технологическая нейтральность (инвариантность к платформе). Архитектурная логика сохраняет семантику в разных технологических контекстах и не содержит деталей конкретных реализаций.
2. Генеративная способность. Из архитектурной логики выводится множество реализаций для различных условий, при этом логика не предписывает конкретную технологию или инструмент.
3. Измеримость свойств. Архитектурная логика содержит (или указывает на) метрики и критерии, обеспечивающие количественное сравнение альтернатив и валидацию соответствия решения заявленным свойствам.
4. Трассируемость реализации. Каждая реализация явно связана с компонентами архитектурной логики через объяснимые проектные выборы, что предотвращает искусственные сопоставления и обеспечивает воспроизводимость.

Критерии корректности трансформируют процесс выбора из качественного сравнения технологий в количественный анализ соответствия характеристик реализаций ограничениям проекта, обеспечивая переносимость архитектурного знания между проектами. ПМ наиболее применим при проектировании подсистем, где существуют: множество технологических альтернатив; нефункциональные характеристики определяются разнородными критериями; компромиссы между вариантами не очевидны, и цена ошибочного выбора высока.

Продемонстрируем применение ПМ на задаче организации подсистемы конфигурирования для промышленного контроллера со встроенной средой исполнения пользовательских программ. Такие программы требуют параметризации (сетевые настройки, пороги срабатывания, режимы работы и т. д.) без модификации внутренней программы (прошивки) контроллера. Рассматриваемая система разработана на базе микроконтроллера без блока управления памятью и механизма страничной памяти и, как следствие, без

механизмов защиты от фрагментации. Начальная реализация подсистемы конфигурирования использовала хранилище типа ключ-значение[1], которое обеспечивало устойчивость к сбоям питания и равномерный износ постоянного хранилища данных. Реализация такого механизма требовала применения динамической памяти, которая подвержена фрагментации (15–20 % по измерениям), а также создавала непредсказуемые паузы из-за работы сборщика мусора (Garbage Collector, GC), которые достигали нескольких миллисекунд при полном заполнении активного хранилища. В сценариях с высокой нагрузкой на чтение (одна запись на 5000 чтений) механизм оказался избыточным. Адаптация через специальные режимы работы, изоляция области динамической памяти, хотя и улучшали ситуацию, но не устраняли фундаментальные ограничения текущей реализации.

Поиск альтернатив сталкивался с проблемами. Текстовые форматы (Tom's Obvious Minimal Language (TOML), JavaScript Object Notation (JSON)) требуют синтаксического анализа при исполнении, нагружая динамическую память. Генерация структур данных из схем данных (Concise Binary Object Representation (CBOR) с использованием Concise Data Definition Language (CDDL)) ограничивает гибкость эволюции и требует постоянных выделений памяти при исполнении. Анализ неудач выявил ключевое противоречие — формат, удобный для написания человеком, оптимален для одной фазы (генерация конфигурации), но не для других (валидация и исполнение). Эти требования не конкурируют во времени, а относятся к разным стадиям процесса конфигурирования устройства. Если требования не пересекаются, они могут быть удовлетворены через разные артефакты. Каждая стадия получает оптимальный для нее формат вместо единого компромиссного решения. Конфигурация существует на трех стадиях (генерация, транспорт и исполнение), и для каждой из них используется представление, оптимизированное под ее задачи. Неизменным должна оставаться семантическая составляющая представления и учитываться ограничения по времени доступа. Это было зафиксировано в виде архитектурной логики для поиска вариантов реализаций такой подсистемы. В результате рассмотрено несколько вариантов, основанных на выявленной архитектурной логике. На этапе подготовки применены форматы, удобные для чтения человеком (JSON и TOML), различия между ними несущественны для настоящей работы, поэтому дальнейший анализ сосредоточен на стадии исполнения. Требования проекта определяют метрики для сравнения реализаций, так, потребление оперативной памяти (ОЗУ) не должно превышать 32 КБ, (выделенный объем памяти для подсистемы), время доступа стремится к $O(1)$ (асимптотическая сложность), необходима статическая валидация (на этапе компиляции) и поддержка прямого использования данных (таблица).

