

# ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

2023, ТОМ 23, НОМЕР 4 (ИЮЛЬ–АВГУСТ)

ISSN 2226-1494 (PRINT), 2500-0373 (ONLINE)

## ОПТИЧЕСКИЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Определение типа действия ингибиторов гидратообразования по их инфракрасным спектрам *Ворожцова Ю.С., Носенко Т.Н., Успенская М.В.* 669

Спектроскопия комбинационного рассеяния света в исследованиях процессов инактивации бактериальных микроорганизмов *Матвеева К.И., Кундалевич А.А., Капитунова А.И., Зозуля А.С., Сухих С.А., Цибульникова А.В., Зюбин А.Ю., Самусев И.Г.* 676

Численное исследование влияния концентрации метгемоглобина в крови на поглощение света в коже человека *Беликов А.В., Чучин В.Ю.* 685

Низкотемпературная ячейка для инфракрасных фурье-спектрометрических исследований углеводородных веществ (на англ. яз.) *Кенбай А.А., Голиков О.Ю., Алдияров А.У., Ережен Д.Е.* 696

## НОВЫЕ МАТЕРИАЛЫ И НАНОТЕХНОЛОГИИ

Особенности выращивания твердых растворов  $Ga_{1-x}In_xAs$  на подложках GaAs в поле температурного градиента через тонкую газовую зону *Девыцкий О.В., Лунин Л.С., Митрофанов Д.В., Сысоев И.А., Никулин Д.А., Чапура О.М.* 703

## КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Усовершенствованный протокол безопасности на основе AES-GCM для защиты связи в интернете вещей (на англ. яз.) *Фероз Хан А.Б., Калтана Деви С., Рама Деви К.* 711

Атаки на основе вредоносных возмущений на системы обработки изображений и методы защиты от них *Есипов Д.А., Бучаев А.Я., Керимбай А., Пузикова Я.В., Сайдумаров С.К., Сулименко Н.С., Попов И.Ю., Кармановский Н.С.* 720

Сверхвысокое разрешение изображения магнитно-резонансной томографии головного мозга с использованием дискретного косинусного преобразования и сверточной нейронной сети (на англ. яз.) *Сингх П., Ганотра Д.* 734

Метод аугментации текстовых данных с сохранением стиля речи и лексики персоны *Матвеева А.А., Махныткина О.В.* 743

Верификация событийно-управляемых программных систем с использованием языка спецификации взаимодействующих автоматных объектов (на англ. яз.) *Афанасьева И.В., Новиков Ф.А., Федорченко Л.Н., Тагирова Л.Ф., Зубкова Т.М.* 750

Интеллектуальная система адаптивного тестирования 757

## ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И КОГНИТИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Нейросетевой метод визуального распознавания голосовых команд водителя с использованием механизма внимания *Аксёнов А.А., Рюмина Е.В., Рюмин Д.А., Иванько Д.В., Карпов А.А.* 767

Сегментация опухоли головного мозга на магнитно-резонансной томографии с использованием нечеткого деформируемого слияния и алгоритма Dolphin-SCA (на англ. яз.) *Типле А.Х., Какаде А.Б.* 776

Оптимизация систем отслеживания человека в виртуальной реальности на основе нейросетевого подхода *Обухов А.Д., Теселкин Д.В.* 786

## МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

Погрешности алгоритма демодуляции с генерируемой фазой несущей, вносимые фильтром низкой частоты *Мирошниченко Г.П., Аржаненкова А.Н., Плотников М.Ю.* 795

Моделирование процесса корректировки формы роторов шаровых гироскопов *Тит М.А., Беляев С.Н., Щербак А.Г., Юльметова О.С.* 803

Метод пространственного мультиплексирования в многоантенных системах связи *Гришенцев А.Ю.* 812

Математическое моделирование теплообменного аппарата с учетом сильной зависимости вязкости нефти от температуры *Курманова Д.Е., Джайчибеков Н.Ж.* 820

Подход к формированию обобщенных параметров технического состояния сложных технических систем с использованием нейросетевых структур *Мальцев Г.Н., Якимов В.Л.* 828

Численное моделирование газодинамики при работе широкодиапазонного ракетного сопла с пористой вставкой *Михайлов М.В., Продан Н.В., Ренев М.Е.* 836

Точное решение задачи отражения ударной волны от стенки, экранированной слоем газозвеси *Садин Д.В., Широкова Е.Н.* 843

## КРАТКИЕ СООБЩЕНИЯ

Адаптивный наблюдатель переменных состояния нелинейной нестационарной системы с неизвестными постоянными параметрами и запаздыванием в канале измерений *Бобцов А.А., Николаев Н.А., Козачёк О.А., Оськина О.В.* 850

RuLegalNER: новый датасет для распознавания именованных юридических сущностей на русском языке (на англ. яз.) *Шахин Э., Муромцев Д.И., Постный И.* 854

Главный редактор — В.О. Никифоров, д.т.н., профессор

Editor-in-Chief — Vladimir O. Nikiforov, D.Sc., Professor

#### РЕДАКЦИОННАЯ КОЛЛЕГИЯ

**Беззатеев С.В.** — д.ф.-м.н., доцент (Россия)  
**Беликов А.В.** — д.ф.-м.н., профессор (Россия)  
**Бессмертный И.А.** — д.т.н., профессор (Россия)  
**Бобцов А.А.** — д.т.н., профессор (Россия)  
**Бухановский А.В.** — д.т.н. (Россия)  
**Волков К. Н.** — д.ф.-м.н. (Великобритания)  
**Дукельский К.В.** — д.т.н., доцент (Россия)  
**Ефимов Д.В.** — PhD (Франция)  
**Заколдаев Д.А.** — к.т.н., доцент (Россия)  
**Иорш И.В.** — д.ф.-м.н. (Россия)  
**Кармановский Н.С.** (заместитель главного редактора) — к.т.н. доцент (Россия)  
**Комаров И.И.** — к.ф.-м.н., доцент (Россия)  
**Матвеев Ю.Н.** — д.т.н. профессор (Россия)  
**Никоноров Н.В.** — д.т.н., профессор (Россия)  
**Полищук Г.С.** — к.т.н. (Россия)  
**Романов А.Е.** — д.ф.-м.н., профессор (Россия)  
**Сидоркина И.Г.** — д.т.н., профессор (Россия)  
**Степанов О.А.** — д.т.н., профессор (Россия)  
**Татарникова Т.М.** — д.т.н., доцент (Россия)  
**Тимофеев А. В.** — д.т.н. (Казахстан)  
**Успенская М.В.** — д.т.н., профессор (Россия)  
**Цыпкин А. Н.** — д.ф.-м.н. (Россия)

Ответственный секретарь — И.В. Малькова

Учредитель: Университет ИТМО

Издание зарегистрировано в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций (свидетельство ПИ №ФС77-67990 от 6 декабря 2016 г.).

Языки журнала: русский, английский.

ISSN 2226-1494 (print version), ISSN 2500-0373 (online version)

Англоязычное название: «Scientific and Technical Journal of Information Technologies, Mechanics and Optics». Транслитерированное название: «Nauchno-Tekhnicheskii Vestnik Informatsionnykh Tekhnologii, Mekhaniki i Optiki». Сокращенное название: «Sci.Tech. J. Inf. Technol. Mech. Opt.»

Журнал включен в каталог периодических изданий Ulrich's Periodical Directory.

Журнал входит в крупнейшие международные реферативные и наукометрические базы Scopus, EBSCO, ProQuest, EastView, ИВИС и другие.

Журнал входит в утвержденный Высшей аттестационной комиссией «Перечень ведущих рецензируемых научных журналов и изданий, выпускаемых в Российской Федерации, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени доктора и кандидата наук».

Периодичность издания — 6 выпусков в год.

**Плата за публикации и редактирование не взимается.**

Редакторы: Ю.В. Михайлова, Л.Н. Точилина, Т.В. Точилина  
Перевод — Н.А. Корнев  
Компьютерная верстка — Е.С. Егорова  
Сайт — К.В. Малькова

Оф. 2136, Университет ИТМО, ул. Ломоносова, д. 9,  
Санкт-Петербург, Российская Федерация, 191002

Телефон (812) 480 02 75  
<http://ntv.ifmo.ru>  
E-mail: [ntvitmo@itmo.ru](mailto:ntvitmo@itmo.ru)

Статьи Журнала доступны по лицензии Creative Commons  
«Attribution-NonCommercial» 4.0 Всемирная.

#### EDITORIAL BOARD

**Sergey V. Bezzateev** — D.Sc., Associate Professor (Russia)  
**Andrey V. Belikov** — D.Sc., Professor (Russia)  
**Igor A. Bessmertny** — D.Sc., Professor (Russia)  
**Alexey A. Bobtsov** — D.Sc., Professor (Russia)  
**Alexandr V. Boukhanovsky** — D.Sc. (Russia)  
**Konstantin N. Volkov** — D.Sc. (Great Britain)  
**Konstantin V. Dukelskii** — D.Sc., Associate Professor (Russia)  
**Denis V. Efimov** — PhD (France)  
**Daniil A. Zakoldaev** — PhD, Associate Professor (Russia)  
**Ivan V. Iorsh** — D.Sc. (Russia)  
**Nikolai S. Karmanovskiy** (Deputy Chief Editor) — PhD, Associate Professor (Russia)  
**Igor I. Komarov** — PhD, Associate Professor (Russia)  
**Iuri N. Matveev** — D.Sc., Professor (Russia)  
**Nikolay V. Nikonov** — D.Sc., Professor (Russia)  
**Grigorii S. Polishchuk** — PhD (Russia)  
**Alexey E. Romanov** — D.Sc., Professor (Russia)  
**Irina G. Sidorkina** — D.Sc., Professor (Russia)  
**Oleg A. Stepanov** — D.Sc., Professor (Russia)  
**Tatiana M. Tatarnikova** — D.Sc., Associate Professor (Russia)  
**Andrey V. Timofeev** — Dr. habil. Sc. (Kazakhstan)  
**Mayya V. Uspenskaya** — D.Sc., Professor (Russia)  
**Anton N. Tsyppkin** — D.Sc., Professor (Russia)

Executive secretary — Irina V. Malkova

Founder: ITMO University

Publication is registered by the Federal Inspectorate Service for Communication, Information Technologies and Communication Media, certificate PI FS77-67 990 dated December, 6th, 2016.

Languages of the journal: Russian, English

ISSN 2226-1494 (print version), ISSN 2500-0373 (online version)

The English title is "Scientific and Technical Journal of Information Technologies, Mechanics and Optics".

Transliterated title is "Nauchno-Tekhnicheskii Vestnik Informatsionnykh Tekhnologii, Mekhaniki i Optiki".

Abbreviated parallel title: "Sci.Tech. J. Inf. Technol. Mech. Opt."

The journal is catalogued in Ulrich's Periodical Directory.

The journal is included in the largest international abstract and scientometric databases Scopus, EBSCO, ProQuest, East View, EVIS and others.

The journal is included in the "List of leading scientific journals and periodicals under review in the Russian Federation, where the main scientific results of theses for the PhD and doctor of sciences degree must be published" approved by the Higher Attestation Commission.

Publication frequency is 6 times a year.

**Publication and editing are free of charge.**

Editors: Yulia V. Mikhailova, Lyubov N. Tochilina, Tatiana V. Tochilina  
English Language Technical Editor — Nikolay A. Kornev  
Computer layout — Ekaterina S. Egorova  
Website — Kseniia V. Malkova

of. 2136, ITMO University, Lomonosova Street, 9,  
St.Petersburg, 191002, Russia

Phone (812) 480 02 75  
<http://ntv.ifmo.ru/en>  
E-mail: [ntvitmo@itmo.ru](mailto:ntvitmo@itmo.ru)



Подписано к печати 01.08.2023 Тираж 350 экз. Заказ № 94 (146)

Издание Университета ИТМО

Отпечатано в ООО «Университетские телекоммуникации»

Адрес: 199034, Санкт-Петербург, Биржевая линия, д. 16

© Университет ИТМО

ОПТИЧЕСКИЕ СИСТЕМЫ И ТЕХНОЛОГИИ  
OPTICAL ENGINEERING

doi: 10.17586/2226-1494-2023-23-4-669-675

УДК 543.421/.424

Определение типа действия ингибиторов гидратообразования  
по их инфракрасным спектрамЮлия Сергеевна Ворожцова<sup>1</sup>✉, Татьяна Николаевна Носенко<sup>2</sup>,  
Майя Валерьевна Успенская<sup>3</sup><sup>1,2,3</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация<sup>1</sup> [leta-x@mail.ru](mailto:leta-x@mail.ru)✉, <https://orcid.org/0009-0006-1359-1235><sup>2</sup> [tata-nostra@yandex.ru](mailto:tata-nostra@yandex.ru), <https://orcid.org/0000-0003-4159-133X><sup>3</sup> [mv\\_uspenskaya@itmo.ru](mailto:mv_uspenskaya@itmo.ru), <https://orcid.org/0000-0003-2510-2639>**Аннотация**

**Введение.** Исследованы отличия инфракрасных спектров ингибиторов гидратообразования термодинамического и кинетического типов действия. Предложен метод определения типа действия ингибитора гидратообразования по его инфракрасному спектру. Актуальность предложенного метода обусловлена его экспрессностью в сравнении с тестированием ингибиторов в рамках лабораторных испытаний. **Метод.** В работе предложено использовать метод инфракрасной спектроскопии с преобразованием Фурье. Метод позволяет получить данные о молекулярных и межмолекулярных взаимодействиях исследуемого вещества. Спектры, полученные в режиме нарушенного полного внутреннего отражения, проанализированы методом главных компонент и регрессионным методом проекции на латентные структуры. Эти методы относятся к хемометрическим методам анализа и позволяют выделить ключевые особенности составов ингибиторов, влияющих на механизм их действия.

**Основные результаты.** Выборки инфракрасных спектров исследуемых ингибиторов разделены на две группы, которые представляют различные типы действия ингибиторов. Для кинетических ингибиторов выявлено характерное соотношение во внутреннем стандарте спектра амплитуд колебаний функциональных групп –ОН и N–H. Показано, что решающим фактором разделения ингибиторов на группы является различие резонансных частот колебаний, соответствующих валентным колебаниям связей C–O, C–N, и резонансных частот колебаний фторорганических соединений. Отмечена схожесть амплитудных значений сигнала в указанных спектральных областях. Для группы термодинамических ингибиторов наиболее значимыми полосами инфракрасного спектра оказались полосы симметричных и асимметричных валентных колебаний связей C–H в группах CH<sub>2</sub> и CH<sub>3</sub>. Отмечено значительное повышение амплитуды в спектральном диапазоне 2950–2750 см<sup>-1</sup> по сравнению с амплитудой сигнала в областях 3300–3400 см<sup>-1</sup> и 1200–1100 см<sup>-1</sup>, также обнаруженных в спектрах данной группы ингибиторов. Методом проекции на латентные структуры построена регрессионная модель для определения типа действия исследуемых ингибиторов. **Обсуждение.** Предложенный метод позволит проводить экспресс-анализ типа действия ингибиторов гидратообразования. Результаты работы могут найти применение в нефтепромышленной химии для определения типа действия ингибиторов гидратообразования, используемых для предотвращения образования газогидратов при добыче, подготовке или транспортировке углеводородного сырья.

**Ключевые слова**

ингибиторы гидратообразования, газогидраты, инфракрасная спектроскопия, метод главных компонент, метод проекции на латентные структуры, хемометрические методы анализа

**Благодарности**

Работа была выполнена на базе Центра химической инженерии Университета ИТМО.

**Ссылка для цитирования:** Ворожцова Ю.С., Носенко Т.Н., Успенская М.В. Определение типа действия ингибиторов гидратообразования по их инфракрасным спектрам // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 669–675. doi: 10.17586/2226-1494-2023-23-4-669-675

## Determination of the action type of hydrate formation inhibitors by their infrared spectra

Iuliia S. Vorozhtsova<sup>1</sup>, Tatiana N. Nosenko<sup>2</sup>, Mayya V. Uspenskaya<sup>3</sup>

<sup>1,2,3</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>1</sup> leta-x@mail.ru, <https://orcid.org/0009-0006-1359-1235>

<sup>2</sup> tata-nostra@yandex.ru, <https://orcid.org/0000-0003-4159-133X>

<sup>3</sup> mv\_ustpenskaya@itmo.ru, <https://orcid.org/0000-0003-2510-2639>

### Abstract

In this paper, differences of infrared spectra of hydrate formation inhibitors of thermodynamic and kinetic action types were investigated. The method was proposed for determining the action type of a hydrate formation inhibitor by its infrared spectrum. The relevance of the proposed method is due to its expressiveness in comparison with the testing of inhibitors in laboratory tests. It is proposed to use the method of Fourier transform infrared spectrometry. The method allows us to obtain data on the molecular and intermolecular interactions of the substance under study. The spectra obtained in the mode of attenuated total reflection were analyzed by the principal component analysis and the regression method of projection on latent structures, which are related to chemometric methods of analysis and make it possible to identify the key features of the inhibitor compositions that affect the mechanism of their action. The separation of the sample of infrared spectra of the studied inhibitors into two subgroups, which represent two different types of inhibitor action, was obtained. The principal component analysis makes it possible to identify the key features of the compositions of reagents that affect the mechanism of their action. For kinetic inhibitors, a characteristic ratio of the amplitudes of vibrations of the –OH and N–H functional groups in the internal standard of the spectrum was revealed. It is shown that the decisive factor in the division of inhibitors into groups is the difference between the resonant vibration frequencies corresponding to the valence vibrations of C–O, C–N bonds and the resonant vibration frequencies of organofluorine compounds. At the same time, the similarity in the amplitude of the indicated spectral regions was noted. For the group of thermodynamic inhibitors, the most influential bands in the IR spectrum were the bands of symmetric and asymmetric stretching vibrations of the C–H bonds in the CH<sub>2</sub> and CH<sub>3</sub> groups. There was a significant increase in the amplitude in the spectral range of 2950–2750 cm<sup>-1</sup> compared with the signal amplitude in the regions of 3300–3400 cm<sup>-1</sup> and 1200–1100 cm<sup>-1</sup>, also found in the spectra of this group of inhibitors. The method of projection on latent structures was used to develop a regression model to determine the mechanism of action of the studied inhibitors. The proposed method allows for express determination of the action type of hydrate formation inhibitors. Results could be used in oilfield chemistry to determine the action type of hydrate formation inhibitors used to prevent the formation of gas hydrates during the production, preparation or transportation of hydrocarbons.

### Keywords

hydrate formation inhibitors, gas hydrates, Fourier-transform infrared spectroscopy, principal component analysis, projection on latent structures, chemometric methods of analysis

### Acknowledgements

The study was facilitated through the provision of materials and resources by the Center for Chemical Engineering of the ITMO University.

**For citation:** Vorozhtsova Iu.S., Nosenko T.N., Uspenskaya M.V. Determination of the action type of hydrate formation inhibitors by their infrared spectra. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 669–675 (in Russian). doi: 10.17586/2226-1494-2023-23-4-669-675

### Введение

Газовые гидраты, кроме нетрадиционного стратегически важного источника углеводородного энергетического сырья [1–3], а также способа хранения и транспортировки природного/попутного нефтяного газа [4, 5], представляют собой один из типов отложений, формирующихся при добыче и транспортировке нефтегазовых продуктов и негативно влияющих на обеспечение стабильности потока углеводородного сырья. Ледоподобные гидратные отложения в виде, похожем на спрессованный снег, образуют пробки в трубопроводах и другом технологическом оборудовании, что снижает технико-экономические показатели производства в целом. К веществам, способным образовывать газовые гидраты, как правило, относят такие низкомолекулярные углеводороды, как метан CH<sub>4</sub>, этилен C<sub>2</sub>H<sub>4</sub>, этан C<sub>2</sub>H<sub>6</sub>, пропан C<sub>3</sub>H<sub>8</sub>, изобутан изо-C<sub>4</sub>H<sub>10</sub> [1, 6], некоторые газы Ar, N<sub>2</sub>, O<sub>2</sub>, Cl<sub>2</sub>, CS<sub>2</sub> и гидрофильные соединения, например, CO<sub>2</sub>, SO<sub>2</sub> и др.,

которые слабо взаимодействуют с водой и не могут препятствовать клатратообразованию.

Среди методов удаления и предотвращения гидратообразования выделяют физические методы, включающие поддержание значений температур и давлений вне зоны стабильности газового гидрата, обогрев и/или изоляция оборудования, осушка газовой фазы [7]. Широкое распространение получили химические методы предотвращения образования и отложения газогидратов, которые реализуются за счет применения термодинамических ингибиторов и ингибиторов гидратообразования низких дозировок. В зависимости от механизма действия ингибиторы гидратообразования низких дозировок делятся на кинетические и антигломеранты. Отличительные особенности механизмов ингибирования образования и отложения газогидратов этих классов веществ заключены в их названиях. Термодинамические ингибиторы влияют на изменение термодинамических свойств среды, тем самым смещая условия стабильности гидрата в область их неустойчи-

вости. Кинетические ингибиторы влияют на скорость образования зародышей газовых гидратов, снижая ее, или замедляют рост уже образовавшихся кристаллов в течение определенного периода времени, зависящего от переохлаждения. Антиагломеранты не позволяют слипаться образовавшимся кристаллам гидратов и собираться им в более крупные массы, которые становятся причиной закупорки оборудования. Актуальны также комплексные составы, способные кроме борьбы с гидратами решать и другие задачи, связанные с осложнениями при добыче углеводородов, например, коррозия нефтепромыслового оборудования. В работах [8–10] Казанского (Приволжского) федерального университета описан перспективный класс кинетических ингибиторов комплексного действия, в составе которых содержатся добавки на основе водорастворимых полиуретанов, природных полимеров и касторового масла. Они предназначены для подавления процессов гидратообразования и коррозии одновременно. В [11, 12] рассмотрен новый тип ингибиторов двойного действия, который отличается способностью к быстрому разложению под действием факторов окружающей среды, а также малой токсичностью, что обуславливается его составом на основе природного полисахарида хитозана.