Выбор осуществлялся последовательным исключением вариантов по критическим ограничениям, TOML и CBOR исключены из-за постоянного высокого потребления ОЗУ на грани лимита, Key-Value хранилище — из-за фрагментации, пауз GC и отсутствия проверки типов данных. Структуры Си исключены из-за необходимости ручной кодогенерации, снижающей гибкость, сложности переноса на другие языки. FlatBuffers является наиболее подходящим под критерии, он обеспечивает прямое использование данных (zero-copy) непосредственно из флэш-памяти в режиме eXecute In Place, константные задержки доступа к данным ($O(1)$) и compile-time валидацию схемы данных. В результате применения FlatBuffers отсутствует этап интерпретации, прямой доступ без поиска, отсутствие пауз GC.

Применимость ПМ не ограничивается подсистемами конфигурирования. Например, авторы работы имеют опыт применения ПМ при проектировании коммуникационных подсистем промышленной автоматики. При выборе протоколов для сетей реального времени ПМ помог структурировать сравнение альтернатив (802.15.4e TSCH, CAN, EtherCAT, Modbus) по единым критериям (гарантии задержки, джиттер, число узлов) вместо качественного сравнения разнородных спецификаций. В распределенных системах сбора данных ПМ фиксирует компромисс между предсказуемостью доставки и энергоэффективностью, сравнивая реализации протоколов (MQTT, AMQP, DDS, LWM2M) по приоритетным для проекта метрикам.

ПМ отличается от абстракций, используемых в методологиях PBD или ADR. Он предназначен для реше-

[1] [Электронный ресурс]. Режим доступа: https://github.com/Infineon/kv-store (дата обращения: 22.09.2025).

Таблица. Сравнение вариантов реализаций подсистемы конфигурирования

Table. Comparison of configuration subsystem implementation options

Реализация	Объем ОЗУ, КБ	Сложность доступа к данным	Проверка типов данных	Прямой доступ к данным (zero-copy)	Соответствие ограничениям
Интерпретация TOML	до 80	$O(n)$	Во время исполнения	Нет	Нет
CBOR с CDDL	до 32	$O(1)$	Во время исполнения	Нет	Нет
Key-Value хранилище	около 20*	$O(n)$	Нет	Нет	Нет
Структуры языка Си	10	$O(1)$	На этапе компиляции	Да	Возможно**
FlatBuffers	0	$O(1)$	На этапе компиляции	Да	Да

Примечание: * — с фрагментацией, по замерам 15–20 % и паузами на сборку мусора; ** — требует ручной кодогенерации, снижает гибкость.

ния проблем проектирования на уровне микроархитектуры. Среди решаемых проблем: явное определение общих проектных принципов, устойчивых к различным реализациям; минимизация привязки к конкретной платформе; фокусировка на архитектурных решениях и их свойствах без погружения в детали конкретной технической реализации.

ПМ наиболее эффективен при проектировании подсистем с высокой внутренней вариативностью и множеством технологических альтернатив. Такой подход избыточен для проектов с однозначным выбором технологии (диктуется стандартом или платформой) или когда функциональные требования полностью определяют реализацию. В таких случаях достаточно традиционного ADR для фиксации решения.

Ограничения подхода связаны с необходимостью явной формулировки метрик, но не все проектные компромиссы поддаются количественной оценке (например, «удобство интеграции» или «зрелость экосистемы»). В этих случаях ПМ дополняется экспертными оценками. Кроме того, извлечение архитектурной логики требует аналитических усилий, что может быть неоправданно для одноразовых решений без перспективы переиспользования.