Как показано в работах [13, 14], ингибиторы разных типов действия проявляют различную эффективность в одних и тех же условиях. Возникает вопрос: какие ключевые особенности состава ингибитора влияют на его отнесение к тому или иному механизму действия? Попытка ответа на данный вопрос предпринята в настоящей работе при помощи инфракрасной (ИК) спектроскопии и хемометрических методов анализа данных.

Во многих научных работах изучены хемометрические методы, которые позволяют проводить качественный и количественный анализ экспериментальных данных, полученных ИК спектроскопией.

При исследовании реагентов в нефтепромысловой химии есть опыт применения хемометрического метода главных компонент (МГК) для исследования корреляции между эффективностью нескольких ингибиторов парафинообразования и составом сырых нефтей, определенным с помощью ИК фурье-спектроскопии [15]. В работе [16] разработаны восемь хемометрических моделей для прогнозирования наличия ароматических соединений, смол и асфальтенов в составах пятидесяти колумбийских нефтей с использованием ИК фурье-спектров, полученных в режиме нарушенного полного внутреннего отражения (НПВО).

Настоящая работа включает в себя исследование различных по составу ингибиторов термодинамического и кинетического типа действия с помощью ИК спектроскопии для решения следующих задач: определения основных функциональных групп в составе ингибиторов гидратообразования, имеющих наибольшее влияние на механизм действия реагентов; получения хемометрической модели по определению типа ингибитора гидратообразования в зависимости от его механизма действия.

## Методы и материалы

В рамках работы проведен анализ состава 17 ингибиторов разных российских производителей, активно применяемых в отечественной промышленности при добыче углеводородов, для выявления структурных особенностей в зависимости от механизма ингибирования образования и отложения газовых гидратов. Среди исследуемых реагентов выбраны 11 образцов кинетического и 6 термодинамического типов действия.

Действие термодинамических ингибиторов гидратообразования проявляется в изменении основных термодинамических свойств системы с жидкими средами в целях сдвига равновесия условий образования газогидратов в сторону более низких температур или высоких давлений. Среди них встречаются как неорганические, так и органические вещества. Основные представители неорганических ингибиторов — соли, например: хлорид натрия, хлорид кальция, формиат калия и ацетат натрия. При поступлении в водный раствор газа соль в процессе ионизации образует ионы с положительными и отрицательными зарядами, притягивающие к себе диполи воды. Вследствие электростатического характера взаимодействия между диполями воды и ионами ингибитора, которое является при этом более сильным, чем взаимодействие между молекулами газа и воды, существующая квазикристаллическая структура воды нарушается, что приводит к снижению растворимости газа. Подобное явление известно как высаливание и возникает в случае отличия полярности компонентов раствора. Органические вещества, используемые в качестве термодинамических ингибиторов гидратообразования, представлены в основном метанолом и гликолями. Механизм действия органических веществ подобен неорганическим. Органические ингибиторы и вода взаимно растворяются, снижая содержание свободной воды в потоке газа, в результате нарушаются благоприятные условия гидратообразования. Такие ингибиторы в водном растворе снижают термодинамическую активность воды, что обеспечивает сдвиг равновесных условий образования гидратов в область их неустойчивого состояния. Ключевые представители органических термодинамических ингибиторов — водорастворимые полярные органические соединения: метанол, моноэтиленгликоль, диэтиленгликоль. Другие химические вещества, исследуемые в качестве термодинамических ингибиторов гидратообразования, включают диметилформамид и N-метилпирролидон, этаноламины и спирты, образующие азеотропы с водой, например, изопропанол.

Кинетические ингибиторы задерживают зарождение газогидратов и рост кристаллов. Такие ингибиторы имеют две ключевые структурные особенности. Во-первых, ингибитору требуются функциональные группы, которые могут установить водородную связь с молекулами воды или поверхностью частиц газогидратов, такие группы, как правило, являются амидными. Вторая ключевая особенность представляет собой гидрофобную группу, прилегающую или непосредственно присоединенную к каждой из амидных

групп. Соответственно этим структурным особенностям предлагаются два основных механизма, работающих в различных условиях. Первый предполагает, что полимеры кинетических ингибиторов гидратообразования через гидрофобные взаимодействия способны возмущать структуру воды до такой степени, что частицы газогидрата не могут расти до критического размера ядра, при котором рост становится спонтанным. Второй механизм предполагает, что полимеры кинетических ингибиторов газогидратов адсорбируются на поверхности растущей частицы гидрата, что ограничивает их рост и может деформировать полости гидрата. Полимерные кинетические ингибиторы могут адсорбироваться различными способами на поверхности гидратов, например, амидные группы закрепляют полимер на поверхности гидратов посредством водородных связей [17].

Особенности состава ингибиторов могут быть определены с помощью оптических методов, например, ИК спектроскопии с преобразованием Фурье, которая регистрирует колебательные спектры веществ.

ИК спектры исследуемых образцов были записаны на фурье-спектрометре BRUKER Tenzor 37, в режиме НПВО; на кристалле KBr с алмазным напылением; со сканированием (32 скана); при разрешении  $2 \text{ см}^{-1}$ .

Специфика состава ингибиторов гидратообразования — высокая степень летучести. При получении ИК спектров таких образцов необходимо учитывать данное свойство. Проведение регистрации ИК спектра связано с затратой небольшого, но достаточного для изменения образца, времени. Для предотвращения данного изменения и получения корректного спектра полного состава образца ингибитора в процессе регистрации спектра ингибитор постоянно добавлялся на поверхность кристалла НПВО. После получения ИК спектра полного состава образец больше не добавляли и каплю образца выдерживали на кристалле до полного испарения летучей части, снимая в процессе его испарения ИК спектры. В анализ включены первые и последние спектры. Полученные ИК спектры были предварительно обработаны: проведена корректировка базовой линии и векторное нормирование.

Для анализа спектральных данных использован МГК, реализованный в программном пакете The Unscrambler 9.7. МГК-анализ — метод, который позволяет уменьшить размерность больших наборов данных без потери информации, идентифицируя доминирующие шаблоны и соотношения между переменными с использованием собственных значений матрицы [18, 19]. МГК позволяет выделить ключевые особенности составов реагентов, влияющих на механизм их действия. Данным методом образцы были проанализированы в среднем ИК диапазоне  $4000\text{--}600 \text{ см}^{-1}$ .

Для возможности определения типа ингибитора по его спектру выполнен регрессионный анализ методом проекций на латентные структуры (ПЛС), который относится к контролируемым методам хемометрического анализа. Этот метод объединяет в себе подходы МГК и множественной регрессии. Таким образом, на первом этапе ПЛС выделяет набор скрытых факторов, которые объясняют максимально возможное количество

ковариаций между независимыми и зависимыми переменными, далее на этапе регрессии предсказываются значения зависимых переменных с использованием декомпозиции независимых переменных.

### Результаты и обсуждение

В процессе оптического исследования ингибиторов ИК спектры полного состава показали преобладание функциональных групп, соответствующих метанолу. Сигнал от данных функциональных групп перекрывает колебания, относящиеся к другим составляющим ингибитора. Для выявления таких колебаний был зарегистрирован ИК спектр ингибитора без метанола посредством его испарения.

Для анализа использованы ИК спектры, которые не содержат летучих органических соединений. По результатам анализа в The Unscrambler нормированных спектральных данных МГК для матрицы счетов получено распределение исследуемых образцов ингибиторов гидратообразования по спектрам в зависимости от их механизма действия в осях главных компонент: первой (ГК1) и второй (ГК2) (рис. 1).

На рис. 1 видно, что в зависимости от информации, содержащейся в спектральных данных и их составов, реагенты распределены по группам, соответствующим их типам действия. На ГК1 приходится 66 % объясненной дисперсии выборки, а на ГК2 — 12 %.

Для интерпретации наблюдаемого разделения с точки зрения спектральных сигналов рассмотрим график матрицы нагрузок, соответствующий главным компонентам ГК1 и ГК2 (рис. 2, а). Матрица нагрузок — матрица, элементами которой являются коэффициентами или факторными нагрузками для соответствующей ГК на некоторую переменную и, фактически, мерой линейной связи между интенсивностью сигнала на конкретном волновом числе и значением координаты образца по оси, соответствующей ГК.

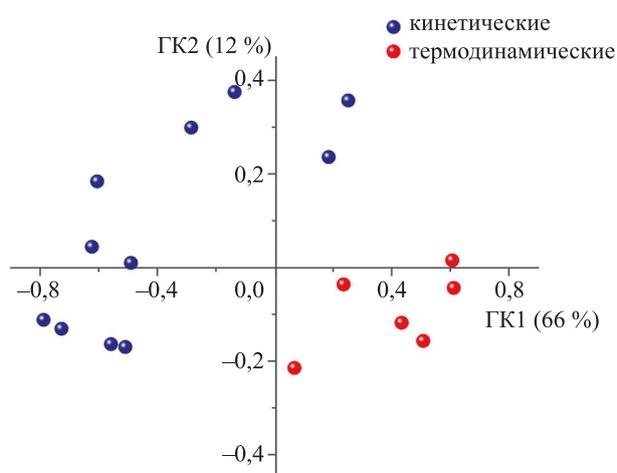


Рис. 1. Матрица счетов для выборки инфракрасных спектров образцов ингибиторов гидратообразования термодинамического и кинетического типов действия

Fig. 1. Score matrix for IR spectra sampling of hydrate formation inhibitors of thermodynamic and kinetic action types

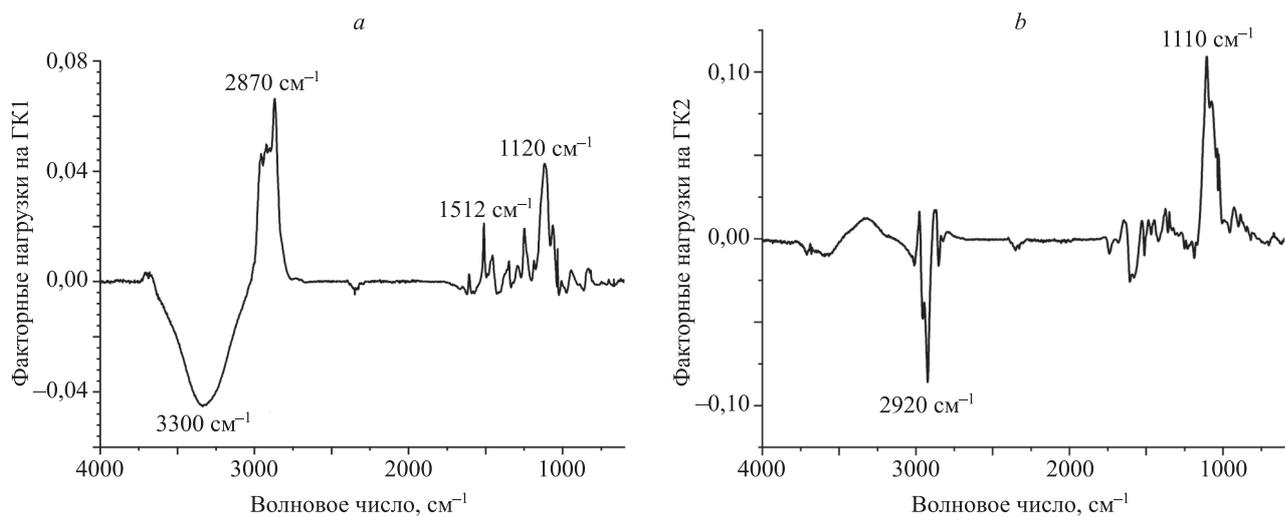


Рис. 2. Спектральные графики матриц нагрузок выборок инфракрасных спектров образцов ингибиторов для ГК1 (66 %) и реагентов для ГК2 (12 %) гидратообразования термодинамического и кинетического типов действия по первой (а) и по второй (б) главным компонентам

Fig. 2. Spectral plots of load matrices for samples of infrared spectra of inhibitor samples for GK1 (66 %) and reagents for GK2 (12 %) samples of hydrate formation of thermodynamic and kinetic action types for the first (a) and second (b) main components

По матрице нагрузок для ГК1, отражающей распределение волновых чисел выборки ИК спектров образцов, можно сделать вывод о наличии двух ключевых областей спектра. Отрицательные нагрузки по ГК1 имеет спектральная область, соответствующая интервалу  $3400\text{--}3300\text{ см}^{-1}$ , в этом диапазоне находятся колебания групп  $\text{--OH}$  и  $\text{N--H}$ . При наличии данных групп в анализируемом веществе сигналы от них могут перекрывать друг друга. Широкая интенсивная полоса в указанном диапазоне соответствует валентному колебанию связи  $\text{--OH}$ , при этом более узкая полоса соответствует симметричным валентным колебаниям групп  $\text{N--H}$  [20]. Основные положительные значения нагрузок по ГК1 приходятся на пики в диапазоне  $2950\text{--}2750\text{ см}^{-1}$ , что соответствует симметричным и асимметричным валентным колебаниям связей  $\text{C--H}$  в группах  $\text{CH}_2$  и  $\text{CH}_3$ . Таким образом, как видно из рис. 2, а, для кинетических ингибиторов характерен пик в интервале  $3400\text{--}3300\text{ см}^{-1}$ , что говорит о более значительном содержании функциональных групп  $\text{--OH}$  в кинетических образцах по сравнению с термодинамическими. Кроме того, в кинетических образцах также, как и в термодинамических, присутствуют группы  $\text{C--H}$ , однако их соотношение с группами  $\text{--OH}$  в среднем меньше, чем в термодинамических.

Для более подробной интерпретации рассмотрим также матрицу нагрузок по ГК2 (рис. 2, б).

Наиболее значительный вклад в значения отрицательных нагрузок по ГК2 вносит диапазон  $2900\text{--}3000\text{ см}^{-1}$ , соответствующий валентным колебаниям связей  $\text{C--H}$ , а в положительные —  $1100\text{--}1200\text{ см}^{-1}$  — полоса валентных колебаний  $\text{C--O}$ ,  $\text{C--N}$  или колебания фторорганических соединений.

Наблюдаемая картина позволяет уточнить выводы о соотношении наиболее влияющих функциональных групп в ингибиторах гидратообразования различного действия.

Группа термодинамических ингибиторов характеризуется наличием более амплитудных колебаний в диапазоне  $2950\text{--}2750\text{ см}^{-1}$  по сравнению с колебаниями в интервалах  $3300\text{--}3400\text{ см}^{-1}$  и  $1200\text{--}1100\text{ см}^{-1}$ .

Кинетические ингибиторы, в свою очередь, характеризуются тремя схожими по значению амплитуды областями, соответствующими колебаниям в диапазонах  $2950\text{--}2750\text{ см}^{-1}$ ,  $3300\text{--}3400\text{ см}^{-1}$  и  $1200\text{--}1100\text{ см}^{-1}$ .

Поскольку МГК показывает наличие в выборке отдельных групп, соответствующих ингибиторам различного типа действия, то возможно построить регрессионную модель. Методом ПЛС построена регрессионная модель, в которой ингибиторам кинетического типа действия было присвоено значение категориальной переменной равное 1, а термодинамическим — 2.

В процессе создания модели валидация выборки ИК спектров ингибиторов осуществлялась с помощью метода Cross validation при исключении двух образцов из обучающей выборки. Для созданной модели получены значения точности  $R^2 = 0,98$ ,  $\text{RMSE} = 0,066$ .

Полученная регрессионная модель была сохранена и использована для прогнозирования механизма действия тестовых образцов — ингибиторов гидратообразования.

Для проверки модели выбраны два образца с условно неопределенным механизмом действия, не включенные в обучающую выборку. Первый образец, реагент 1 (P1), имел термодинамический механизм ингибирования, второй — реагент 2 (P2) — кинетический. Результаты работы модели показаны на рис. 3.

Как видно по результатам работы модели, один из тестируемых ингибиторов, P1, относится к типу 2, т. е. к термодинамическим ингибиторам, его значение по результатам прогнозирования 1,953 с доверительным интервалом 0,116. Механизм действия второго тестируемого реагента, P2, определен как кинетический (тип 1

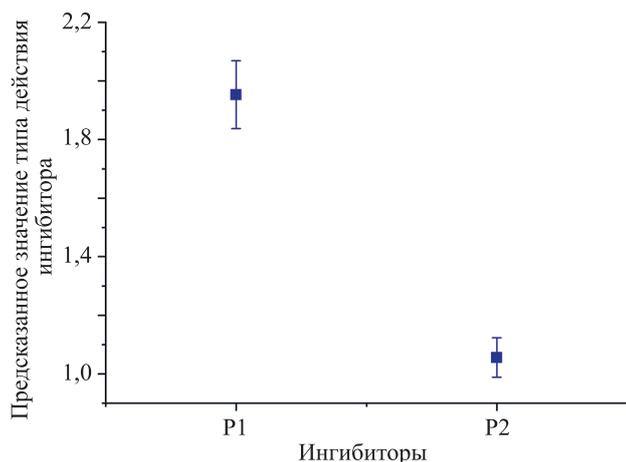


Рис. 3. Результаты работы модели для образцов ингибиторов гидратообразования с неизвестными типами действия.

Образцы ингибиторов: P1 — реагент 1, P2 — реагент 2

Fig. 3. Model performance results for hydrate inhibitors samples with unknown action types.

Inhibitor samples: P1 — Reagent 1, P2 — Reagent 2

в регрессионной модели), со значением 1,056 с доверительным интервалом 0,067.

Таким образом, выделены ключевые особенности составов ингибиторов образования газогидратов, влияющих на механизм их действия, создана регрессионная модель, определяющая термодинамический и кинетический механизмы действия ингибиторов гидратообразования. Разработанный метод может применяться в

нефтепромышленной химии для определения типа действия ингибиторов гидратообразования.

### Заключение

По результатам работы можно сделать следующие выводы.

Инфракрасные спектры кинетических ингибиторов гидратообразования имеют колебания, соотносимые с функциональными группами –ОН и N–H. Данные спектры характеризуются наличием в анализируемых образцах колебаний в диапазонах, соответствующих полосам валентных колебаний C–O и/или C–N, либо колебаниям фторорганических соединений. При этом отмечена схожесть в амплитудах сигнала в интервалах, соответствующих колебаниям, соотносимым с функциональными группами –ОН и N–H, и соответствующих полосам валентных колебаний C–O и/или C–N, либо колебаниям фторорганических соединений.

Для группы термодинамических ингибиторов характерно наличие в инфракрасных спектрах выраженных симметричных и асимметричных валентных колебаний связей C–H в группах CH<sub>2</sub> и CH<sub>3</sub>. Отмечено значительное повышение амплитуды в спектральном диапазоне 2950–2750 см<sup>-1</sup> по сравнению с амплитудой сигнала в интервалах 3300–3400 см<sup>-1</sup> и 1200–1100 см<sup>-1</sup>. Указанные интервалы обнаружены в спектрах данной группы ингибиторов, однако значительно менее выраженные.

Определена возможность создания регрессионной модели, которая может быть использована для определения механизма действия ингибиторов гидратообразования.

### Литература

1. Макогон Ю.Ф. Природные газовые гидраты: распространение, модели образования, ресурсы // Российский химический журнал. 2003. Т. 48. № 3. С. 70–79.
2. Соловьёв В.А. Природные газовые гидраты как потенциальное полезное ископаемое // Российский химический журнал. 2003. Т. 48. № 3. С. 59–69.
3. Makogon Y.F. Natural gas hydrates — A promising source of energy // Journal of Natural Gas Science and Engineering. 2010. V. 2. N 1. P. 49–59. <https://doi.org/10.1016/j.jngse.2009.12.004>
4. Dong H., Wang J., Xie Z., Wang B., Zhang L., Shi Q. Potential applications based on the formation and dissociation of gas hydrates // Renewable and Sustainable Energy Reviews. 2021. V. 143. P. 110928. <https://doi.org/10.1016/j.rser.2021.110928>
5. Thakre N., Jana A.K. Physical and molecular insights to Clathrate hydrate thermodynamics // Renewable and Sustainable Energy Reviews. 2021. V. 135. P. 110150. <https://doi.org/10.1016/j.rser.2020.110150>
6. Якусени В.П. Газогидраты — нетрадиционное газовое сырье, их образование, свойства, распространение и геологические ресурсы // Нефтегазовая геология. Теория и практика. 2013. Т. 8. № 4. С. 12.
7. Грицишин Д.Н., Квон В.Г., Истомин В.А., Минигулов Р.М. Технологии предупреждения гидратообразования в промысловых системах: проблемы и перспективы // Газохимия. 2009. № 10. С. 32–40.
8. Farhadian A., Kudbanov A., Varfolomeev M.A., Dalmazzone D. Waterborne polyurethanes as a new and promising class of kinetic inhibitors for methane hydrate formation // Scientific Reports. 2019. V. 9. P. 9797. <https://doi.org/10.1038/s41598-019-46274-w>
9. Farhadian A., Varfolomeev M.A., Kudbanov A., Rezaeisadat M., Nurgaliev D.K. Waterborne polymers as kinetic/anti-agglomerant methane hydrate and corrosion inhibitors: A new and promising

### References

1. Makogon Yu.F. Natural gas hydrates: scope of distribution, formation models, resources. *Rossiiskij himicheskij zhurnal*, 2003, vol. 48, no. 3, pp. 70–79. (in Russian)
2. Solov'ev V.A. Natural gas hydrates as a potential mineral resource. *Rossiiskij himicheskij zhurnal*, 2003, vol. 48, no. 3, pp. 59–69. (in Russian)
3. Makogon Y.F. Natural gas hydrates — A promising source of energy. *Journal of Natural Gas Science and Engineering*, 2010, vol. 2, no. 1, pp. 49–59. <https://doi.org/10.1016/j.jngse.2009.12.004>
4. Dong H., Wang J., Xie Z., Wang B., Zhang L., Shi Q. Potential applications based on the formation and dissociation of gas hydrates. *Renewable and Sustainable Energy Reviews*, 2021, vol. 143, pp. 110928. <https://doi.org/10.1016/j.rser.2021.110928>
5. Thakre N., Jana A.K. Physical and molecular insights to Clathrate hydrate thermodynamics. *Renewable and Sustainable Energy Reviews*, 2021, vol. 135, pp. 110150. <https://doi.org/10.1016/j.rser.2020.110150>
6. Yakutseni V.P. Gas hydrates — unconventional gas sources, their formation, properties, distribution and geological resources. *Petroleum Geology — Theoretical and Applied Studie*, 2013, vol. 8, no. 4, pp. 12. (in Russian)
7. Gritcishin D.N., Kvon V.G., Istomin V.A., Minigulov R.M. Technologies for the prevention of hydrate build-up in field-based systems: challenges and opportunities. *Gazohimiya*, 2009, no. 10, pp. 32–40. (in Russian)
8. Farhadian A., Kudbanov A., Varfolomeev M.A., Dalmazzone D. Waterborne polyurethanes as a new and promising class of kinetic inhibitors for methane hydrate formation. *Scientific Reports*, 2019, vol. 9, pp. 9797. <https://doi.org/10.1038/s41598-019-46274-w>
9. Farhadian A., Varfolomeev M.A., Kudbanov A., Rezaeisadat M., Nurgaliev D.K. Waterborne polymers as kinetic/anti-agglomerant methane hydrate and corrosion inhibitors: A new and promising

- strategy for flow assurance // *Journal of Natural Gas Science and Engineering*. 2020. V. 77. P. 103235. <https://doi.org/10.1016/j.jngse.2020.103235>
10. Farhadian A., Varfolomeev M.A., Kudbanov A., Gallyamova S.R. A new class of promising biodegradable kinetic/anti-agglomerant methane hydrate inhibitors based on castor oil // *Chemical Engineering Science*. 2019. V. 206. P. 507–517. <https://doi.org/10.1016/j.ces.2019.05.055>
  11. Farhadian A., Varfolomeev M.A., Shaabani A., Zaripova Y.F., Yarkovoi V.V., Khayarov K.R. Inhibition performance of chitosan-graft-polyacrylamide as an environmentally friendly and high-cloud-point inhibitor of nucleation and growth of methane hydrate // *Crystal Growth & Design*. 2020. V. 20. N 3. P. 1771–1778. <https://doi.org/10.1021/acs.cgd.9b01500>
  12. Farhadian A., Varfolomeev M.A., Shaabani A., Nasiri S., Vakhitov I., Zaripova Y.F., Yarkovoi V.V., Sukhov A.V. Sulfonated chitosan as green and high cloud point kinetic methane hydrate and corrosion inhibitor: Experimental and theoretical studies // *Carbohydrate Polymers*. 2020. V. 236. P. 116035. <https://doi.org/10.1016/j.carbpol.2020.116035>
  13. Фаресов А.В., Пономарев А.И. Исследование эффективности ингибиторов гидратообразования кинетического типа // Нефтегазовое дело. 2013. Т. 11. № 4. С. 86–95.
  14. Kunakova A.M., Usmanova F.G., Vorozhtsova I.S., Lanchuk I.V. Approaches to the selection of effective inhibitors of gas hydrate formation // *Proc. of the SPE Russian Petroleum Technology Conference*. 2019. P. SPE-196781-MS. <https://doi.org/10.2118/196781-MS>
  15. Perez P., Gurnon K., Chichak K., McDermott J., de Paulo J., Peng W., Xie X. Mitigating wax deposition from crude oils: correlations between physical-chemical properties of crude oils and the performance of wax inhibitors // *Proc. of the Offshore Technology Conference*. 2016. P. OTC-27255-MS. <https://doi.org/10.4043/27255-MS>
  16. Meléndez L.V., Lache A., Orrego-Ruiz J.A., Pachón Z., Mejía-Ospino E. Prediction of the SARA analysis of colombian crude oils using ATR-FTIR Spectroscopy and chemometric methods // *Journal of Petroleum Science and Engineering*. 2012. V. 90–91. P. 56–60. <https://doi.org/10.1016/j.petrol.2012.04.016>
  17. Келланд М. Промысловая химия в нефтегазовой отрасли. СПб.: ЦОП «Профессия», 2015. 608 с.
  18. Wold S., Esbensen K., Geladi P. Principal component analysis // *Chemometrics and Intelligent Laboratory Systems*. 1987. V. 2. N 1-3. P. 37–52. [https://doi.org/10.1016/0169-7439\(87\)80084-9](https://doi.org/10.1016/0169-7439(87)80084-9)
  19. Родионова О.Е., Померанцев А.Л. Хемометрика: достижения и перспективы // *Успехи химии*. 2006. Т. 75. № 4. С. 302–321.
  20. Тарасевич Б.Н. ИК спектры основных классов органических соединений. Справочные материалы / МГУ имени М.В. Ломоносова, химический факультет, кафедра органической химии. М., 2012. 54 с.
  - strategy for flow assurance. *Journal of Natural Gas Science and Engineering*, 2020, vol. 77, pp. 103235. <https://doi.org/10.1016/j.jngse.2020.103235>
  10. Farhadian A., Varfolomeev M.A., Kudbanov A., Gallyamova S.R. A new class of promising biodegradable kinetic/anti-agglomerant methane hydrate inhibitors based on castor oil. *Chemical Engineering Science*, 2019, vol. 206, pp. 507–517. <https://doi.org/10.1016/j.ces.2019.05.055>
  11. Farhadian A., Varfolomeev M.A., Shaabani A., Zaripova Y.F., Yarkovoi V.V., Khayarov K.R. Inhibition performance of chitosan-graft-polyacrylamide as an environmentally friendly and high-cloud-point inhibitor of nucleation and growth of methane hydrate. *Crystal Growth & Design*, 2020, vol. 20, no. 3, pp. 1771–1778. <https://doi.org/10.1021/acs.cgd.9b01500>
  12. Farhadian A., Varfolomeev M.A., Shaabani A., Nasiri S., Vakhitov I., Zaripova Y.F., Yarkovoi V.V., Sukhov A.V. Sulfonated chitosan as green and high cloud point kinetic methane hydrate and corrosion inhibitor: Experimental and theoretical studies. *Carbohydrate Polymers*, 2020, vol. 236, pp. 116035. <https://doi.org/10.1016/j.carbpol.2020.116035>
  13. Faresov A.V., Ponomarev A.I. Studies on the efficiency of kinetic hydrate inhibitors. *Petroleum Engineering*, 2013, vol. 11, no. 4, pp. 86–95. (in Russian)
  14. Kunakova A.M., Usmanova F.G., Vorozhtsova I.S., Lanchuk I.V. Approaches to the selection of effective inhibitors of gas hydrate formation. *Proc. of the SPE Russian Petroleum Technology Conference*, 2019, pp. SPE-196781-MS. <https://doi.org/10.2118/196781-MS>
  15. Perez P., Gurnon K., Chichak K., McDermott J., de Paulo J., Peng W., Xie X. Mitigating wax deposition from crude oils: correlations between physical-chemical properties of crude oils and the performance of wax inhibitors. *Proc. of the Offshore Technology Conference*, 2016, pp. OTC-27255-MS. <https://doi.org/10.4043/27255-MS>
  16. Meléndez L.V., Lache A., Orrego-Ruiz J.A., Pachón Z., Mejía-Ospino E. Prediction of the SARA analysis of colombian crude oils using ATR-FTIR Spectroscopy and chemometric methods. *Journal of Petroleum Science and Engineering*, 2012, vol. 90–91, pp. 56–60. <https://doi.org/10.1016/j.petrol.2012.04.016>
  17. Kelland M.A. *Production Chemicals for the Oil and Gas Industry*. Taylor & Francis, 2014, 454 p.
  18. Wold S., Esbensen K., Geladi P. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, 1987, vol. 2, no. 1-3, pp. 37–52. [https://doi.org/10.1016/0169-7439\(87\)80084-9](https://doi.org/10.1016/0169-7439(87)80084-9)
  19. Rodionova O.Ye., Pomerantsev A.L. Chemometrics: achievements and prospects. *Russian Chemical Reviews*, 2006, vol. 75, no. 4, pp. 271–287. <https://doi.org/10.1070/rc2006v075n04abeh003599>
  20. Tarasevich B.N. *IR Spectra of Major Classes of Organic Compounds*. Reference Source. Moscow, Lomonosov Moscow State University, 2012, 54 p. (in Russian)

### Авторы

**Ворожцова Юлия Сергеевна** — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57215118092](https://orcid.org/0009-0006-1359-1235), <https://orcid.org/0009-0006-1359-1235>, [leta-x@mail.ru](mailto:leta-x@mail.ru)

**Носенко Татьяна Николаевна** — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57190940294](https://orcid.org/0000-0003-4159-133X), <https://orcid.org/0000-0003-4159-133X>, [tata-nostra@yandex.ru](mailto:tata-nostra@yandex.ru)

**Успенская Майя Валерьевна** — доктор технических наук, профессор, директор центра химической инженерии, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 6602189454](https://orcid.org/0000-0003-2510-2639), <https://orcid.org/0000-0003-2510-2639>, [mv\\_uspenskaya@itmo.ru](mailto:mv_uspenskaya@itmo.ru)

### Authors

**Iuliia S. Vorozhtsova** — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57215118092](https://orcid.org/0009-0006-1359-1235), <https://orcid.org/0009-0006-1359-1235>, [leta-x@mail.ru](mailto:leta-x@mail.ru)

**Tatiana N. Nosenko** — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57190940294](https://orcid.org/0000-0003-4159-133X), <https://orcid.org/0000-0003-4159-133X>, [tata-nostra@yandex.ru](mailto:tata-nostra@yandex.ru)

**Mayya V. Uspenskaya** — D.Sc., Professor, Head of Chemical Engineering Center, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 6602189454](https://orcid.org/0000-0003-2510-2639), <https://orcid.org/0000-0003-2510-2639>, [mv\\_uspenskaya@itmo.ru](mailto:mv_uspenskaya@itmo.ru)

Статья поступила в редакцию 06.04.2023  
Одобрена после рецензирования 21.05.2023  
Принята к печати 16.07.2023

Received 06.04.2023  
Approved after reviewing 21.05.2023  
Accepted 16.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-676-684

УДК 535.8

## Спектроскопия комбинационного рассеяния света в исследованиях процессов инактивации бактериальных микроорганизмов

Карина Игоревна Матвеева<sup>1</sup>✉, Анна Анатольевна Кундалевич<sup>2</sup>,  
Анастасия Ивановна Капитунова<sup>3</sup>, Александр Сергеевич Зозуля<sup>4</sup>,  
Станислав Алексеевич Сухих<sup>5</sup>, Анна Владимировна Цибульникова<sup>6</sup>,  
Андрей Юрьевич Зюбин<sup>7</sup>, Илья Геннадьевич Самусев<sup>8</sup>

<sup>1,2,3,4,5,6,7,8</sup> Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация

<sup>1</sup> matveeva.k.i@inbox.ru ✉, <https://orcid.org/0000-0003-2585-163X>

<sup>2</sup> Kundalevich3@mail.ru, <https://orcid.org/0000-0001-7428-6441>

<sup>3</sup> AIKapitunova@mail.ru, <https://orcid.org/0000-0003-0958-6338>

<sup>4</sup> zozul97@mail.ru, <https://orcid.org/0000-0001-5717-8625>

<sup>5</sup> SSukhikh@kantiana.ru, <https://orcid.org/0000-0001-7910-8388>

<sup>6</sup> memorgold@mail.ru, <https://orcid.org/0000-0001-8578-0701>

<sup>7</sup> azubin@mail.ru, <https://orcid.org/0000-0002-9766-1408>

<sup>8</sup> ISamusev@kantiana.ru, <https://orcid.org/0000-0001-5026-7510>

### Аннотация

**Введение.** Спектроскопия комбинационного рассеяния света является одним из перспективных подходов для структурно-функциональных исследований биологических объектов, в том числе бактериальных микроорганизмов. Для бактериального анализа применяются традиционные биохимические тесты и генетические методы. Генетические методы требуют дорогостоящих реагентов, расходных материалов и являются времязатратными. Спектроскопические методы исследования позиционируются как неинвазивные, высокочувствительные и требующие минимальной пробоподготовки образцов. В представленной работе исследована возможность применения метода комбинационного рассеяния света, использующего оптические сенсоры на основе золотых анизотропных наночастиц. Применимость метода продемонстрирована на изучении влияния антибиотика цефалоспоринового ряда широкого спектра действия и экстракта калины (*Viburnum opulus L*) на колонии кишечной палочки (*Escherichia coli*). **Метод.** Применен метод спектроскопии комбинационного рассеяния света с использованием спектрометра Virsa (Renishaw). Усиление сигнала комбинационного рассеяния света проводилось с использованием предложенных авторами двух оригинальных оптических сенсоров. Для создания сенсоров выбраны химический метод осаждения нанозвезд золота на APTES-модифицированные кварцевые стекла и физический метод создания сенсоров, основанный на анодировании титановых поверхностей. **Основные результаты.** Полученные результаты показали высокую чувствительность и информативность предложенного метода. Подтверждена возможность использования метода комбинационного рассеяния света для исследования инактивации бактериальных микроорганизмов. Определены и идентифицированы спектральные полосы комбинационного рассеяния света *Escherichia coli* до и после воздействия экстракта калины и антибиотика в качестве контроля. Установлено уменьшение интенсивности спектральных мод, соответствующих аминокислотам и пуриновым метаболитам, в усредненном спектре комбинационного рассеяния света *Escherichia coli* после воздействия экстракта калины. **Обсуждение.** Впервые проведено исследование антимикробного эффекта водного экстракта плодов калины методом комбинационного рассеяния света. Показано, что применение растительных экстрактов, в том числе экстрактов плодов калины, для инактивации жизнедеятельности колоний бактерий является перспективным подходом по поиску новых альтернативных антибактериальных средств. Полученные результаты хорошо согласуются с уже известными научными исследованиями и подтверждают эффективность предложенного метода.

© Матвеева К.И., Кундалевич А.А., Капитунова А.И., Зозуля А.С., Сухих С.А., Цибульникова А.В., Зюбин А.Ю., Самусев И.Г., 2023

### Ключевые слова

плазмонный резонанс, нанозвезды, КРС, *Viburnum opulus L.*, кишечная палочка, антибиотики

### Благодарности

Работа выполнена при поддержке проекта государственного задания Минобрнауки Российской Федерации № FZWM-2020-0003 «Исследование новых материалов и методов плазмо- и фототерапии онкологических заболеваний, дерматитов и септических осложнений» 2020–2023 гг.

**Ссылка для цитирования:** Матвеева К.И., Кундалевич А.А., Капитунова А.И., Зозуля А.С., Сухих С.А., Цибульникова А.В., Зюбин А.Ю., Самусев И.Г. Спектроскопия комбинационного рассеяния света в исследованиях процессов инактивации бактериальных микроорганизмов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 676–684. doi: 10.17586/2226-1494-2023-23-4-676-684

## Application of Raman spectroscopy to study the inactivation process of bacterial microorganisms

Karina I. Matveeva<sup>1</sup>, Anna A. Kundalevich<sup>2</sup>, Anastasia I. Kapitonova<sup>3</sup>, Aleksandr S. Zozulya<sup>4</sup>, Stanislav A. Sukhikh<sup>5</sup>, Anna V. Tsibulnikova<sup>6</sup>, Andrey Yu. Zyubin<sup>7</sup>, Ilya G. Samusev<sup>8</sup>

<sup>1,2,3,4,5,6,7,8</sup> Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation

<sup>1</sup> matveeva.k.i@inbox.ru, <https://orcid.org/0000-0003-2585-163X>

<sup>2</sup> Kundalevich3@mail.ru, <https://orcid.org/0000-0001-7428-6441>

<sup>3</sup> AIKapitonova@mail.ru, <https://orcid.org/0000-0003-0958-6338>

<sup>4</sup> zozul97@mail.ru, <https://orcid.org/0000-0001-5717-8625>

<sup>5</sup> SSukhikh@kantiana.ru, <https://orcid.org/0000-0001-7910-8388>

<sup>6</sup> memorgold@mail.ru, <https://orcid.org/0000-0001-8578-0701>

<sup>7</sup> azubin@mail.ru, <https://orcid.org/0000-0002-9766-1408>

<sup>8</sup> ISamusev@kantiana.ru, <https://orcid.org/0000-0001-5026-7510>

### Abstract

Raman spectroscopy (RS) is one of the promising approaches for structural and functional studies of various biological objects, including bacterial microorganisms. Both traditional biochemical tests and genetic methods which require expensive reagents, consumables and are time-consuming are used for bacterial analysis. Spectroscopic methods are positioned as noninvasive, highly sensitive, and requiring minimal sample preparation. In this work we investigated the possibility of using the RS method using optical sensors based on gold anisotropic nanoparticles. The applicability of the method was demonstrated by studying the effect of a broad-spectrum cephalosporin antibiotic and an extract of *Viburnum opulus L (VO)* on *Escherichia coli (E. Coli)* colonies. The studies were performed by Raman spectroscopy using a Virsa spectrometer (Renishaw). Raman signal amplification was carried out using two original optical sensors proposed by the authors. To create sensors, we used a chemical method of depositing gold nanostars on APTES-modified quartz glasses and a physical method for creating sensors based on anodizing titanium surfaces. The results of the study showed the high sensitivity and information content of the proposed method. The possibility of using the RS method for studying the inactivation of bacterial microorganisms is shown. Spectral Raman bands of *E. Coli* were determined and identified before and after exposure to *VO* extract and antibiotic as a control. A decrease in the intensity of spectral modes corresponding to amino acids and purine metabolites was found in the average Raman spectrum of *E. Coli* after exposure to *VO* extract. For the first time, a study of the antimicrobial effect of an aqueous extract of *VO* fruits was carried out by the method of Raman scattering. It has been shown that the use of plant extracts, including *VO* fruit extracts, to inactivate the vital activity of bacterial colonies is a promising approach to the search for new alternative antibacterial agents. The results obtained are in good agreement with the already known scientific studies and confirm the effectiveness of the proposed method.

### Keywords

plasmon resonance, nanostars, SERS, *Viburnum opulus L.*, *Escherichia coli*, antibiotics

### Acknowledgements

The work was supported by the Ministry of Education and Science of the Russian Federation State Assignment no. FZWM-2020-0003 “Research of new materials and methods of plasma and phototherapy of cancer, dermatitis, and septic complications” 2020-2023.

**For citation:** Matveeva K.I., Kundalevich A.A., Kapitonova A.I., Zozulya A.S., Sukhikh S.A., Tsibulnikova A.V., Zyubin A.Yu., Samusev I.G. Application of Raman spectroscopy to study the inactivation process of bacterial microorganisms. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 676–684 (in Russian). doi: 10.17586/2226-1494-2023-23-4-676-684

### Введение

Появление в современной медицине противомикробных препаратов (антибиотики, противовирусные и противогрибковые препараты) позволило взять под кон-

троль развитие вирусных и бактериальных инфекций. С 1930 г. по 1962 г. было разработано более 20 классов антибиотиков [1], которые направлены на остановку синтеза белков и метаболитов, нарушение целостности клеточной стенки и процесса бинарного деления бак-

териальных микроорганизмов [2]. В настоящее время темпы разработок новых классов антибиотиков замедлились, однако их потребление увеличилось, что привело к появлению и быстрому распространению устойчивых к антибиотикам патогенов, особенно бактерий с множественной лекарственной устойчивостью [3]. Один из основных подходов к получению новых антимикробных агентов — поиск и адаптация натуральных продуктов, например растительных экстрактов, богатых фенолами и флавоноидами [4, 5]. По литературным данным, Всемирной организацией здравоохранения зарегистрированы названия более 20 000 видов лекарственных растений с различными потенциальными применениями, в том числе в качестве антибактериальных агентов [2]. Отметим, что известно применение экстрактов растений против бактерий, плесеней, насекомых, вирусов и опухолевых клеток в качестве фотосенсибилизаторов активной формы кислорода [6, 7].

Среди перспективных растительных антибактериальных агентов выделяются агенты растительного происхождения. Одним из перспективных природных агентов ввиду потенциальных антибактериальных свойств является *Viburnum opulus L.* Перспективность применения экстрактов на основе *Viburnum opulus L.* обусловлена наличием таких биологически активных компонентов как витамин С, каротиноиды, иридоиды, эфирные масла и фенольные соединения. Благодаря такому составу экстракт *Viburnum opulus L.* обладает антиоксидантной активностью, антимикробным потенциалом (особенно в отношении грамположительных бактерий), имеет доказанные противовоспалительные, противодействующие ожирению, антидиабетические, остеогенные, кардио- и цитопротекторные свойства [8–10].

Существует множество методов скрининга и оценки антимикробной активности, например: проточная цитофлуориметрия, биолюминесцентные и различные диффузионные методы [11]. В настоящее время активно развиваются оптические методы неинвазивного анализа чувствительности бактериальных клеток к противомикробным препаратам *in vitro*. Один из таких методов — метод комбинационного рассеяния света (КРС) и его частные случаи, как гигантское комбинационное рассеяние света (ГКРС) [12]. Данный метод — неинвазивен и позволяет получить информацию о химической структуре микробных клеток и биопленок без пробоподготовки исследуемых образцов. Метод ГКРС основан на явлении плазмонного резонанса, который можно наблюдать как на шероховатых металлических поверхностях [13], так и на металлических наночастицах [14]. Изменяя параметры наночастиц (материал, размер, геометрия) можно контролировать генерацию плазмонного резонанса и усиление сигнала КРС, что позволяет увеличить предел обнаружения исследуемых веществ. Теоретические и экспериментальные исследования показали, что максимум напряженности электрического поля формируется на углах или вершинах анизотропных наночастиц с максимальной кривизной [15].