Абстракции проектирования, осознанно применяемые разработчиками встраиваемых систем на всех этапах проектирования, играют определяющую роль в повышении производительности труда. Пространство проектных решений, с которым работает инженер в области компьютинга, объективно трудно формализуемо, особенно на этапах архитектурного (микроархитектурного) проектирования. Следует отметить, что процесс освоения технологий проектирования на основе такого рода абстракций, как представленный ПМ, в качестве альтернативы «привычному» манипулированию техническими решениями (конструированию) уровня реализации требует значительных усилий и желания, и должен активнее демонстрироваться в образовательных учреждениях.

Литература

1. Кольчурин М.В., Пинкевич В.Ю., Платунов А.Е. Усиление роли микроархитектурных этапов проектирования встраиваемых систем // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22. № 4. С. 716–724. https://doi.org/10.17586/2226-1494-2022-22-4-716-724
2. Kolchurin M., Platunov A., Asminkin F., Pinkevich V. Architectural abstractions in the design of distributed embedded systems platforms // Proc. of the 24th International Multidisciplinary Scientific GeoConference SGEM. 2024. V. 24. N 2.1. P. 19–26. https://doi.org/10.5593/sgem2024/2.1/s07.03
3. Pinkevich V., Platunov A., Gorbachev Ya. Design of embedded and cyber-physical systems using a cross-level microarchitectural pattern of the computational process organization // CEUR Workshop Proceedings. 2020. V. 2893. P. 1–14.
4. Technical Report: SysML v1 to SysML v2 Model Conversion Approach. Office of Systems Engineering and Architecture. Washington, D.C., 2024. 49 p.
5. Alenazi M., Niu N., Savolainen J. SysML modeling mistakes and their impacts on requirements // Proc. of the IEEE 27th International Requirements Engineering Conference Workshops (REW). 2019. P. 14–23. https://doi.org/10.1109/REW.2019.00010
6. Neider D., Roy R. What is formal verification without specifications? A survey on mining LTL specifications // Lecture Notes in Computer Science. 2025. V. 15262. P. 109–125. https://doi.org/10.1007/978-3-031-75778-5_6
7. Formal Methods Specification and Verification Guidebook for Software and Computer Systems. Volume I: Planning and Technology Insertion. Technical report NASA-GB-002-95. Washington, DC: NASA, 1995. 227 p.
8. Kotusev S. TOGAF-based enterprise architecture practice: An exploratory case study // Communications of the Association for Information Systems. 2018. V. 43. N 1. P. 321–359. https://doi.org/10.17705/1CAIS.04320
9. Sangiovanni-Vincentelli A., Carloni L., De Bernardinis F., Sgroi M. Benefits and challenges for platform-based design // Proc. of the 41st annual Design Automation Conference. 2004. P. 409–414. https://doi.org/10.1145/996566.996684
10. Keutzer K., Malik S., Newton A.R., Rabaey J.M., Sangiovanni-Vincentelli A. System-level design: orthogonalization of concerns and platform-based design // IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2000. V. 19. N 12. P. 1523–1543. https://doi.org/10.1109/43.898830
11. Buchgeher G., Schöberl S., Geist V., Dorninger B., Haindl P., Weinreich R. Using architecture decision records in open source projects — an MSR study on GitHub // IEEE Access. 2023. V. 11. P. 63725–63740. https://doi.org/10.1109/ACCESS.2023.3287654