Цель настоящей работы — уставить влияние экстракта плодов калины красной и антибиотиков на бак-

териальные колонии методом спектроскопии комбинационного рассеяния в присутствии коллоидных и планарных структур, усиливающих сигнал комбинационного рассеяния света.

### Используемые методы и подходы

**Микробиологическая пробоподготовка.** В работе использован штамм (*Dh5a*) *Escherichia coli* (*E. Coli*), предоставленный Балтийским федеральным университетом имени Иммануила Канта НОЦ «Прикладная биотехнология». В качестве среды для роста использовался лизогенный бульон (*Lysogeny broth, LB*). Штамм бактерий *E. Coli* культивирован из глицеринового стока при температуре минус 80 °С в бульоне LB и инкубирован в течение 12 ч при 250 об/мин и температуре 37 ± 0,5 °С. Полученные культуры бактерий были разбавлены бульоном LB для достижения оптической плотности 0,2 при 600 нм и помещены в чашки Петри на 12 ч для дальнейшего роста. Выращенные колонии бактерий были переданы для дальнейшего исследования НОЦ «Фундаментальная и прикладная фотоника. Нанопотоника».

Для приготовления экстрактов из плодов калины использованы заранее размороженные ягоды. Разморозка ягод была произведена в течении 30 мин при комнатной температуре (около 22 °С) с дальнейшим извлечением косточек и выжимкой сока. Для удаления мякоти полученный сок был центрифугирован в течении 10 мин при 10 000 об/мин и очищен фильтрованием надосадочный слой, в результате получен экстракт калины. Бактерии *E. Coli* (40 мкл) подвергали стрессу в течение 10 мин, смешивая их с приготовленным экстрактом калины (20 мкл). Далее полученная смесь была нанесена на анодированный титан [16] объемом 2,5 мкл для регистрации спектров КРС.

В экспериментах использовался антибиотик широкого спектра — цефтриаксон, рабочей концентрации 7,5 мг/мл.

**Подготовка оптического сенсора.** Для спектральной съемки бактерий в присутствии экстракта плодов калины использована методология, реализованная на анодированных титановых поверхностях, апробированная ранее, но без нанесенных золотых наночастиц [17]. Выбор данной поверхности для усиления сигнала КРС обусловлен наличием плазмонного резонанса, генерируемого в результате фотовозбуждений шероховатой поверхности анодированного титана. Для съемки антибиотиков была применена альтернативная, новая, менее ресурсоемкая технология, основанная на функционализации поверхности кварцевых стекол нанозвездами. Для этого применены оптически прозрачные кварцевые стекла марки КУ-1. Очистка стекол осуществлена многостадийным способом. На первом этапе пластины из стекла КУ-1 были промыты в ультразвуковой ванне в течение 30 мин, в которой через каждые 10 мин происходила замена жидкости по следующей схеме: Milli-Q — изопропиловый спирт — Milli-Q. Очищенные пластины были просушены при температуре 90 °С до полного испарения жидкости на поверхности. Высохшие стекла располагали вертикаль-

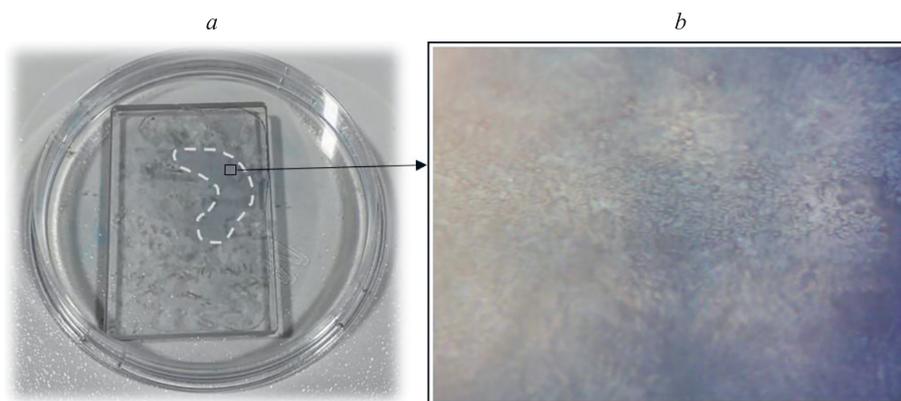


Рис. 1. Бактерии, нанесенные мазком на поверхность оптического сенсора (а), и изображение бактерий *E. Coli* на поверхности кварцевого стекла при увеличении 100 $\times$  (b)

Fig. 1. Bacteria smeared on the surface of the optical sensor (a) and the image of *E. Coli* cells on the surface of the quartz glass at magnification 100 $\times$  (b)

но в термостойкой посуде с раствором пираньи (30 %  $\text{H}_2\text{O}_2$  и 96 %  $\text{H}_2\text{SO}_4$ , 1:3) и подогреты до 70 °С. Через 30 мин пластины были промыты три раза по 10 мин в ультразвуковой ванне с Milli-Q и просушены при 90 °С до полного высыхания. Очищенные стекла располагали вертикально в 5 % растворе АРТЕС ((3-Аминопропил) триэтоксисилан) и безводного толуола на 24 ч. Для очистки от излишек АРТЕС-образцы последовательно погружали в ультразвуковую ванны с безводным этанолом (два раза по 10 мин) и с сверхчистой водой (три раза по 10 мин). На последнем этапе на АРТЕС-модифицированные поверхности адсорбировали золотые нанозвезды. Для этого пластины располагали горизонтально в коллоидных растворах и оставляли на 12 ч. По истечению данного временного промежутка стекла промывали водой и использовали по назначению. Оптические свойства таких сенсоров были подробно изучены в работе [18]. Бактерии были помещены на поверхность полученной структуры мазком и оказались различимы при стократном увеличении (рис. 1).

### Экспериментальные условия регистрации спектров

Исследуемый образец был помещен на держатель спектрометра Virsa (Renishaw). С помощью цифровой видеокамеры получено изображение образца на оптическом столе. Выполнена юстировка спектрометра и произведено позиционирование образцов. Наблюдение изображения осуществлялось на экран компьютера с помощью программного обеспечения Wire 5.4. С помощью Wire 5.4 выполнялись: настройка параметров прибора и источника лазерного излучения; управление и получение данных с детекторов; обработка данных и съемка образцов.

Спектральная съемка на анодированном титане бактерий до воздействия антибиотика и экстракта (контрольные спектры) и после воздействия экстракта была проведена при условиях: динамическая съемка в диапазоне волновых чисел 350–3200  $\text{см}^{-1}$ ; накопление сигнала 30 с; длина волны лазера  $\lambda = 785$  нм; мощность лазера варьировалась от 30 до 60 мВт и времени нако-

пления сигнала на детекторе от 30 до 60 с. Выбранные, оптимальные условия съемки были обусловлены факторами неповреждения образца и отсутствия засветки детектора.

После регистрации спектров бактерий (контрольные спектры) с помощью автоматической пипетки Eppendorf Research на бактерии была нанесена капля выбранного антибиотика объемом 1 мкл. В течение 5 мин происходило высыхание лекарственного препарата, после чего проведена спектральная съемка с антибиотиком при условиях, отличающихся от съемки на анодированном титане диапазоном мощности лазера: от 45 до 15 мВт. Измерения колебательных спектров выполнялись каждые 5 мин с целью наблюдения за динамикой изменения бактериальной стенки под воздействием антибиотика. Съемка была прекращена, когда на протяжении трех изменений спектральная картина не изменялась.

Полученные графики были сохранены в формате .txt для дальнейшей обработки в пакете программ OriginPro.

### Результаты исследования

В результате проведения экспериментальной части получены спектры КРС и ГКРС бактерий *E. Coli* до и после воздействия антибиотиков и выжимки плодов калины. Совокупность спектров, полученных при естественном высыхании бактерий, показана на рис. 2.

На рис. 2, а наблюдается постепенное увеличение интенсивности спектральных полос КРС бактерий *E. Coli* (спектры № 1–№ 3). Спектры получены при следующих условиях съемки: мощность лазера — 60 мВт; время накопления — 60 с. При регистрации спектров № 4–№ 13 экспериментальные условия были изменены (мощность лазера — 30 мВт, время накопления — 30 с) по причине резкого усиления сигнала КРС вследствие приближения (высыхания) исследуемого образца к поверхности анодированного титана. Для лучшей идентификации спектральных мод осуществлено усреднение данных по 13 спектрам (рис. 1, b). Для дальнейшего исследования антибактериального эффекта выжим-

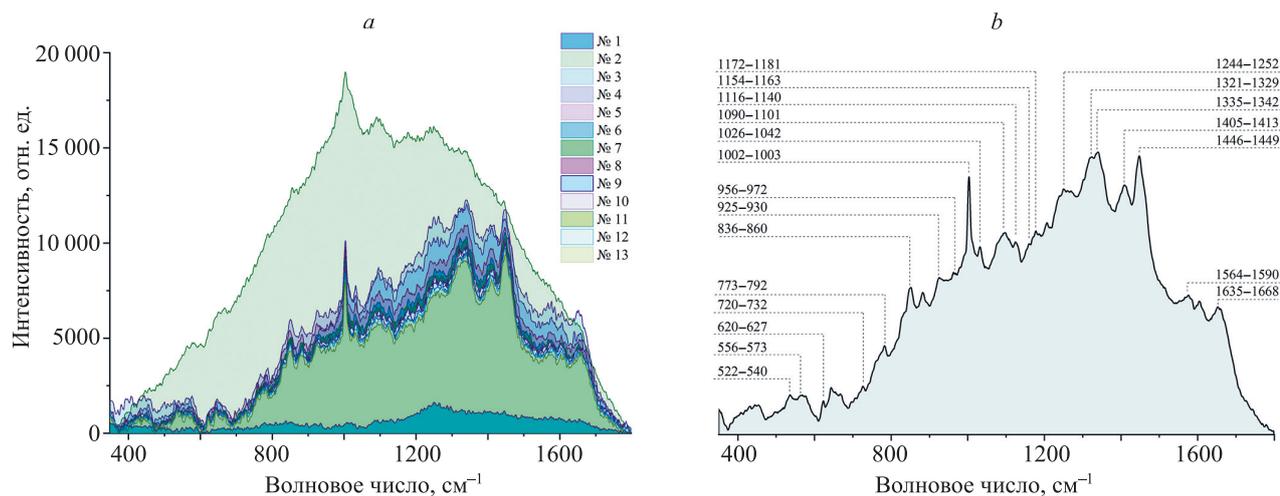


Рис. 2. Спектры комбинационного рассеяния света бактерий *E. Coli*, зарегистрированные при естественном высыхании бактерий (а), и усредненный спектр (б), являющийся их комбинацией

Fig. 2. Raman spectra of *E. Coli* recorded during natural drying of bacteria (a), and the averaged spectrum (b) which is their combination

ки плодов калины выделены спектральные полосы, в которых наблюдались аминокислоты и пуриновые производные (аденин, гуанин, гипоксантин, ксантин) [19, 20], которые представлены в таблице.

Таблица. Спектральные полосы, наблюдаемые в усредненном спектре бактерий *E. Coli* и их интерпретация

Table. Spectral bands observed in the averaged spectrum of *E. Coli* and their interpretation

Спектральные полосы, $\text{cm}^{-1}$	Интерпретация [19, 20]
522–540	гуанин
556–573	ксантин, гуанин, аденин
620–627	гипоксантин, аденин, гуанин
720–732	аденин, гипоксантин
773–792	гипоксантин
836–860	тирозин
925–930	гипоксантин
956–972	ксантин, гуанин, гипоксантин, аденин
1002–1003	фенилаланин, белок
1026–1042	гуанин, аденин, гипоксантин
1090–1101	гипоксантин
1116–1140	ксантин, гуанин, аденин
1154–1163	гипоксантин
1172–1181	гуанин, аденин
1244–1252	гипоксантин
1321–1329	гипоксантин
1335–1342	аденин
1405–1413	лейцин
1446–1449	гипоксантин, гуанин, аденин
1564–1590	гуанин, аденин, ксантин, гипоксантин
1635–1668	гуанин, аденин, ксантин, гипоксантин

Инактивация бактериальных микроорганизмов является нетривиальной задачей. В зависимости от способа (ультрафиолетовое облучение, спирты, температура, антибиотики [19, 21–23], и т. д.) и времени воздействия наблюдаются разные спектральные изменения. Изменения касаются интенсивности и положения мод, соответствующих нуклеиновым кислотам, белкам и липидам, наблюдаемым в спектрах КРС бактерий *E. Coli*. В случае исследования воздействия выжимки плодов калины на бактериальные микроорганизмы методом КРС, научных данных не представлено. Однако сравнение спектров КРС, зарегистрированных от живых и мертвых бактерий, характеризуется как общим уменьшением интенсивности спектра, так и спектральных полос.

На рис. 3 представлен спектр КРС выжимки плодов калины, усредненные спектры КРС бактерий *E. Coli* до и после воздействия экстракта. Из-за перекрытия спектральных полос характерных как для плодов калины, так и для бактерий *E. Coli*, трудно проанализировать изменения во всех модах, относящихся *E. Coli*. Однако на усредненном спектре *E. Coli* после воздействия выжимки плодов калины были определены моды ( $853 \text{ cm}^{-1}$ ,  $965 \text{ cm}^{-1}$ ,  $1003 \text{ cm}^{-1}$ ,  $1031 \text{ cm}^{-1}$ ,  $1410 \text{ cm}^{-1}$ ,  $1450 \text{ cm}^{-1}$ ,  $1658 \text{ cm}^{-1}$ ), которые не проявляются на спектре КРС выжимки плодов калины. При сравнении мод  $853 \text{ cm}^{-1}$ ,  $1003 \text{ cm}^{-1}$ ,  $1410 \text{ cm}^{-1}$ ,  $1450 \text{ cm}^{-1}$ ,  $1658 \text{ cm}^{-1}$  на спектрах *E. Coli* до и после воздействия экстракта наблюдается уменьшение интенсивности спектральных полос, относящихся к пуриновым метаболитам, аминокислотам, что предварительно указывает на инактивацию жизнедеятельности бактерий, вследствие воздействия флавоноида кверцетина.

На рис. 4 представлены экспериментальные спектры чувствительного штамма бактерий *E. Coli* (черный спектр), антибиотика (красный спектр) и спектр бактерий после взаимодействия с цефтриаксоном (синий спектр). Видно несколько характерных особенностей воздействия цефтриаксона на чувствительный штамм

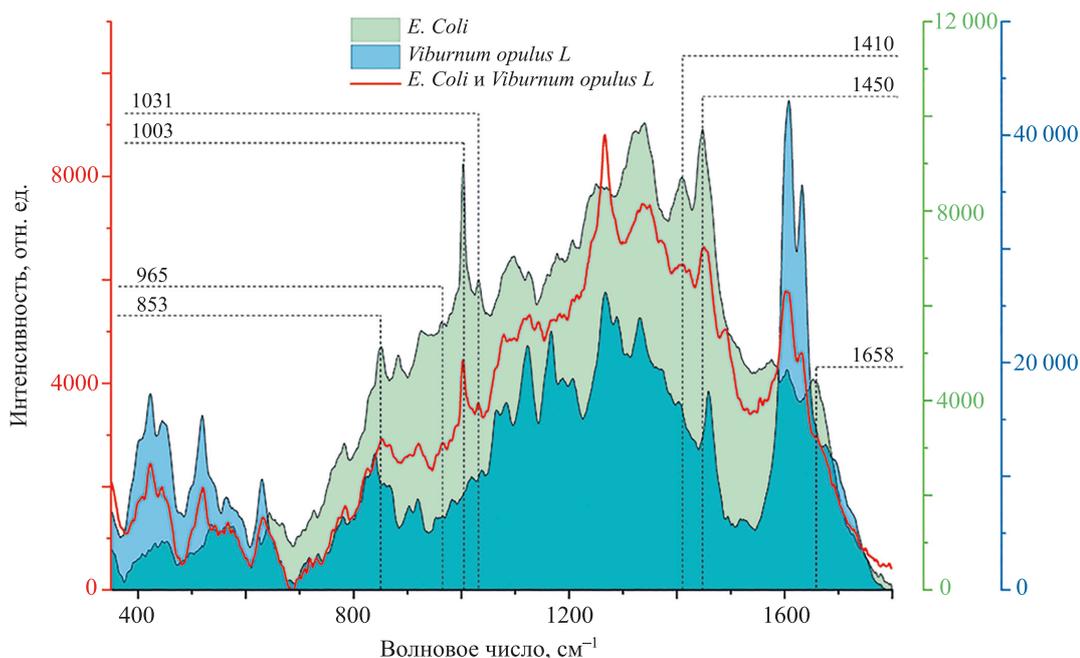


Рис. 3. Спектры комбинационного рассеяния света выжимки плодов калины и усредненные спектры *E. Coli* до и после воздействия выжимки плодов калины

Fig. 3. Raman spectrum of *Viburnum opulus L* pomace, averaged Raman spectra of *E. Coli* before and after exposure to *Viburnum opulus L* pomace

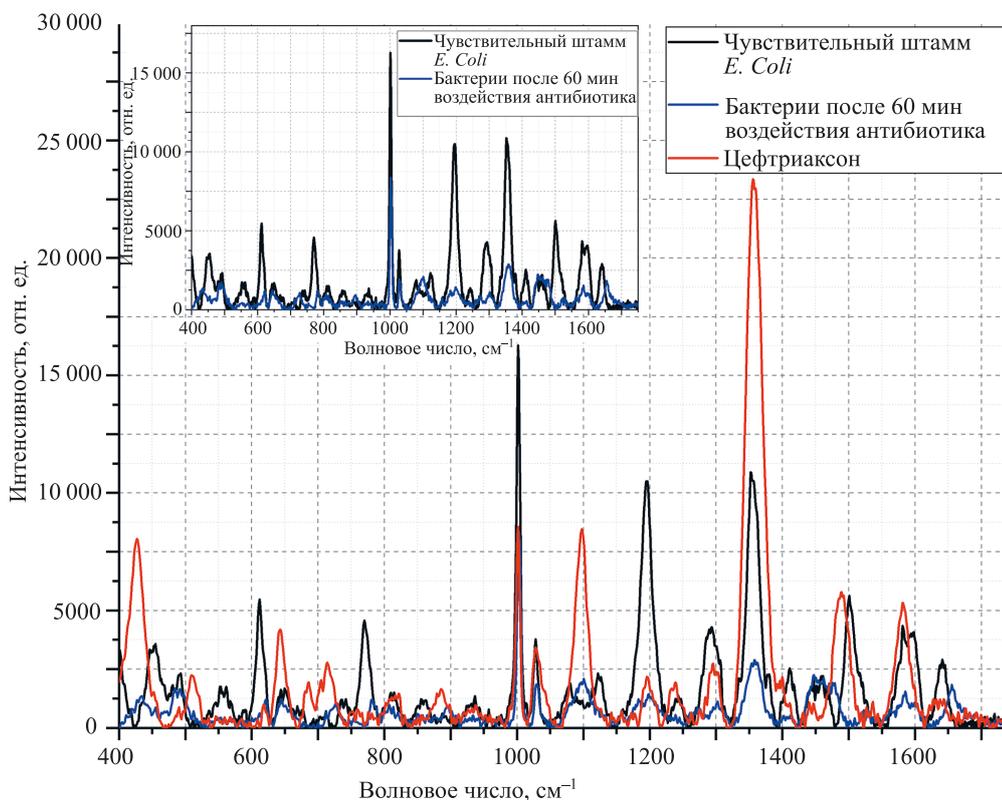


Рис. 4. Спектры гигантского комбинационного рассеяния бактерий *E. Coli* и антибиотика цефтриаксон. На вставке графика изображено сравнение спектров чувствительного штамма бактерий и штамма бактерий после 60 мин воздействия антибиотика

Fig. 4. SERS spectra of bacteria *E. Coli* and antibiotic ceftriaxone.

The inset of the graph shows a comparison of the spectra of the susceptible bacterial strain and the bacterial strain after 60 min of antibiotic exposure

*E. Coli* для органических соединений, формирующих бактериальную стенку. Для нуклеиновых кислот наблюдается сдвиг максимумов с  $735\text{ см}^{-1}$  на  $727\text{ см}^{-1}$ , с  $771\text{ см}^{-1}$  на  $785\text{ см}^{-1}$ , а также уменьшение интенсивности максимумов нуклеиновых кислот на  $672\text{ см}^{-1}$ ,  $812\text{ см}^{-1}$  и  $1575\text{ см}^{-1}$ , однако на значении  $1483\text{ см}^{-1}$  наблюдается увеличение интенсивности. Для белковых полос заметно снижение интенсивности максимумов на значениях:  $855\text{ см}^{-1}$ ,  $1003\text{ см}^{-1}$ ,  $1031\text{ см}^{-1}$  и  $1246\text{ см}^{-1}$ . Для липидной полосы определено уменьшение интенсивности на  $1064\text{ см}^{-1}$ , а также уменьшение интенсивности и сдвиг с  $1295\text{ см}^{-1}$  на  $1300\text{ см}^{-1}$ . Отдельного обсуждения заслуживает полоса фенилаланина на частоте обратного сдвига  $1003\text{ см}^{-1}$ , обозначенная как биомаркер клеточной гибели. Стоит отметить, что бактерии *E. Coli* широко используются для производства L-фенилаланина и других ароматических соединений [24, 25], являясь простым и изученным организмом с преимуществами четкого генетического фона и возможностью простых генетических манипуляций, а также быстрого роста в дешевых средах. В связи с этим

можно сделать вывод о его наличии в клетке *E. Coli* и связи с жизненным циклом клетки. В связи с этим предложено использование колебания на частоте обратного сдвига  $1003\text{ см}^{-1}$  в качестве маркера клеточной гибели.

### Заключение

Проведена идентификация спектральных мод, наблюдаемых в усредненном спектре комбинационного рассеяния колоний бактерий *Escherichia coli*. Успешно выполнена идентификация спектральных мод колоний бактерий *Escherichia coli* после воздействия *Viburnum opulus L.* Показано уменьшение интенсивности спектральных полос, относящиеся к пуриновым метаболитам, аминокислотам, что указывает на инактивацию жизнедеятельности бактерий и свидетельствует об эффективной антибактериальной активности *Viburnum opulus L.* По результатам исследования установлено, что маркером клеточной гибели может являться полоса фенилаланина на частоте обратного сдвига  $1003\text{ см}^{-1}$ .

### Литература

- Coates A.R.M., Halls G., Hu Y. Novel classes of antibiotics or more of the same? // *British Journal of Pharmacology*. 2011. V. 163. N 1. P. 184–194. <https://doi.org/10.1111/j.1476-5381.2011.01250.x>
- Cheesman M.J., Ilanko A., Blonk B. Developing new antimicrobial therapies: are synergistic combinations of plant extracts/compounds with conventional antibiotics the solution? // *Pharmacognosy Reviews*. 2017. V. 11. N 22. P. 57. [https://doi.org/10.4103/phrev.phrev\\_21\\_17](https://doi.org/10.4103/phrev.phrev_21_17)
- Aminov R.I. The role of antibiotics and antibiotic resistance in nature // *Environmental Microbiology*. 2009. V. 11. N 12. P. 2970–2988. <https://doi.org/10.1111/j.1462-2920.2009.01972.x>
- Njimoh D.L., Assob J.C.N., Mokake S.E., Nyhalah D.J., Yinda C.K., Sandjon B. Antimicrobial activities of a plethora of medicinal plant extracts and hydrolates against human pathogens and their potential to reverse antibiotic resistance // *International Journal of Microbiology*. 2015. V. 2015. P. 1–5. <https://doi.org/10.1155/2015/547156>
- Kothari V., Gupta A., Naraniwal M. Comparative study of various methods for extraction of antioxidant and antibacterial compounds from plant seeds // *Journal of Natural Remedies*. 2012. V. 12. N 2. P. 162–173. <https://doi.org/10.18311/jnr/2012/271>
- Chiode M.M.M., Colonello G.P., Kabadayan F., de Souza Silva J., Suffredini I.B., Saraceni C.H.C. Plant extract incorporated into glass ionomer cement as a photosensitizing agent for antimicrobial photodynamic therapy on *Streptococcus mutans* // *Photodiagnosis and Photodynamic Therapy*. 2022. V. 38. P. 102788. <https://doi.org/10.1016/j.pdpdt.2022.102788>
- Sarker M.A.R., Ahn Y.H. Green phytoextracts as natural photosensitizers in LED-based photodynamic disinfection of multidrug-resistant bacteria in wastewater effluent // *Chemosphere*. 2022. V. 297. P. 134157. <https://doi.org/10.1016/j.chemosphere.2022.134157>
- Ozkan G., Kostka T., Dräger G., Capanoglu E., Esatbeyoglu T. Bioaccessibility and transepithelial transportation of cranberrybush (*Viburnum opulus*) phenolics: Effects of non-thermal processing and food matrix // *Food Chemistry*. 2022. V. 380. P. 132036. <https://doi.org/10.1016/j.foodchem.2021.132036>
- Kajszczyk D., Zakłós-Szyda M., Podśędek A. *Viburnum opulus L.* A review of phytochemistry and biological effects // *Nutrients*. 2020. V. 12. N 11. P. 3398. <https://doi.org/10.3390/nu12113398>
- Perova I.B., Zhogova A.A., Cherkashina A.V., Èller K.I., Ramenskaya G.V., Samylina I.A. Biologically active substances from European guelder berry fruits // *Pharmaceutical Chemistry Journal*. 2014, vol. 48, no. 5, pp. 332–339. <https://doi.org/10.1007/s11094-014-1105-8>

### References

- Coates A.R.M., Halls G., Hu Y. Novel classes of antibiotics or more of the same? *British Journal of Pharmacology*, 2011, vol. 163, no. 1, pp. 184–194. <https://doi.org/10.1111/j.1476-5381.2011.01250.x>
- Cheesman M.J., Ilanko A., Blonk B. Developing new antimicrobial therapies: are synergistic combinations of plant extracts/compounds with conventional antibiotics the solution? *Pharmacognosy Reviews*, 2017, vol. 11, no. 22, pp. 57. [https://doi.org/10.4103/phrev.phrev\\_21\\_17](https://doi.org/10.4103/phrev.phrev_21_17)
- Aminov R.I. The role of antibiotics and antibiotic resistance in nature. *Environmental Microbiology*, 2009, vol. 11, no. 12, pp. 2970–2988. <https://doi.org/10.1111/j.1462-2920.2009.01972.x>
- Njimoh D.L., Assob J.C.N., Mokake S.E., Nyhalah D.J., Yinda C.K., Sandjon B. Antimicrobial activities of a plethora of medicinal plant extracts and hydrolates against human pathogens and their potential to reverse antibiotic resistance. *International Journal of Microbiology*, 2015, vol. 2015, pp. 1–5. <https://doi.org/10.1155/2015/547156>
- Kothari V., Gupta A., Naraniwal M. Comparative study of various methods for extraction of antioxidant and antibacterial compounds from plant seeds. *Journal of Natural Remedies*, 2012, vol. 12, no. 2, pp. 162–173. <https://doi.org/10.18311/jnr/2012/271>
- Chiode M.M.M., Colonello G.P., Kabadayan F., de Souza Silva J., Suffredini I.B., Saraceni C.H.C. Plant extract incorporated into glass ionomer cement as a photosensitizing agent for antimicrobial photodynamic therapy on *Streptococcus mutans*. *Photodiagnosis and Photodynamic Therapy*, 2022, vol. 38, pp. 102788. <https://doi.org/10.1016/j.pdpdt.2022.102788>
- Sarker M.A.R., Ahn Y.H. Green phytoextracts as natural photosensitizers in LED-based photodynamic disinfection of multidrug-resistant bacteria in wastewater effluent. *Chemosphere*, 2022, vol. 297, pp. 134157. <https://doi.org/10.1016/j.chemosphere.2022.134157>
- Ozkan G., Kostka T., Dräger G., Capanoglu E., Esatbeyoglu T. Bioaccessibility and transepithelial transportation of cranberrybush (*Viburnum opulus*) phenolics: Effects of non-thermal processing and food matrix. *Food Chemistry*, 2022, vol. 380, pp. 132036. <https://doi.org/10.1016/j.foodchem.2021.132036>
- Kajszczyk D., Zakłós-Szyda M., Podśędek A. *Viburnum opulus L.* A review of phytochemistry and biological effects. *Nutrients*, 2020, vol. 12, no. 11, pp. 3398. <https://doi.org/10.3390/nu12113398>
- Perova I.B., Zhogova A.A., Cherkashina A.V., Èller K.I., Ramenskaya G.V., Samylina I.A. Biologically active substances from European guelder berry fruits. *Pharmaceutical Chemistry Journal*, 2014, vol. 48, no. 5, pp. 332–339. <https://doi.org/10.1007/s11094-014-1105-8>

2014. V. 48. N 5. P. 332–339. <https://doi.org/10.1007/s11094-014-1105-8>
11. Balouiri M., Sadiki M., Ibsouda S.K. Methods for in vitro evaluating antimicrobial activity: A review // *Journal of Pharmaceutical Analysis*. 2016. V. 6. N 2. P. 71–79. <https://doi.org/10.1016/j.jppha.2015.11.005>
  12. Jung G.B., Nam S.W., Choi S., Lee G.-J., Park H.-K. Evaluation of antibiotic effects on *Pseudomonas aeruginosa* biofilm using Raman spectroscopy and multivariate analysis // *Biomedical Optics Express*. 2014. V. 5. N 9. P. 3238–3251. <https://doi.org/10.1364/boe.5.003238>
  13. Macias G., Alba M., Marsal L.F., Mihi A. Surface roughness boosts the SERS performance of imprinted plasmonic architectures // *Journal of Materials Chemistry C*. 2016. V. 4. N 18. P. 3970–3975. <https://doi.org/10.1039/c5tc02779a>
  14. Barbillon G. Latest novelties on plasmonic and non-plasmonic nanomaterials for SERS sensing // *Nanomaterials*. 2020. V. 10. N 6. P. 1200. <https://doi.org/10.3390/nano10061200>
  15. Khoury C.G., Vo-Dinh T. Gold nanostars for surface-enhanced Raman scattering: synthesis, characterization and optimization // *The Journal of Physical Chemistry C*. 2008. V. 112. N 48. P. 18849–18859. <https://doi.org/10.1021/jp8054747>
  16. Tcibulnikova A.V., Khankaev A.A., Samusev I.G., Slezhkin V.A., Bryukhanov V.V., Demin M.V., Myslitskaya N.A., Lyatun I.I., Medvedskaya P.N. Polarized plasmon resonance spectra of electrochemically modified titanium surfaces with gold nanoparticles // *Materials Research Express*. 2020. V. 7. N 12. P. 125802. <https://doi.org/10.1088/2053-1591/abce9e>
  17. Zyubin A., Rafalskiy V., Tcibulnikova A., Moiseeva E., Matveeva K., Tsapkova A., Lyatun I., Medvedskaya P., Samusev I., Demin M. Surface-enhanced Raman spectroscopy for antiplatelet therapy effectiveness assessment // *Laser Physics Letters*. 2020. V. 17. N 4. P. 045601. <https://doi.org/10.1088/1612-202x/ab7be5>
  18. Зюбин А.Ю., Кон И.И., Кундалевич А.А., Демискевич Е.А., Матвеева К.И., Зозуля А.С., Евтифеев Д.О., Полторабатыко Д.А., Самусев И.Г. Оптические свойства планарных плазмон-активных поверхностей, модифицированных золотыми нанозвездами // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22. № 5. С. 824–831. <https://doi.org/10.17586/2226-1494-2022-22-5-824-831>
  19. Durovich E., Evtushenko E., Senko O., Stepanov N., Efremenko E., Eremenko A., Kurochkin I. Molecular origin of surface-enhanced Raman spectra of *E. coli* suspensions excited at 532 and 785 nm using silver nanoparticle sols as sers substrates // *Bulletin of Russian State Medical University*. 2018. N 6. P. 25–32. <https://doi.org/10.24075/brsmu.2018.088>
  20. Noothalapati Venkata H.N. Raman Micro-spectroscopic and Imaging Studies of *Escherichia coli* Biofilm in situ and Intracellular Dynamics of Fission Yeasts in vivo Using Stable Isotope Labelling: PhD Thesis / National Chiao Tung University, Department of Applied Chemistry, 2013. 68 p.
  21. Teng L., Wang X., Wang X., Gou H., Ren L., Wang T., Wang Y., Ji Y., Huang W.E., Xu J. Label-free, rapid and quantitative phenotyping of stress response in *E. coli* via ramanome // *Scientific Reports*. 2016. V. 6. N 1. P. 34359. <https://doi.org/10.1038/srep34359>
  22. Moritz T.J., Polage C.R., Taylor D.S., Krol D.M., Lane S.M., Chan J.W. Evaluation of *Escherichia coli* cell response to antibiotic treatment by use of Raman spectroscopy with laser tweezers // *Journal of Clinical Microbiology*. 2010. V. 48. N 11. P. 4287–4290. <https://doi.org/10.1128/jcm.01565-10>
  23. Li R., Dhankhar D., Chen J., Krishnamoorthi A., Cesario Th.C., Rentzepis P.M. Identification of live and dead bacteria: A Raman spectroscopic study // *IEEE Access*. 2019. V. 7. P. 23549–23559. <https://doi.org/10.1109/access.2019.2899006>
  24. Gu P., Yang F., Kang J., Wang Q., Qi Q. One-step of tryptophan attenuator inactivation and promoter swapping to improve the production of L-tryptophan in *Escherichia coli* // *Microbial Cell Factories*. 2012. V. 11. N 1. P. 1–9. <https://doi.org/10.1186/1475-2859-11-30>
  25. Gu P., Su T., Wang Q., Liang Q., Qi Q. Tunable switch mediated shikimate biosynthesis in an engineered non-auxotrophic *Escherichia coli* // *Scientific Reports*. 2016. V. 6. N 1. P. 29745. <https://doi.org/10.1038/srep29745>
  11. Balouiri M., Sadiki M., Ibsouda S.K. Methods for in vitro evaluating antimicrobial activity: A review. *Journal of Pharmaceutical Analysis*, 2016, vol. 6, no. 2, pp. 71–79. <https://doi.org/10.1016/j.jppha.2015.11.005>
  12. Jung G.B., Nam S.W., Choi S., Lee G.-J., Park H.-K. Evaluation of antibiotic effects on *Pseudomonas aeruginosa* biofilm using Raman spectroscopy and multivariate analysis. *Biomedical Optics Express*, 2014, vol. 5, no. 9, pp. 3238–3251. <https://doi.org/10.1364/boe.5.003238>
  13. Macias G., Alba M., Marsal L.F., Mihi A. Surface roughness boosts the SERS performance of imprinted plasmonic architectures. *Journal of Materials Chemistry C*, 2016, vol. 4, no. 18, pp. 3970–3975. <https://doi.org/10.1039/c5tc02779a>
  14. Barbillon G. Latest novelties on plasmonic and non-plasmonic nanomaterials for SERS sensing. *Nanomaterials*, 2020, vol. 10, no. 6, pp. 1200. <https://doi.org/10.3390/nano10061200>
  15. Khoury C.G., Vo-Dinh T. Gold nanostars for surface-enhanced Raman scattering: synthesis, characterization and optimization. *The Journal of Physical Chemistry C*, 2008, vol. 112, no. 48, pp. 18849–18859. <https://doi.org/10.1021/jp8054747>
  16. Tcibulnikova A.V., Khankaev A.A., Samusev I.G., Slezhkin V.A., Bryukhanov V.V., Demin M.V., Myslitskaya N.A., Lyatun I.I., Medvedskaya P.N. Polarized plasmon resonance spectra of electrochemically modified titanium surfaces with gold nanoparticles. *Materials Research Express*, 2020, vol. 7, no. 12, pp. 125802. <https://doi.org/10.1088/2053-1591/abce9e>
  17. Zyubin A., Rafalskiy V., Tcibulnikova A., Moiseeva E., Matveeva K., Tsapkova A., Lyatun I., Medvedskaya P., Samusev I., Demin M. Surface-enhanced Raman spectroscopy for antiplatelet therapy effectiveness assessment. *Laser Physics Letters*, 2020, vol. 17, no. 4, pp. 045601. <https://doi.org/10.1088/1612-202x/ab7be5>
  18. Zyubin A.Yu., Kon I.I., Kundalevich A.A., Demishkevich E.A., Matveeva K.I., Zozulya A.S., Evtifeev D.O., Poltorabatkо D.A., Samusev I.G. Optical properties of planar plasmon active surfaces modified with gold nanostars. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 5, pp. 824–831. (in Russian). <https://doi.org/10.17586/2226-1494-2022-22-5-824-831>
  19. Durovich E., Evtushenko E., Senko O., Stepanov N., Efremenko E., Eremenko A., Kurochkin I. Molecular origin of surface-enhanced Raman spectra of *E. coli* suspensions excited at 532 and 785 nm using silver nanoparticle sols as sers substrates. *Bulletin of Russian State Medical University*, 2018, no. 6, pp. 25–32. <https://doi.org/10.24075/brsmu.2018.088>
  20. Noothalapati Venkata H.N. *Raman Micro-spectroscopic and Imaging Studies of Escherichia coli Biofilm in situ and Intracellular Dynamics of Fission Yeasts in vivo Using Stable Isotope Labelling*. PhD Thesis. National Chiao Tung University, Department of Applied Chemistry, 2013, 68 p.
  21. Teng L., Wang X., Wang X., Gou H., Ren L., Wang T., Wang Y., Ji Y., Huang W.E., Xu J. Label-free, rapid and quantitative phenotyping of stress response in *E. coli* via ramanome. *Scientific Reports*, 2016, vol. 6, no. 1, pp. 34359. <https://doi.org/10.1038/srep34359>
  22. Moritz T.J., Polage C.R., Taylor D.S., Krol D.M., Lane S.M., Chan J.W. Evaluation of *Escherichia coli* cell response to antibiotic treatment by use of Raman spectroscopy with laser tweezers. *Journal of Clinical Microbiology*, 2010, vol. 48, no. 11, pp. 4287–4290. <https://doi.org/10.1128/jcm.01565-10>
  23. Li R., Dhankhar D., Chen J., Krishnamoorthi A., Cesario Th.C., Rentzepis P.M. Identification of live and dead bacteria: A Raman spectroscopic study. *IEEE Access*, 2019, vol. 7, pp. 23549–23559. <https://doi.org/10.1109/access.2019.2899006>
  24. Gu P., Yang F., Kang J., Wang Q., Qi Q. One-step of tryptophan attenuator inactivation and promoter swapping to improve the production of L-tryptophan in *Escherichia coli*. *Microbial Cell Factories*, 2012, vol. 11, no. 1, pp. 1–9. <https://doi.org/10.1186/1475-2859-11-30>
  25. Gu P., Su T., Wang Q., Liang Q., Qi Q. Tunable switch mediated shikimate biosynthesis in an engineered non-auxotrophic *Escherichia coli*. *Scientific Reports*, 2016, vol. 6, no. 1, pp. 29745. <https://doi.org/10.1038/srep29745>

## Авторы

**Матвеева Карина Игоревна** — кандидат физико-математических наук, научный сотрудник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 57200305996](https://orcid.org/0000-0003-2585-163X), <https://orcid.org/0000-0003-2585-163X>, [matveeva.k.i@inbox.ru](mailto:matveeva.k.i@inbox.ru)

**Кундалевич Анна Анатольевна** — инженер-исследователь, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 57219604837](https://orcid.org/0000-0001-7428-6441), <https://orcid.org/0000-0001-7428-6441>, [Kundalevich3@mail.ru](mailto:Kundalevich3@mail.ru)

**Капитунова Анастасия Ивановна** — инженер-исследователь, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 57952671000](https://orcid.org/0000-0003-0958-6338), <https://orcid.org/0000-0003-0958-6338>, [AIKapitunova@mail.ru](mailto:AIKapitunova@mail.ru)

**Зозуля Александр Сергеевич** — техник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 57952671000](https://orcid.org/0000-0001-5717-8625), <https://orcid.org/0000-0001-5717-8625>, [Zozul97@mail.ru](mailto:Zozul97@mail.ru)

**Сухих Станислав Алексеевич** — доктор технических наук, доцент, заведующий лабораторией, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 57200609482](https://orcid.org/0000-0001-7910-8388), <https://orcid.org/0000-0001-7910-8388>, [SSukhikh@kantiana.ru](mailto:SSukhikh@kantiana.ru)

**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), <https://orcid.org/0000-0001-8578-0701>, [memorgold@mail.ru](mailto:memorgold@mail.ru)

**Зюбин Андрей Юрьевич** — кандидат физико-математических наук, старший научный сотрудник, Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 57193159520](https://orcid.org/0000-0002-9766-1408), <https://orcid.org/0000-0002-9766-1408>, [azubin@mail.ru](mailto:azubin@mail.ru)

**Самусев Илья Геннадьевич** — кандидат физико-математических наук, директор Научно-образовательного центра «Фундаментальная и прикладная фотоника. Нанопотоника», Балтийский федеральный университет имени Иммануила Канта, Калининград, 236041, Российская Федерация, [sc 12779220200](https://orcid.org/0000-0001-5026-7510), <https://orcid.org/0000-0001-5026-7510>, [ISamusev@kantiana.ru](mailto:ISamusev@kantiana.ru)

Статья поступила в редакцию 12.05.2023  
Одобрена после рецензирования 08.06.2023  
Принята к печати 18.07.2023

## Authors

**Karina I. Matveeva** — PhD (Physics & Mathematics), Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, [sc 57200305996](https://orcid.org/0000-0003-2585-163X), <https://orcid.org/0000-0003-2585-163X>, [matveeva.k.i@inbox.ru](mailto:matveeva.k.i@inbox.ru)

**Anna A. Kundalevich** — Research Edgineer, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, [sc 57219604837](https://orcid.org/0000-0001-7428-6441), <https://orcid.org/0000-0001-7428-6441>, [Kundalevich3@mail.ru](mailto:Kundalevich3@mail.ru)

**Anastasia I. Kapitunova** — Research Engineer, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, [sc 57952671000](https://orcid.org/0000-0003-0958-6338), <https://orcid.org/0000-0003-0958-6338>, [AIKapitunova@mail.ru](mailto:AIKapitunova@mail.ru)

**Aleksandr S. Zozulya** — Technician, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, [sc 57952671000](https://orcid.org/0000-0001-5717-8625), <https://orcid.org/0000-0001-5717-8625>, [Zozul97@mail.ru](mailto:Zozul97@mail.ru)

**Stanislav A. Sukhikh** — D.Sc., Associate Professor, Head of Laboratory, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, [sc 57200609482](https://orcid.org/0000-0001-7910-8388), <https://orcid.org/0000-0001-7910-8388>, [SSukhikh@kantiana.ru](mailto:SSukhikh@kantiana.ru)