References

1. Kolchurin M.V., Pinkevich V.Yu., Platunov A.E. Strengthening the role of microarchitectural stages of mbedded systems design. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 4, pp. 716–724. (in Russian). https://doi.org/10.17586/2226-1494-2022-22-4-716-724
2. Kolchurin M., Platunov A., Asminkin F., Pinkevich V. Architectural abstractions in the design of distributed embedded systems platforms. *Proc. of the 24th International Multidisciplinary Scientific GeoConference SGEM*, 2024, vol. 24, no. 2.1, pp. 19–26. https://doi.org/10.5593/sgem2024/2.1/s07.03
3. Pinkevich V., Platunov A., Gorbachev Ya. Design of embedded and cyber-physical systems using a cross-level microarchitectural pattern of the computational process organization. *CEUR Workshop Proceedings*, 2020, vol. 2893, pp. 1–14
4. *Technical Report: SysML v1 to SysML v2 Model Conversion Approach*. Office of Systems Engineering and Architecture. Washington, D.C., 2024, 49 p.
5. Alenazi M., Niu N., Savolainen J. SysML modeling mistakes and their impacts on requirements. *Proc. of the IEEE 27th International Requirements Engineering Conference Workshops (REW)*, 2019, pp. 14–23. https://doi.org/10.1109/REW.2019.00010
6. Neider D., Roy R. What is formal verification without specifications? A survey on mining LTL specifications. *Lecture Notes in Computer Science*, 2025, vol. 15262, pp. 109–125. https://doi.org/10.1007/978-3-031-75778-5_6
7. *Formal Methods Specification and Verification Guidebook for Software and Computer Systems. Volume I: Planning and Technology Insertion*. Technical report NASA-GB-002-95. Washington, DC: NASA, 1995. 227 p.
8. Kotusev S. TOGAF-based enterprise architecture practice: An exploratory case study. *Communications of the Association for Information Systems*, 2018, vol. 43, no. 1, pp. 321–359. https://doi.org/10.17705/1CAIS.04320
9. Sangiovanni-Vincentelli A., Carloni L., De Bernardinis F., Sgroi M. Benefits and challenges for platform-based design. *Proc. of the 41st annual Design Automation Conference*, 2004, pp. 409–414. https://doi.org/10.1145/996566.996684
10. Keutzer K., Malik S., Newton A.R., Rabaey J.M., Sangiovanni-Vincentelli A. System-level design: orthogonalization of concerns and platform-based design. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2000, vol. 19, no. 12, pp. 1523–1543. https://doi.org/10.1109/43.898830
11. Buchgeher G., Schöberl S., Geist V., Dorninger B., Haindl P., Weinreich R. Using architecture decision records in open source projects — an MSR study on GitHub. *IEEE Access*, 2023, vol. 11, pp. 63725–63740. https://doi.org/10.1109/ACCESS.2023.3287654

Авторы

Кольчурин Максим Вячеславович — инженер, ООО «ЛМТ», Санкт-Петербург, 199034, Российская Федерация, sc 57225126062, https://orcid.org/0000-0002-7061-9357, maxim.kolchurin@gmail.com
Платунов Алексей Евгеньевич — доктор технических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, sc 35318291200, https://orcid.org/0000-0003-3003-3949, aeplatunov@itmo.ru

Authors

Maxim V. Kolchurin — Engineer, LMT Ltd., Saint Petersburg, 199034, Russian Federation, sc 57225126062, https://orcid.org/0000-0002-7061-9357, maxim.kolchurin@gmail.com
Alexey E. Platunov — D.Sc., Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, sc 35318291200, https://orcid.org/0000-0003-3003-3949, aeplatunov@itmo.ru

Статья поступила в редакцию 30.09.2025
Одобрена после рецензирования 16.10.2025
Принята к печати 17.11.2025

Received 30.09.2025
Approved after reviewing 16.10.2025
Accepted 17.11.2025

Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

Уважаемые подписчики научно-технической литературы!

Журнал выходит 6 раз в год.

На журнал «Научно-технический вестник информационных технологий, механики и оптики» можно оформить подписку в почтовых отделениях по объединенному каталогу «Пресса России», подписные индексы 47197 (полугодовая подписка),

а также online
по объединенному каталогу «Пресса России» и по каталогу «Пресса по подписке», подписные индексы Э47197 (полугодовая подписка)
и по электронному каталогу Почты России, подписной индекс ПС543.

Корпоративная подписка и подписка физических лиц возможна
по каталогу компаний «Урал-Пресс»,
подписные индексы 47197 (полугодовая подписка) и 70522 (годовая подписка).

Сведения о подписке можно уточнить в редакции журнала по адресу:

Санкт-Петербург, ул. Ломоносова, д.9, литера А, комн. 2136.

Тел.: +7(812) 480 02 75