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

**Andrey Yu. Zyubin** — PhD (Physics & Mathematics), Senior Researcher, Immanuel Kant Baltic Federal University, Kaliningrad, 236041, Russian Federation, [sc 57193159520](https://orcid.org/0000-0002-9766-1408), <https://orcid.org/0000-0002-9766-1408>, [azubin@mail.ru](mailto:azubin@mail.ru)

**Ilya G. Samusev** — PhD (Physics & Mathematics), 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), <https://orcid.org/0000-0001-5026-7510>, [ISamusev@kantiana.ru](mailto:ISamusev@kantiana.ru)

Received 12.05.2023  
Approved after reviewing 08.06.2023  
Accepted 18.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-685-695

УДК 535.3: 535.015: 535.34: 535.36: 577.342

## Численное исследование влияния концентрации метгемоглобина в крови на поглощение света в коже человека

Андрей Вячеславович Беликов<sup>1</sup>, Виктор Юрьевич Чучин<sup>2</sup>✉

<sup>1,2</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>1</sup> Первый Санкт-Петербургский государственный медицинский университет имени академика И.П. Павлова, Санкт-Петербург, 197022, Российская Федерация

<sup>2</sup> ООО «Научно-производственное предприятие волоконно-оптического и лазерного оборудования», Санкт-Петербург, 199034, Российская Федерация

<sup>1</sup> [avbelikov@gmail.com](mailto:avbelikov@gmail.com), <https://orcid.org/0000-0001-8733-4860>

<sup>2</sup> [treasure\\_planet@mail.ru](mailto:treasure_planet@mail.ru) ✉, <https://orcid.org/0009-0001-5062-1265>

### Аннотация

**Введение.** Лазеры широко применяются в дерматологии для удаления телеангиэктазий. Повышение эффективности склерозирования глуболежащих и крупных телеангиэктазий лазерным излучением, возможно за счет изменения оптического пропускания кожи при ее нагреве и преобразовании гемоглобина, содержащегося в крови кожи, в метгемоглобин. Влияние концентрации метгемоглобина в крови на поглощение света в коже человека мало изучено, что определяет актуальность данной работы в контексте поиска путей повышения эффективности лазерного удаления телеангиэктазий. **Метод.** Разработаны численные семислойные оптические модели кожи человека без и с телеангиэктазией. В диапазоне длин волн 400–1600 нм рассчитаны спектры коэффициента экстинкции и степени изменения оптического пропускания цельной крови и слоев кожи в моделях кожи без и с артериолярной и веноулярной телеангиэктазиями при различных концентрациях метгемоглобина в крови. На основании анализа спектров выбраны длины волн, на которых при преобразовании гемоглобина в метгемоглобин происходит наибольшее изменение оптического пропускания цельной крови и слоев кожи. Для выбранных длин волн методом Монте-Карло смоделирован ход лучей и получены распределения поглощенной оптической мощности в каждом слое кожи без и с телеангиэктазией при различных концентрациях метгемоглобина. **Основные результаты.** Обнаружено, что при преобразовании гемоглобина в метгемоглобин наибольшее уменьшение степени изменения оптического пропускания цельной крови происходит на длинах волн 629 нм и 1105 нм, а наибольшее увеличение — на 447 нм и 578 нм. Показано, что с ростом концентрации метгемоглобина с 0 до 100 % в коже доля поглощенной оптической мощности в слое поверхностного сплетения сосудов без и с телеангиэктазией на длинах волн 629 нм и 1105 нм повышается, а на 441 нм и 574 нм — снижается. При этом в слое глубокого сплетения сосудов на длинах волн 441 нм, 574 нм и 1105 нм доля поглощенной оптической мощности повышается, а на длине волны 629 нм сначала с ростом концентрации метгемоглобина до 25 % повышается, а затем снижается, но до значений, превышающих значение доли поглощенной оптической мощности без метгемоглобина. **Обсуждение.** Изменение оптического пропускания, связанное с заменой гемоглобина крови на метгемоглобин, в большей степени проявляется для слоя поверхностного сплетения сосудов, что связано с высоким содержанием крови в нем и ограниченным вкладом вышележащих слоев кожи в деформацию спектра падающего на этот слой света. В коже с телеангиэктазией изменение концентрации метгемоглобина изменяет долю поглощенной оптической мощности на большую величину, чем в коже без телеангиэктазии, что можно связать с увеличением объемной концентрации крови слоев кожи с телеангиэктазией и увеличением их толщины. Полученные результаты могут быть применены при разработке лазерных систем и технологий лечения заболеваний кожи, в том числе для лазерного склерозирования телеангиэктазий.

### Ключевые слова

коэффициент поглощения, коэффициент рассеяния, кожа, сосуд, телеангиэктазии, оптическая модель кожи, лазерное склерозирование

**Ссылка для цитирования:** Беликов А.В., Чучин В.Ю. Численное исследование влияния концентрации метгемоглобина в крови на поглощение света в коже человека // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 685–695. doi: 10.17586/2226-1494-2023-23-4-685-695

## Numerical study of the effect of methemoglobin concentration in the blood on the absorption of light by human skin

Andrey V. Belikov<sup>1</sup>, Viktor Yu. Chuchin<sup>2</sup>✉

<sup>1,2</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>1</sup> Pavlov University, Saint Petersburg, 197022, Russian Federation

<sup>2</sup> Limited Liability Company “NPP VOLO”, Saint Petersburg, 199034, Russian Federation

<sup>1</sup> avbelikov@gmail.com, <https://orcid.org/0000-0001-8733-4860>

<sup>2</sup> treasure\_planet@mail.ru✉, <https://orcid.org/0009-0001-5062-1265>

### Abstract

Lasers are widely used in dermatology to remove telangiectasias. Increasing the efficiency of sclerosis of deep-lying and large telangiectasias with laser radiation is possible by changing the optical transmission of the skin when it is heated and converting the hemoglobin of the blood contained in it into methemoglobin. The influence of the concentration of methemoglobin in the blood on the absorption of light in human skin is poorly understood, which determines the relevance of this study in the context of finding ways to improve the efficiency of laser removal of telangiectasias. Seven-layer optical models of human skin without telangiectasia and with it for numerical simulation were developed. The extinction coefficients and the degree of change in the optical transmission of whole blood and skin layers were calculated in the range of wavelength from 400 to 1600 nm for skin model without and with arteriolar and venular telangiectasias at various concentrations of methemoglobin in the blood. Based on the analysis of these data, the wavelengths with the biggest change in the optical transmission of whole blood and skin layers occurred during the transformation of hemoglobin to methemoglobin were selected. At the selected wavelengths, the Monte Carlo method was used in optical modelling to get the distribution of the absorbed optical power in each layer of the skin model without and with telangiectasia at various concentrations of methemoglobin. It has been shown that the spectra of extinction coefficients for arteriolar and venular telangiectasias do not differ significantly. During the transformation of hemoglobin to methemoglobin, the largest decrease in the degree of change in the optical transmission of whole blood occurs at wavelengths of 629 nm and 1105 nm, and the largest increase occurs at wavelengths of 447 nm and 578 nm. The part of absorbed optical power in the layer of superficial vascular plexus without and with telangiectasia at wavelengths of 629 nm and 1105 nm increases, and at wavelengths of 441 nm and 574 nm it decreases with a growth of the methemoglobin concentration from 0 to 100 % in the skin model. At the same time, in the layer of deep vascular plexus the value of part of absorbed optical power increases at wavelengths of 441 nm, 574 nm and 1105 nm, but at a wavelength of 629 nm first increases with a growth of the methemoglobin concentration up to 25 %, and then decreases, but to values exceeding the value of part of absorbed optical power without methemoglobin. The change in optical transmission associated with the replacement of blood hemoglobin with methemoglobin is more pronounced for the superficial vascular plexus layer, which is associated with high blood content in it and a limited contribution of the overlying skin layers to the deformation of the spectrum of light incident on this layer. In skin with telangiectasia, a change in the concentration of methemoglobin changes the proportion of absorbed optical power by a greater amount than in skin without telangiectasia, which can be associated with an increase in the volume concentration of blood in skin layers with telangiectasia and an increase in their thickness. The obtained results can be applied in the development of laser systems and technologies for the treatment of skin diseases, including laser sclerosis of telangiectasias.

### Keywords

absorption coefficient, scattering coefficient, skin, vessel, telangiectasias, skin optical model, laser sclerosing

**For citation:** Belikov A.V., Chuchin V.Yu. Numerical study of the effect of methemoglobin concentration in the blood on the absorption of light by human skin. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 685–695 (in Russian). doi: 10.17586/2226-1494-2023-23-4-685-695

### Введение

Лазерные источники широко применяются в медицинской практике [1, 2], в том числе в дерматологии для лечения телеангиэктазий, представляющих собой стойкое расширение кровеносных сосудов микроциркуляторного русла дермы (венулы, капилляры, артериолы), которое может появляться на коже из-за различных причин, включая наследственность, воздействие ультрафиолетового излучения, возрастные изменения и травмы. Хотя телеангиэктазии не представляют угрозы для здоровья, они могут вызывать эстетический дискомфорт у пациентов [2, 3]. В настоящее время для лечения телеангиэктазий проводят склеротерапию (вве-

дение в просвет сосуда препаратов, повреждающих внутреннюю оболочку сосуда, в результате чего наступает слипание его стенок) [2], электрокоагуляцию [3] или лазерное склерозирование [4–6].

Лазерное склерозирование позволяет быстро и эффективно удалять телеангиэктазии без хирургического вмешательства с минимальным временем восстановления. В связи с этим метод получил широкое распространение в дерматологической практике [4, 5]. Оптимальным лазерным источником для лечения телеангиэктазий можно считать тот, чья длина волны излучения будет склерозировать патологически измененные сосуды, эффективно поглощаясь кровью сосудов и нагревая их до температуры порядка 80 °С, но при этом

не будет вызывать некротических изменений в окружающих эти сосуды тканях [2, 6]. Известно, что кровь сосудов эффективно поглощает излучение, лежащее в видимом диапазоне электромагнитного спектра [7]. В этой связи для лазерного склерозирования наиболее часто используется излучение видимого спектра (неодимовый лазер с генерацией второй гармоники, криптоновый, на парах меди, на красителях с длиной волны  $\lambda = 500\text{--}600$  нм), а также александритовый ( $\lambda = 755$  нм), диодный ( $\lambda = 800\text{--}1000$  нм) и неодимовый ( $\lambda = 1064$  нм) лазеры [2]. Одна из проблем применения лазерного излучения видимого спектра для склерозирования телеангиэктазий — малая глубина проникновения излучения в кожу, поэтому глуболежащие телеангиэктазии не поддаются данному лечению. Для эффективного лечения крупных телеангиэктазий требуется высокая плотность энергии, что вызывает нагрев, нежелательное повреждение окружающих сосудов тканей и увеличивает сроки заживления [8]. Для повышения эффективности склерозирования глуболежащих и крупных телеангиэктазий необходимо изменить оптические свойства кожи для получения максимального ее оптического просветления в видимом диапазоне спектра и в результате добиться эффективного склерозирования патологических сосудов.

Изменение оптических свойств тканей в процессе лечения имеет большое значение для лазерной медицины. Во многих техниках выполнения процедур в лазерной хирургии, терапии и диагностике используется сжатие и растяжение биологических тканей, которые помогают доставить лазерное излучение в глубокие слои тканей [9]. Так как оптическая прозрачность биоткани зависит от коэффициентов поглощения и рассеяния, фактора анизотропии, а также показателя преломления ее компонентов, таких как мембраны, цитоплазма, ядро и другие органеллы и окружающей их жидкости [10, 11], то ее можно эффективно контролировать, изменяя разность показателей преломления компонентов биоткани и окружающего их материала [12]. Для этого можно применять различные физические или химические методы, например облучение светом ультрафиолетового диапазона, коагуляцию, воздействие низких температур, пропитку химическими веществами, гелями и маслами, в том числе обладающими гиперосмотическими свойствами и пр. [9, 12–16].

В работах [17, 18] описан способ увеличения оптического пропускания крови человека на некоторых длинах волн до 50 % путем ее нагрева до температуры порядка  $65\text{ }^{\circ}\text{C}$  [17], что может быть связано с преобразованием гемоглобина в метгемоглобин [18]. В работе [6] трансформация гемоглобина в метгемоглобин принимается как мгновенная. В [19–21] описаны такие способы нагрева кожи человека, как нанесение горячих жидкостей на поверхность тела, воздействие электрического тока высокой частоты, контакт кожи с электрическим нагревательным элементом, элементом Пельтье, а также воздействие на кожу интенсивным светом или лазерным излучением. Очевидно, что оптимизация лазерного склерозирования телеангиэктазий путем изменения оптического пропускания кожи в результате ее нагрева с последующим преобразованием

гемоглобина содержащейся в ней крови в метгемоглобин невозможна без знаний о влиянии концентрации метгемоглобина в крови на поглощение света сосудами кожи. Для теоретической оценки этого влияния прежде всего необходимо создание адекватной оптической модели кожи с телеангиэктазией.

На сегодняшний день актуальной является семислойная оптическая модель кожи, где выделены слои с наибольшим скоплением кровеносных сосудов [22–26]. К сожалению, в рамках этой модели не описан вклад процесса преобразования гемоглобина в метгемоглобин, приводящий к изменению концентрации метгемоглобина в крови, например в результате нагрева кожи, при поглощении света слоями кожи с телеангиэктазией.

Таким образом, задача настоящей работы — построение численной оптической модели кожи с телеангиэктазией, расширяющей знания о взаимодействии света с тканями кожи. А также оценка в рамках полученной модели влияния концентрации метгемоглобина в крови на поглощение света в коже, что может быть использовано при разработке новых перспективных лазерных систем и технологий лечения кожных заболеваний, в том числе телеангиэктазий.

## Материалы и методы

В результате анализа работ [27–29], посвященных строению кожи и ее оптическому моделированию, построены численные оптические модели кожи с и без телеангиэктазий. Для обоих случаев модель кожи состояла из семи слоев: роговой слой эпидермиса, слой живых клеток эпидермиса, сосочковый слой дермы, слой поверхностного сплетения сосудов, сетчатый слой дермы, слой глубокого сплетения сосудов, слой подкожно-жировой клетчатки [22–26]. Значения показателей преломления слоев кожи были взяты из следующих работ: рогового слоя — [30], слоя живых клеток эпидермиса и слоев дермы (сосочкового слоя, слоя поверхностного сплетения сосудов, сетчатого слоя и слоя глубокого сплетения сосудов) — [31], слоя подкожно-жировой клетчатки — [32]. Толщины слоев в моделях без (норма) и с телеангиэктазией приведены в табл. 1.

Телеангиэктазии локализованы преимущественно в слое поверхностного сплетения сосудов, а размер большинства телеангиэктазий составляет  $26\text{--}225$  мкм [33]. Как видно из табл. 1, при моделировании толщина слоя поверхностного сплетения сосудов в модели без телеангиэктазии была принята равной  $80$  мкм, а в модели с телеангиэктазией —  $160$  мкм. Толщина слоя глубокого сплетения сосудов в модели без телеангиэктазии равна  $200$  мкм, а с телеангиэктазией —  $400$  мкм, при этом толщина сетчатого слоя уменьшалась с  $1620$  мкм до  $1540$  мкм при моделировании кожи с телеангиэктазией в поверхностном сплетении сосудов и до  $1420$  мкм — в глубоком.

Объемные концентрации хромофоров в слоях кожи человека без телеангиэктазии приведены в табл. 2.

Известно, что объемная концентрация крови в слоях с варикозно расширенными сосудами больше на 10 % по сравнению с нормой [49]. В этой связи объемная

Таблица 1. Толщины слоев кожи с и без телеангиэктазией  
 Table 1. Thicknesses of skin layers with and without telangiectasia

Наименование слоя кожи	Толщина слоя в модели кожи без (норма) телеангиэктазии, мкм [22]	Толщина слоя в модели кожи с телеангиэктазией, мкм	
		в поверхностном сплетении сосудов	в глубоком сплетении сосудов
Роговой слой эпидермиса	20	20	20
Слой живых клеток эпидермиса	80	80	80
Сосочковый слой дермы	100	100	100
Слой поверхностного сплетения сосудов	80	160	80
Сетчатый слой дермы	1620	1540	1420
Слой глубокого сплетения сосудов	200	200	400
Слой подкожно-жировой клетчатки	6000	6000	6000

концентрация крови в модели кожи с телеангиэктазией в слое поверхностного сплетения сосудов была увеличена до 33 % (при этом объемная концентрация воды в этом слое была пропорционально уменьшена до 57 %), а в слое глубокого сплетения сосудов до 11 % (при этом объемная концентрация воды в этом слое была пропорционально уменьшена до 69 %).

Рассчитаем коэффициенты поглощения:  
 — рогового слоя эпидермиса кожи в модели [24]:

$$\mu_{aSC} = W\mu_{a_{water}} + L\mu_{a_{lipid}} + ((0,1 - 0,3 \cdot 10^{-4}\lambda) + 0,125\mu_{a_{baseline}})Bl, \quad (1)$$

где  $W$ ,  $L$  и  $Bl$  — объемные концентрации воды, липидов и форменных компонентов;  $\mu_{a_{water}}$  [50],  $\mu_{a_{lipid}}$  [7] и  $\mu_{a_{baseline}} = 0,0244 + 8,53 \exp\left(\frac{-(\lambda - 154)}{66,2}\right)$  [39] — коэффициенты поглощения воды, липидов, и форменных компонентов;

— остальных слоев кожи в модели [51]:

$$\mu_{a_{layer}} = B\mu_{a_{vessel}} + W\mu_{a_{water}} + L\mu_{a_{lipid}} + M\mu_{a_{melanin}} + C\mu_{a_{collagen}} + Br\mu_{a_{bile}} + 2,3V\epsilon_{Bc} + Bl\mu_{a_{baseline}}, \quad (2)$$

где  $B$ ,  $W$ ,  $L$ ,  $M$ ,  $C$ ,  $Br$ ,  $V\epsilon_{Bc}$  и  $Bl$  — объемные концентрации крови, воды, липидов, меланина, коллагена, билирубина, бета-каротина и форменных компонентов;  $\mu_{a_{vessel}}$ ,  $\mu_{a_{water}}$  [50],  $\mu_{a_{lipid}}$  [7],  $\mu_{a_{melanin}} = 5 \cdot 10^9 \lambda^{-3,33}$  [52],  $\mu_{a_{collagen}}$  [53, 54],  $\mu_{a_{bile}}$  [7] и  $\mu_{a_{baseline}}$  [39] — коэффициенты поглощения цельной крови в сосудах, воды, липидов, меланина, коллагена, билирубина и форменных компонентов;  $\epsilon_{Bc}$  [55] — коэффициент молярной экстинкции бета-каротина;

— цельной крови, локализованной в сосудах [51]:

$$\mu_{a_{vessel}} = C_{corr}\mu_{a_{blood}}, \quad (3)$$

Таблица 2. Объемные концентрации (%) хромофоров в слоях кожи человека без телеангиэктазии  
 Table 2. Volume concentrations (%) of chromophores in layers of human skin without telangiectasia

Наименование слоя кожи	Хромофоры							
	Кровь (B)	Вода (W)	Липиды (L)	Меланин (M)	Коллаген (C)	Билирубин (Br)	Бета-каротин (Bc)	Форменные компоненты (Bl)
Роговой слой эпидермиса	0 [24]	5,0 [24]	26,20690 [34–36]	0 [25]	0 [37]	0 [38]	0 [24]	68,7931 [39]
Слой живых клеток эпидермиса	0 [24]	20,0 [24]	2,37069 [35, 36, 40]	10,0 [25]	0 [37]	0 [38]	0 [24]	67,62931 [39]
Сосочковый слой дермы	4,0 [24]	50,0 [24]	0,83908 [35, 36, 41]	0 [25]	24,76923 [36, 42, 43]	$4 \cdot 10^{-5}$ [36, 38, 44]	$2,8 \cdot 10^{-4}$ [36, 45, 46]	20,39137 [39]
Слой поверхностного сплетения сосудов	30,0 [24]	60,0 [24]	0,83908 [35, 36, 41]	0 [25]	5,38461 [36, 42, 43]	$2,7 \cdot 10^{-4}$ [36, 38, 44]	$2,8 \cdot 10^{-4}$ [36, 45, 46]	3,77576 [39]
Сетчатый слой дермы	4,0 [24]	70,0 [24]	2,74712 [35, 36, 41]	0 [25]	14,00000 [36, 42, 43]	$4 \cdot 10^{-5}$ [36, 38, 44]	$2,8 \cdot 10^{-4}$ [36, 45, 46]	9,25256 [39]
Слой глубокого сплетения сосудов	10,0 [24]	70,0 [24]	2,74712 [35, 36, 41]	0 [25]	10,76923 [36, 42, 43]	$9 \cdot 10^{-5}$ [36, 38, 44]	$2,8 \cdot 10^{-4}$ [36, 45, 46]	6,48328 [39]
Слой подкожно-жировой клетчатки	5,0 [24]	15,0 [47]	70,0 [47]	0 [25]	5,46 [48]	$5 \cdot 10^{-5}$ [36, 38,44]	0 [45]	4,53995 [39]

где  $C_{corr} = \frac{1 - e^{-2\mu a_{blood}}}{2\mu a_{blood}r}$  [51] – коэффициент коррекции для упакованной в сосуды крови;  $\mu a_{blood} = htc(S\mu a_{HbO} + (1 - S)\mu a_{Hb}) + (1 - htc)\mu a_{water}$  [56] — поглощение цельной крови;  $\mu a_{HbO}$  [7] и  $\mu a_{Hb}$  [7] — коэффициенты поглощения оксигемоглобина (HbO) и гемоглобина (Hb);  $S$  — сатурация крови;  $htc = 0,45$  [56] — гематокрит;  $r$  [51] — средний радиус сосудов в слое.

В расчетах предположено, что кровь в слоях поверхностного и глубокого сплетения сосудов кожи с артериальной телеангиэктазией имеет сатурацию  $S = 76\%$ , с веноулярной —  $S = 69\%$  [57], а в других слоях кровь на 50 % по объему была артериальной и на 50 % – венозной и  $S = 80\%$  [23].

Рассчитаем коэффициенты рассеяния каждого слоя [23]:

$$\mu s_{layer} = BC_{corr}\mu s_{blood} + (1 - B)\mu s_0 \frac{577}{\lambda}, \quad (4)$$

где  $\mu s_{blood}$  — коэффициент рассеяния цельной крови [56];  $\mu s_0$  — коэффициент рассеяния обескровленной биоткани на длине волны 577 нм [23];  $\lambda$  — длина волны.

Фактор анизотропии каждого слоя [23] имеет вид:

$$g_{layer} = \frac{BC_{corr}\mu s_{blood}g_{blood} + (1 - B)\mu s_0g_0 \frac{577}{\lambda}}{\mu s_{layer}}, \quad (5)$$

где  $g_{blood}$  — фактор анизотропии цельной крови [56];  $g_0 = 0,7645 + 0,2355 \left(1 - \exp\left(-\frac{\lambda - 500}{729,1}\right)\right)$  — фактор анизотропии обескровленной биоткани [23].

Для цельной крови и слоев кожи, содержащих кровеносные сосуды (сосочковый слой дермы, слой поверхностного сплетения сосудов, сетчатый слой дермы и слой глубокого сплетения сосудов), по формулам (1)–(5) рассчитаны коэффициенты поглощения, рассеяния и фактор анизотропии при замене гемоглобина на метгемоглобин [58, 59] в концентрации метгемоглобина  $c_{MetHb}$  равной 25 %, 50 %, 75 % и 100 % от всего объема гемоглобина. Далее для этих случаев получены коэффициенты экстинкции [60]:

$$\varepsilon = \mu a + (1 - g)\mu s. \quad (6)$$

Степень изменения оптического пропускания (ОТ) цельной крови и слоев кожи, содержащих кровеносные сосуды, происходящего в результате преобразования гемоглобина в метгемоглобин, оценивалось по формуле:

$$OT = \frac{\varepsilon_0\%}{\varepsilon_{c_{MetHb}}}, \quad (7)$$

где  $\varepsilon_0\%$  — коэффициент экстинкции при  $c_{MetHb} = 0\%$  (при концентрации гемоглобина равной 100 %);  $\varepsilon_{c_{MetHb}}$  — коэффициент экстинкции слоя при текущей концентрации метгемоглобина в крови.

В результате расчетов по формулам (1)–(7) получены спектры коэффициентов экстинкции и степени

изменения оптического пропускания цельной крови и слоев кожи без и с телеангиэктазией (артериолярной, веноулярной) при различных концентрациях метгемоглобина в крови. На основании анализа спектров выбраны длины волн, на которых при превращении гемоглобина в метгемоглобин происходит наибольшее изменение ОТ цельной крови и слоев кожи. Для выбранных таким образом длин волн в оптической модели кожи без и с телеангиэктазией методом Монте-Карло в программе «TracePro 7.0.1» (Lambda Research Corporation, США) построен ход лучей. Световое (лазерное) воздействие осуществлено параллельным пучком диаметром 4 мм перпендикулярно поверхности кожи. При расчете использовано 10 000 лучей. В результате моделирования были получены распределения поглощенной оптической мощности в каждом слое кожи без и с телеангиэктазией при различных концентрациях метгемоглобина и построена зависимость доли поглощенной оптической мощности ( $P$ ) в каждом слое кожи от концентрации метгемоглобина в крови.

### Результаты и обсуждения

Рассчитанные спектры коэффициента экстинкции цельной крови ( $\varepsilon_{blood}$ ) при различной концентрации метгемоглобина ( $c_{MetHb}$ ) в ее составе и степени изменения оптического пропускания цельной крови  $OT_{blood}$ , полученные в результате замены гемоглобина на метгемоглобин, приведены на рис. 1. На приведенных спектрах в видимой области спектра цельной крови наблюдаются пики поглощения гемоглобина и оксигемоглобина, а в инфракрасной области спектра — пики поглощения воды. Отметим, что представленный на рис. 1 спектр коэффициента экстинкции цельной крови без метгемоглобина (при концентрации гемоглобина в ней равной 100 %) соответствует представленным спектрам цельной крови в работе [56]. Поглощение метгемоглобина имеет существенные отличия от поглощения гемоглобина в видимой области в диапазоне 415–630 нм и в инфракрасной области в диапазоне 800–1200 нм, что существенно влияет на спектр экстинкции цельной крови, в которой гемоглобин замещается метгемоглобином. При этом замещении наибольшее уменьшение коэффициента экстинкции цельной крови наблюдается на длинах волн 441 нм и 574 нм, а наибольшее увеличение — на 629 нм и 1105 нм.  $OT_{blood}$  отличается от максимального значения не более чем на 5 % в диапазонах длин волн  $441 \pm 5$  нм,  $574 \pm 5$  нм,  $629 \pm 5$  нм и  $1105 \pm 15$  нм. Наиболее близко к данным длинам волн излучение полупроводниковых InGaN/GaN ( $\lambda = 450$  нм), InGaAlP/GaAs ( $\lambda = 560$ – $635$  нм) лазеров и волоконного иттербиевого лазера ( $\lambda = 1030$ – $1120$  нм). Кроме этих лазеров, можно также обратить внимание на широко используемые в дерматологии неодимовые лазеры с генерацией второй гармоники ( $\lambda = 532$  нм), импульсные лазеры на красителях ( $\lambda = 585$ – $600$  нм) и Nd:YAG лазеры ( $\lambda = 1064$  нм), но для данных лазеров значение  $OT_{blood}$  меньше.

Степень изменения оптического пропускания цельной крови при  $c_{MetHb} = 100\%$  на длине волны 441 нм составила  $OT_{blood} = 1,96$ , на 574 нм —  $OT_{blood} = 2,50$ , на

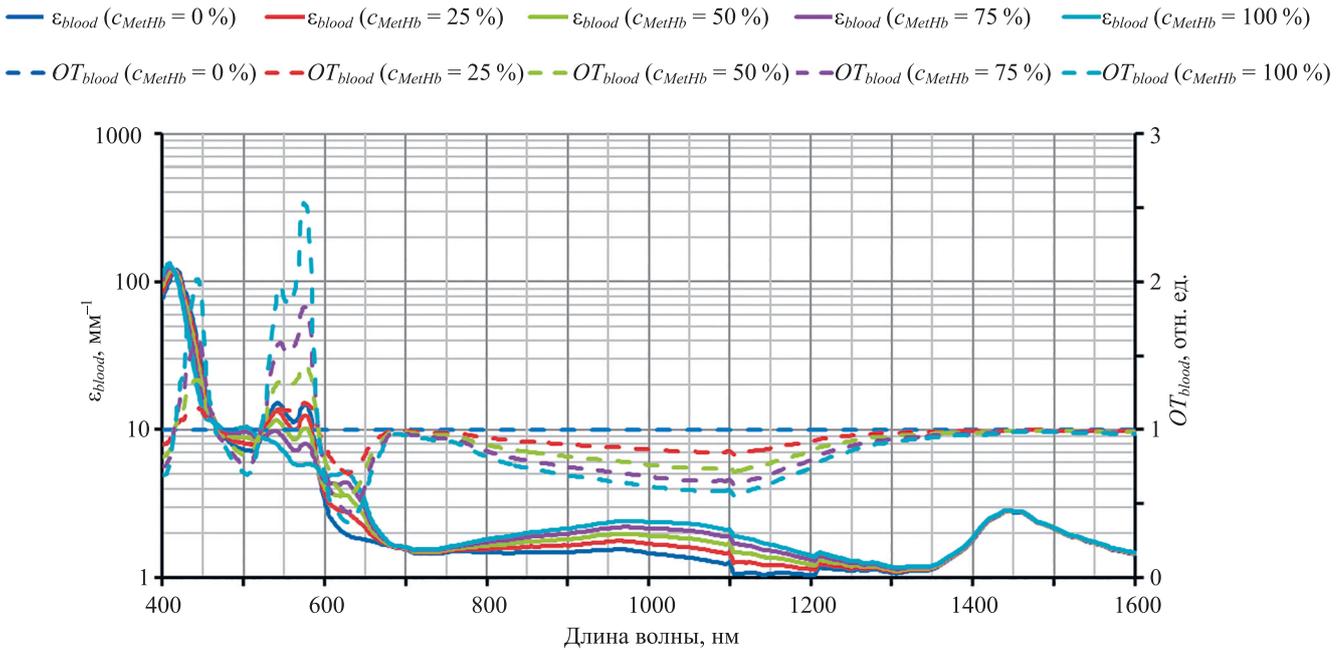


Рис. 1. Спектры коэффициента экстинкции ( $\epsilon_{blood}$ ) и степени изменения оптического пропускания ( $OT_{blood}$ ) цельной крови при различных концентрациях метгемоглобина

Fig. 1. Extinction coefficients of whole blood ( $\epsilon_{blood}$ ) and the optical clearing degree ( $OT_{blood}$ ) at various concentrations of methemoglobin in it

629 нм —  $OT_{blood} = 0,37$ , а на 1105 нм —  $OT_{blood} = 0,55$ . При этом в области пика поглощения воды на длине волны 1440 нм заметных изменений коэффициента экстинкции не произошло. При изменении концентрации метгемоглобина длины волн, соответствующие наибольшим изменениям коэффициента экстинкции, не изменились, а  $OT_{blood}$  линейно увеличился с ростом концентрации метгемоглобина.

Расчет коэффициента экстинкции и  $OT_{pl}$  в сосочковом слое дермы показал, что для этого слоя наибольшие значения  $OT_{pl}$  наблюдаются на длинах волн 440 нм, 578 нм, а наименьшие — на 1071 нм, что связано с существенным влиянием коэффициента рассеяния. В сосочковом слое степень изменения оптического пропускания при полной замене гемоглобина на метгемоглобин ( $c_{MetHb} = 100\%$ ) на длине волны 441 нм составила  $OT_{pl} = 1,10$ , на 574 нм —  $OT_{pl} = 1,11$ , на 629 нм —  $OT_{pl} = 0,94$ , а на 1105 нм —  $OT_{pl} = 0,95$ . Такое незначительное изменение  $OT_{pl}$  при замене гемоглобина на метгемоглобин связано с низкой объемной концентрацией крови в этом слое.

Расчет коэффициента экстинкции и  $OT_{svpl}$  в слое поверхностного сплетения сосудов показал, что для этого слоя наибольшие значения  $OT_{svpl}$  наблюдаются на длине волны 444 нм, что связано с вкладом коэффициента рассеяния. В слое поверхностного сплетения сосудов степень изменения оптического пропускания при полной замене гемоглобина на метгемоглобин ( $c_{MetHb} = 100\%$ ) на длине волны 441 нм составила  $OT_{svpl} = 1,43$ , на 574 нм —  $OT_{svpl} = 1,64$ , на 629 нм —  $OT_{svpl} = 0,69$ , а на 1105 нм —  $OT_{svpl} = 0,76$ .

Расчет коэффициента экстинкции и  $OT_{rl}$  в сетчатом слое дермы показал, что для этого слоя наибольшие значения  $OT_{rl}$  наблюдаются на длинах волн 440 нм,

578 нм, а наименьшие — на 1071 нм, что связано с существенным влиянием коэффициента рассеяния. В сетчатом слое степень изменения оптического пропускания при полной замене гемоглобина на метгемоглобин ( $c_{MetHb} = 100\%$ ) на длине волны 441 нм составила  $OT_{rl} = 1,10$ , на 574 нм —  $OT_{rl} = 1,11$ , на 629 нм —  $OT_{rl} = 0,94$ , а на 1105 нм —  $OT_{rl} = 0,95$ . Такое незначительное изменение  $OT_{rl}$  при замене гемоглобина на метгемоглобин связано с низкой объемной концентрацией крови в этом слое.

Расчет коэффициента экстинкции и  $OT_{dvpl}$  в слое глубокого сплетения сосудов показал, что в этом слое степень изменения оптического пропускания при полной замене гемоглобина на метгемоглобин ( $c_{MetHb} = 100\%$ ) на длине волны 441 нм составила  $OT_{dvpl} = 1,22$ , на 574 нм —  $OT_{dvpl} = 1,25$ , на 629 нм —  $OT_{dvpl} = 0,88$ , а на 1105 нм —  $OT_{dvpl} = 0,89$ .

Таким образом, наибольшее изменение оптического пропускания, связанное с заменой гемоглобина крови на метгемоглобин, наблюдается для слоя поверхностного сплетения сосудов, что связано с высоким содержанием крови в нем и ограниченным вкладом вышележащих слоев кожи в деформацию спектра падающего на этот слой света. Для слоя поверхностного сплетения сосудов были рассчитаны спектры коэффициента экстинкции с артериолярной и венулярной телеангиэктазиями. Существенного различия в этих спектрах не было установлено. Исходя из этого, далее все данные приведены для кожи с артериолярной телеангиэктазией. Для всех слоев в инфракрасной области спектра коэффициента экстинкции присутствует пик поглощения воды на длине волны 1440 нм, изменения коэффициента экстинкции, при замене гемоглобина на метгемоглобин, не происходит.

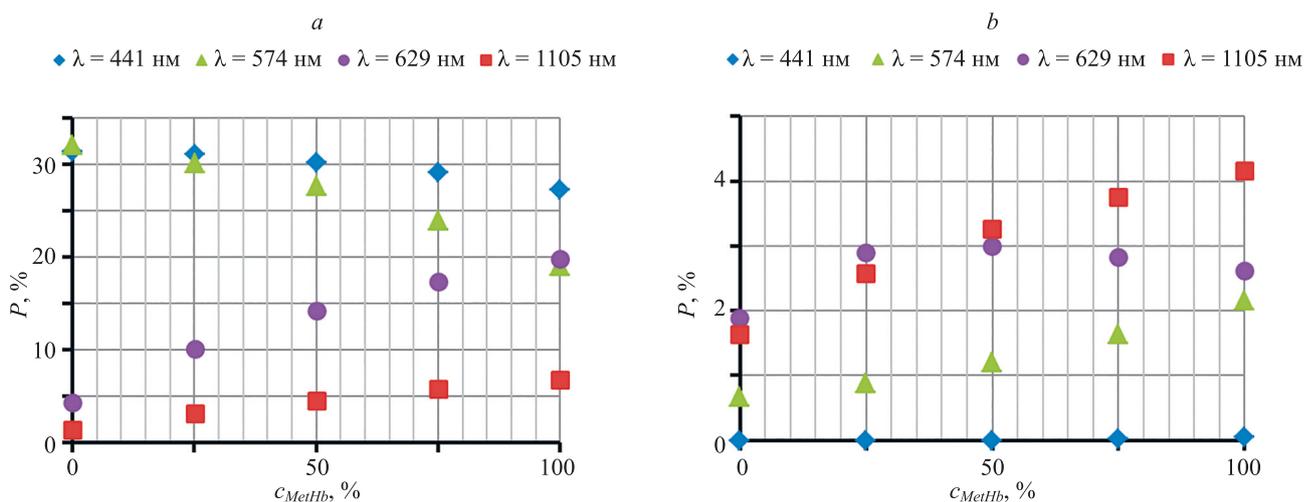


Рис. 2. Зависимость доли поглощенной оптической мощности ( $P$ ) на длинах волн 447 нм, 574 нм, 629 нм и 1105 нм в слое поверхностного сплетения сосудов (а) и в слое глубокого сплетения сосудов (б) в модели кожи без (норма) телеангиэктазии от концентрации метгемоглобина в коже

Fig. 2. Relation of the absorbed optical power ( $P$ ) in the upper blood net layer (a) and in the deep blood net layer (b) at wavelengths 447 nm, 574 nm, 629 nm and 1105 nm in the skin model without telangiectasia (normal skin) from the methemoglobin concentration in skin model

В результате оптического моделирования получены зависимости доли поглощенной оптической мощности ( $P$ ) на длинах волн 441, 574, 629 и 1105 нм от концентрации метгемоглобина в слое поверхностного (рис. 2, а) и слое глубокого (рис. 2, б) сплетения сосудов в модели кожи без (норма) телеангиэктазии.

Установлено, что в коже без телеангиэктазии на длинах волн 441 нм и 574 нм в результате замены гемоглобина метгемоглобином в коже без телеангиэктазии происходит снижение  $P$  в слое поверхностного сплетения сосудов, а в слое глубокого сплетения сосудов на этих длинах волн  $P$  увеличивается. На длине волны 1105 нм видно увеличение  $P$  в слое поверхностного и глубокого сплетений сосудов. На длине волны 629 нм

в слое поверхностного сплетения сосудов также происходит увеличение  $P$ , однако в слое глубокого сплетения сосудов при увеличении концентрации метгемоглобина в коже с 0 % до 25 % заметно увеличение  $P$ , а при  $c_{MetHb}$  выше 25 % происходит снижение  $P$ , которое можно связать с ростом поглощения в вышележащих слоях кожи.

Также были получены зависимости доли поглощенной оптической мощности ( $P$ ) на длинах волн 441, 574, 629 и 1105 нм от концентрации метгемоглобина в слое поверхностного (рис. 3, а) и глубокого (рис. 3, б) сплетений сосудов в модели кожи с артериолярной телеангиэктазией.

Видно, что в коже с артериолярной телеангиэктазией начальная экстинкция рассмотренных слоев выше,

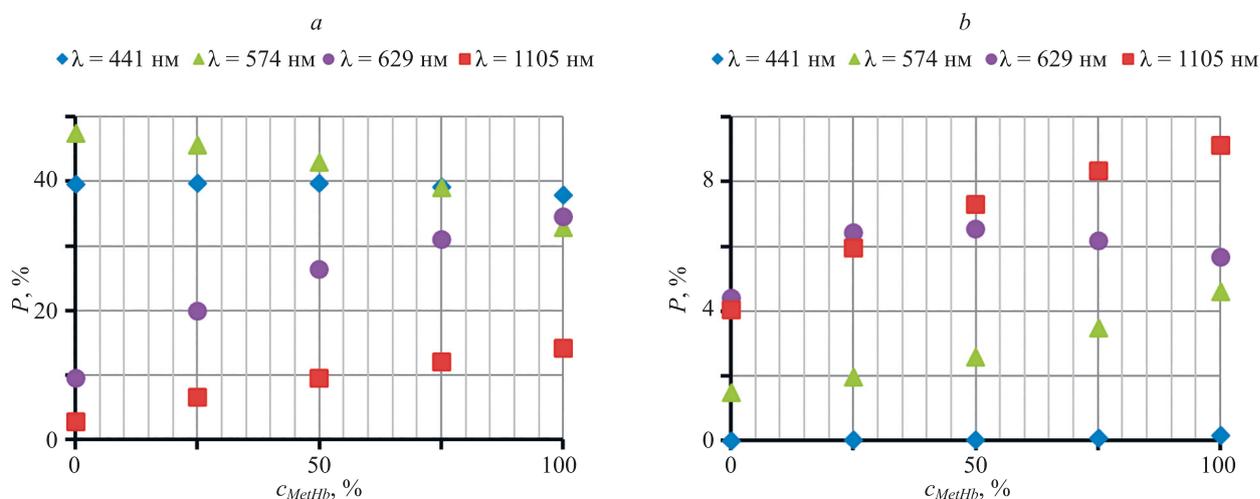


Рис. 3. Зависимость доли поглощенной оптической мощности ( $P$ ) на длинах волн 447 нм, 574 нм, 629 нм и 1105 нм в слое поверхностного (а) и глубокого (б) сплетений сосудов в модели кожи с артериолярной телеангиэктазией от концентрации метгемоглобина в коже

Fig. 3. Relation of the proportion of absorbed optical power ( $P$ ) at wavelengths of 447 nm, 574 nm, 629 nm and 1105 nm in the upper blood net layer (a) and in the deep blood net layer (b) in a skin model with arteriolar telangiectasia vs. the methemoglobin concentration in the skin model

чем в коже без телеангиэктазии, что, очевидно, связано с увеличением объемной концентрации крови в этих слоях и увеличением их толщины.

Установлено, что в слоях поверхностного и глубокого сплетения сосудов в коже с артериолярной телеангиэктазией в результате замены гемоглобина метгемоглобином  $P$  на длинах волн 441 нм, 574 нм, 629 нм и 1105 нм изменяется аналогично изменениям  $P$ , наблюдаемым для кожи без телеангиэктазии. Однако в случае кожи с телеангиэктазией изменение концентрации метгемоглобина изменяет  $P$  на большую величину, что так же, как и для начальной экстинкции, можно связать с увеличением объемной концентрации крови слоев кожи с телеангиэктазией и увеличением их толщины.

### Заключение

В работе рассмотрена возможность применения нагрева как источника метгемоглобина для управления

оптическими свойствами биоткани. Впервые численными методами исследовано влияние концентрации метгемоглобина на коэффициент экстинкции и степень изменения оптического пропускания кожи. Установлено, что метгемоглобин в составе кожи приводит к наибольшему изменению ее оптического пропускания на длинах волн вблизи 441 нм и 578 нм и максимально увеличивает коэффициент экстинкции на длинах волн вблизи 629 нм и 1105 нм. Показано, что различия между артериолярными и веноулярными телеангиэктазиями с точки зрения степени изменения оптического пропускания кожи, вызванного заменой гемоглобина на метгемоглобин, не существенны. Продемонстрировано, что замена гемоглобина на метгемоглобин приводит к изменению поглощенной оптической мощности в слоях кожи, что можно использовать для управления оптическими свойствами кожи при создании лазерных систем и технологий лечения заболеваний кожи, в том числе для лазерного склерозирования телеангиэктазий.

### Литература

1. Беликов А.В., Храмов В.Ю. Перспективы развития инновационных направлений исследований в области лазерных систем и биомедицинских оптических технологий // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2010. № 5(69). С. 110–114.
2. Goldman M.P., Weiss R.A. Sclerotherapy E-Book: Treatment of Varicose and Telangiectatic Leg Veins (Expert Consult). Elsevier Health Sciences, 2016. 455 p.
3. Ливандовский Ю.А., Павлова О.Ю. Телеангиэктазии // Клиническая дерматология и венерология. 2010. Т. 8. № 5. С. 6–15.
4. Karsai S., Roos S., Raulin C. Treatment of facial telangiectasia using a dual-wavelength laser system (595 and 1,064 nm): a randomized controlled trial with blinded response evaluation // *Dermatologic Surgery*. 2008. V. 34. N 5. P. 702–708. <https://doi.org/10.1111/j.1524-4725.2008.34131.x>
5. Sadick N., Sorhaindo L. Laser treatment of telangiectasias and reticular veins // *The Vein Book*. Academic Press, 2007. P. 157–166. <https://doi.org/10.1016/B978-012369515-4/50019-3>
6. Ross E.V., Domankevitz Y. Laser treatment of leg veins: physical mechanisms and theoretical considerations // *Lasers in Surgery and Medicine*. 2005. V. 36. N 2. P. 105–116. <https://doi.org/10.1002/lsm.20141>
7. Nachabé R., Evers D.J., Hendriks B.H.W., Lucassen G.W., Van der Voort M., Wesseling J., Ruers T.J.M. Effect of bile absorption coefficients on the estimation of liver tissue optical properties and related implications in discriminating healthy and tumorous samples // *Biomedical Optics Express*. 2011. V. 2. N 3. P. 600–614. <https://doi.org/10.1364/boe.2.000600>
8. Barton J.K., Frangineas G., Pummer H., Black J.F. Cooperative phenomena in two-pulse, two-color laser photocoagulation of cutaneous blood vessels // *Photochemistry and Photobiology*. 2001. V. 73. N 6. P. 642–650. [https://doi.org/10.1562/0031-8655\(2001\)0730642cpitpt2.0.co2](https://doi.org/10.1562/0031-8655(2001)0730642cpitpt2.0.co2)
9. Chan E.K., Sorg B., Protsenko D., O'Neil M., Motamedi M., Welch A.J. Effects of compression on soft tissue optical properties // *IEEE Journal of Selected Topics in Quantum Electronics*. 1996. V. 2. N 4. P. 943–950. <https://doi.org/10.1109/2944.577320>
10. Tuchin V.V. *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*. SPIE Press, 2015. 988 p.
11. Drezek R., Dunn A., Richards-Kortum R. Light scattering from cells: finite-difference time-domain simulations and goniometric measurements // *Applied Optics*. 1999. V. 38. N 16. P. 3651–3661. <https://doi.org/10.1364/ao.38.003651>
12. Tuchin V.V. Optical clearing of tissues and blood using the immersion method // *Journal of Physics D: Applied Physics*. 2005. V. 38. N 15. P. 2497–2518. <https://doi.org/10.1088/0022-3727/38/15/001>

### References

1. Belikov A., Khramov V. Perspective for the development of innovative research areas in the field of laser systems and biomedical optical technologies. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2010, no. 5(69), pp. 110–114. (in Russian)
2. Goldman M.P., Weiss R.A. *Sclerotherapy E-Book: Treatment of Varicose and Telangiectatic Leg Veins (Expert Consult)*. Elsevier Health Sciences, 2016, 455 p.
3. Livandovskii Yu.A., Pavlova O.Yu. Telangiectasia. *Klinicheskaya Dermatologiya i Venerologiya*, 2010, vol. 8, no. 5, pp. 6–15. (in Russian)
4. Karsai S., Roos S., Raulin C. Treatment of facial telangiectasia using a dual-wavelength laser system (595 and 1,064 nm): a randomized controlled trial with blinded response evaluation. *Dermatologic Surgery*, 2008, vol. 34, no. 5, pp. 702–708. <https://doi.org/10.1111/j.1524-4725.2008.34131.x>
5. Sadick N., Sorhaindo L. Laser treatment of telangiectasias and reticular veins. *The Vein Book*. Academic Press, 2007, pp. 157–166. <https://doi.org/10.1016/B978-012369515-4/50019-3>
6. Ross E.V., Domankevitz Y. Laser treatment of leg veins: physical mechanisms and theoretical considerations. *Lasers in Surgery and Medicine*, 2005, vol. 36, no. 2, pp. 105–116. <https://doi.org/10.1002/lsm.20141>
7. Nachabé R., Evers D.J., Hendriks B.H.W., Lucassen G.W., Van der Voort M., Wesseling J., Ruers T.J.M. Effect of bile absorption coefficients on the estimation of liver tissue optical properties and related implications in discriminating healthy and tumorous samples. *Biomedical Optics Express*, 2011, vol. 2, no. 3, pp. 600–614. <https://doi.org/10.1364/boe.2.000600>
8. Barton J.K., Frangineas G., Pummer H., Black J.F. Cooperative phenomena in two-pulse, two-color laser photocoagulation of cutaneous blood vessels. *Photochemistry and Photobiology*, 2001, vol. 73, no. 6, pp. 642–650. [https://doi.org/10.1562/0031-8655\(2001\)0730642cpitpt2.0.co2](https://doi.org/10.1562/0031-8655(2001)0730642cpitpt2.0.co2)
9. Chan E.K., Sorg B., Protsenko D., O'Neil M., Motamedi M., Welch A.J. Effects of compression on soft tissue optical properties. *IEEE Journal of Selected Topics in Quantum Electronics*, 1996, vol. 2, no. 4, pp. 943–950. <https://doi.org/10.1109/2944.577320>
10. Tuchin V.V. *Tissue Optics: Light Scattering Methods and Instruments for Medical Diagnosis*. SPIE Press, 2015, 988 p.
11. Drezek R., Dunn A., Richards-Kortum R. Light scattering from cells: finite-difference time-domain simulations and goniometric measurements. *Applied Optics*, 1999, vol. 38, no. 16, pp. 3651–3661. <https://doi.org/10.1364/ao.38.003651>
12. Tuchin V.V. Optical clearing of tissues and blood using the immersion method. *Journal of Physics D: Applied Physics*, 2005, vol. 38, no. 15, pp. 2497–2518. <https://doi.org/10.1088/0022-3727/38/15/001>

13. Sterenborg H.J., Van der Leun J.C. Change in epidermal transmission due to UV-induced hyperplasia in hairless mice: a first approximation of the action spectrum // *Photo-dermatology*. 1988. V. 5. N 2. P. 71–82.
14. Yaroslavsky A.N., Schulze P.C., Yaroslavsky I.V., Schober R., Ulrich F., Schwarzmaier H.-J. Optical properties of selected native and coagulated human brain tissues in vitro in the visible and near infrared spectral range // *Physics in Medicine & Biology*. 2002. V. 47. N 12. P. 2059–2073. <https://doi.org/10.1088/0031-9155/47/12/305>
15. Maier J.S., Walker S.A., Fantini S., Franceschini M.A., Gratton E. Possible correlation between blood glucose concentration and the reduced scattering coefficient of tissues in the near infrared // *Optics Letters*. 1994. V. 19. N 24. P. 2062–2064. <https://doi.org/10.1364/ol.19.002062>
16. Wen X., Mao Z., Han Z., Tuchin V.V., Zhu D. In vivo skin optical clearing by glycerol solutions: mechanism // *Journal of Biophotonics*. 2010. V. 3. N 1-2. P. 44–52. <https://doi.org/10.1002/jbio.200910080>
17. Baranov V.Y., Chekhov D.I., Leonov A.G., Leonov P.G., Ryaboshapka O.M., Semenov S.Y., Tatsis G.P. Heat-induced changes in optical properties of human whole blood in vitro // *Proceedings of SPIE*. 1999. V. 3599. P. 180–187. <https://doi.org/10.1117/12.348378>
18. Seto Y., Kataoka M., Tsuge K. Stability of blood carbon monoxide and hemoglobins during heating // *Forensic Science International*. 2001. V. 121. N 1-2. P. 144–150. [https://doi.org/10.1016/s0379-0738\(01\)00465-0](https://doi.org/10.1016/s0379-0738(01)00465-0)
19. Wilcox G.L., Giesler Jr G.J. An instrument using a multiple layer Peltier device to change skin temperature rapidly // *Brain Research Bulletin*. 1984. V. 12. N 1. P. 143–146. [https://doi.org/10.1016/0361-9230\(84\)90227-2](https://doi.org/10.1016/0361-9230(84)90227-2)
20. Бурлева Е.П., Эктова М.В., Беленцов С.М., Чукин С.А., Макаров С.Е., Веселов Б.А. Лечение телеангиэктазий нижних конечностей методом термокоагуляции с использованием аппарата ТС-3000 // *Стационарорзамещающие технологии: Амбулаторная хирургия*. 2018. № 1-2. С. 72–79. <https://doi.org/10.21518/1995-14772018-1-2-72-79>
21. Trelles M.A., Weiss R., Moreno-Moragas J., Romero C., Vélez M., Álvarez X. Treatment of leg veins with combined pulsed dye and Nd:YAG lasers: 60 patients assessed at 6 months // *Lasers in Surgery and Medicine*. 2010. V. 42. N 9. P. 769–774. <https://doi.org/10.1002/lsm.20972>
22. Dremín V., Zhrebtsov E., Bykov A., Popov A., Doronin A., Meglinski I. Influence of blood pulsation on diagnostic volume in pulse oximetry and photoplethysmography measurements // *Applied Optics*. 2019. V. 58. N 34. P. 9398–9405. <https://doi.org/10.1364/ao.58.009398>
23. Bashkatov A.N., Genina E.A., Tuchin V.V., Altshuler G.B., Yaroslavsky I.V. Monte Carlo study of skin optical clearing to enhance light penetration in the tissue: implications for photodynamic therapy of acne vulgaris // *Proceedings of SPIE*. 2008. V. 7022. P. 80–91. <https://doi.org/10.1117/12.803909>
24. Meglinski I.V., Matcher S.J. Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions // *Physiological Measurement*. 2002. V. 23. N 4. P. 741–753. <https://doi.org/10.1088/0967-3334/23/4/312>
25. Kim O., McMurdy J., Lines C., Duffy S., Crawford G., Alber M. Reflectance spectrometry of normal and bruised human skins: experiments and modeling // *Physiological Measurement*. 2012. V. 33. N 2. P. 159–175. <https://doi.org/10.1088/0967-3334/33/2/159>
26. Meglinski I.V., Matcher S.J. Computer simulation of the skin reflectance spectra // *Computer Methods and Programs in Biomedicine*. 2003. V. 70. N 2. P. 179–186. [https://doi.org/10.1016/s0169-2607\(02\)00099-8](https://doi.org/10.1016/s0169-2607(02)00099-8)
27. Walters K.A., Roberts M.S. The structure and function of skin // *Dermatological and Transdermal Formulations*. CRC press, 2002. P. 19–58.
28. Shirshin E.A., Gurfinkel Y.I., Priezzhev A.V., Fadeev V.V., Lademann J., Darvin M.E. Two-photon autofluorescence lifetime imaging of human skin papillary dermis in vivo: assessment of blood capillaries and structural proteins localization // *Scientific Reports*. 2017. V. 7. N 1. P. 1171. <https://doi.org/10.1038/s41598-017-01238-w>
29. Wang S., Yu X.-Y., Fan W., Li C.-X., Fei W.-M., Li S., Zhou J., Hu R., Lui M., Xu F., Xu J., Cui Y. Detection of skin thickness and density in healthy Chinese people by using high-frequency ultrasound // *Skin Research and Technology*. 2023. V. 29. N 1. P. e13219. <https://doi.org/10.1111/srt.13219>
13. Sterenborg H.J., Van der Leun J.C. Change in epidermal transmission due to UV-induced hyperplasia in hairless mice: a first approximation of the action spectrum. *Photo-dermatology*, 1988, vol. 5, no. 2, pp. 71–82.
14. Yaroslavsky A.N., Schulze P.C., Yaroslavsky I.V., Schober R., Ulrich F., Schwarzmaier H.-J. Optical properties of selected native and coagulated human brain tissues in vitro in the visible and near infrared spectral range. *Physics in Medicine & Biology*, 2002, vol. 47, no. 12, pp. 2059–2073. <https://doi.org/10.1088/0031-9155/47/12/305>
15. Maier J.S., Walker S.A., Fantini S., Franceschini M.A., Gratton E. Possible correlation between blood glucose concentration and the reduced scattering coefficient of tissues in the near infrared. *Optics Letters*, 1994, vol. 19, no. 24, pp. 2062–2064. <https://doi.org/10.1364/ol.19.002062>
16. Wen X., Mao Z., Han Z., Tuchin V.V., Zhu D. In vivo skin optical clearing by glycerol solutions: mechanism. *Journal of Biophotonics*, 2010, vol. 3, no. 1-2, pp. 44–52. <https://doi.org/10.1002/jbio.200910080>
17. Baranov V.Y., Chekhov D.I., Leonov A.G., Leonov P.G., Ryaboshapka O.M., Semenov S.Y., Tatsis G.P. Heat-induced changes in optical properties of human whole blood in vitro. *Proceedings of SPIE*, 1999, vol. 3599, pp. 180–187. <https://doi.org/10.1117/12.348378>
18. Seto Y., Kataoka M., Tsuge K. Stability of blood carbon monoxide and hemoglobins during heating. *Forensic Science International*, 2001, vol. 121, no. 1-2, pp. 144–150. [https://doi.org/10.1016/s0379-0738\(01\)00465-0](https://doi.org/10.1016/s0379-0738(01)00465-0)
19. Wilcox G.L., Giesler Jr G.J. An instrument using a multiple layer Peltier device to change skin temperature rapidly. *Brain Research Bulletin*, 1984, vol. 12, no. 1, pp. 143–146. [https://doi.org/10.1016/0361-9230\(84\)90227-2](https://doi.org/10.1016/0361-9230(84)90227-2)
20. Burleva E.P., Ektova M.V., Belentsov S.M., Chukin S.A., Makarov S.E., Veselov B.A. Treatment of lower extremity telangiectasias by thermocoagulation method using TS-3000 apparatus. *Ambulatory Surgery (Russia)*, 2018, no. 1-2, pp. 72–79. (in Russian). <https://doi.org/10.21518/1995-14772018-1-2-72-79>
21. Trelles M.A., Weiss R., Moreno-Moragas J., Romero C., Vélez M., Álvarez X. Treatment of leg veins with combined pulsed dye and Nd:YAG lasers: 60 patients assessed at 6 months. *Lasers in Surgery and Medicine*, 2010, vol. 42, no. 9, pp. 769–774. <https://doi.org/10.1002/lsm.20972>
22. Dremín V., Zhrebtsov E., Bykov A., Popov A., Doronin A., Meglinski I. Influence of blood pulsation on diagnostic volume in pulse oximetry and photoplethysmography measurements. *Applied Optics*, 2019, vol. 58, no. 34, pp. 9398–9405. <https://doi.org/10.1364/ao.58.009398>
23. Bashkatov A.N., Genina E.A., Tuchin V.V., Altshuler G.B., Yaroslavsky I.V. Monte Carlo study of skin optical clearing to enhance light penetration in the tissue: implications for photodynamic therapy of acne vulgaris. *Proceedings of SPIE*, 2008, vol. 7022, pp. 80–91. <https://doi.org/10.1117/12.803909>
24. Meglinski I.V., Matcher S.J. Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions. *Physiological Measurement*, 2002, vol. 23, no. 4, pp. 741–753. <https://doi.org/10.1088/0967-3334/23/4/312>
25. Kim O., McMurdy J., Lines C., Duffy S., Crawford G., Alber M. Reflectance spectrometry of normal and bruised human skins: experiments and modeling. *Physiological Measurement*, 2012, vol. 33, no. 2, pp. 159–175. <https://doi.org/10.1088/0967-3334/33/2/159>
26. Meglinski I.V., Matcher S.J. Computer simulation of the skin reflectance spectra. *Computer Methods and Programs in Biomedicine*, 2003, vol. 70, no. 2, pp. 179–186. [https://doi.org/10.1016/s0169-2607\(02\)00099-8](https://doi.org/10.1016/s0169-2607(02)00099-8)
27. Walters K.A., Roberts M.S. The structure and function of skin. *Dermatological and Transdermal Formulations*. CRC press, 2002, pp. 19–58.
28. Shirshin E.A., Gurfinkel Y.I., Priezzhev A.V., Fadeev V.V., Lademann J., Darvin M.E. Two-photon autofluorescence lifetime imaging of human skin papillary dermis in vivo: assessment of blood capillaries and structural proteins localization. *Scientific Reports*, 2017, vol. 7, no. 1, pp. 1171. <https://doi.org/10.1038/s41598-017-01238-w>
29. Wang S., Yu X.-Y., Fan W., Li C.-X., Fei W.-M., Li S., Zhou J., Hu R., Lui M., Xu F., Xu J., Cui Y. Detection of skin thickness and density in healthy Chinese people by using high-frequency ultrasound // *Skin Research and Technology*. 2023. V. 29. N 1. P. e13219. <https://doi.org/10.1111/srt.13219>

30. Bashkatov A.N., Genina E.A., Tuchin V.V. Optical properties of skin, subcutaneous, and muscle tissues: a review // *Journal of Innovative Optical Health Sciences*. 2011. V. 4. N 1. P. 9–38. <https://doi.org/10.1142/S1793545811001319>
31. Ding H., Lu J.Q., Wooden W.A., Kragel P.J., Hu X.-H. Refractive indices of human skin tissues at eight wavelengths and estimated dispersion relations between 300 and 1600 nm // *Physics in Medicine & Biology*. 2006. V. 51. N 6. P. 1479–1489. <https://doi.org/10.1088/0031-9155/51/6/008>
32. Müller G.J., Roggan A. *Laser-Induced Interstitial Thermotherapy*. SPIE Press, 1995. 549 p.
33. Wokalek H., Vanscheidt W., Martay K., Leder O. Morphology and localization of sunburst varicosities: an electron microscopic and morphometric study // *The Journal of Dermatologic Surgery and Oncology*. 1989. V. 15. N 2. P. 149–154. <https://doi.org/10.1111/j.1524-4725.1989.tb03021.x>
34. Lee S.H., Jeong S.K., Ahn S.K. An update of the defensive barrier function of skin // *Yonsei Medical Journal*. 2006. V. 47. N 3. P. 293–306. <https://doi.org/10.3349/ymj.2006.47.3.293>
35. Groen D., Gooris G.S., Barlow D.J., Lawrence M.J., Van Mechelen J.B., Demé B., Bouwstra J.A. Disposition of ceramide in model lipid membranes determined by neutron diffraction // *Biophysical Journal*. 2011. V. 100. N 6. P. 1481–1489. <https://doi.org/10.1016/j.bpj.2011.02.001>
36. Сетейкин А.Ю. Модель расчета температурных полей, возникающих при воздействии лазерного излучения на многослойную биоткань // *Оптический журнал*. 2005. Т. 72. № 7. С. 42–47.
37. Labib R.S., Anhalt G.J., Patel H.P., Diaz L.A. Epidermal proteins. I. Differential extraction and quantitative polyacrylamide gel-electrophoretic analysis of basal spinous-cell proteins of neonatal mouse epidermis // *Archives of Dermatological Research*. 1985. V. 277. N 4. P. 253–263. <https://doi.org/10.1007/BF00509077>
38. Chaudhary P., Kumar A., Saxena N., Biswal U.C. Hyperbilirubinemia as a predictor of gangrenous/perforated appendicitis: a prospective study // *Annals of Gastroenterology: Quarterly Publication of the Hellenic Society of Gastroenterology*. 2013. V. 26. N 4. P. 325–331.
39. Malskies C.R., Eibenberger E., Angelopoulou E. The Recognition of ethnic groups based on histological skin properties // *Proc. of the Vision, Modeling, and Visualization Workshop*. Berlin, Germany, 2011. P. 353–360. <https://doi.org/10.2312/PE/VMV/VMV11/353-360>
40. Nicolaides N. Skin lipids. II. Lipid class composition of samples from various species and anatomical sites // *Journal of the American Oil Chemists' Society*. 1965. V. 42. N 8. P. 691–702. <https://doi.org/10.1007/BF02540042>
41. McMaster J.D., Jenkinson D.M., Noble R.C., Elder H.Y. The lipid composition of bovine sebum and dermis // *British Veterinary Journal*. 1985. V. 141. N 1. P. 34–41. [https://doi.org/10.1016/0007-1935\(85\)90124-1](https://doi.org/10.1016/0007-1935(85)90124-1)
42. Waller J.M., Maibach H.I. Age and skin structure and function, a quantitative approach (II): protein, glycosaminoglycan, water, and lipid content and structure // *Skin Research and Technology*. 2006. V. 12. N 3. P. 145–154. <https://doi.org/10.1111/j.0909-752X.2006.00146.x>
43. Yannas I.V., Burke J.F., Gordon P.L., Huang C., Rubenstein R.H. Design of an artificial skin. II. Control of chemical composition // *Journal of Biomedical Materials Research*. 1980. V. 14. N 2. P. 107–132. <https://doi.org/10.1002/jbm.820140203>
44. Bonnett R., Davies J.E., Hursthouse M.B. Structure of bilirubin // *Nature*. 1976. V. 262. N 5566. P. 326–328. <https://doi.org/10.1038/262326a0>
45. Ermakov I.V., Gellermann W. Dermal carotenoid measurements via pressure mediated reflection spectroscopy // *Journal of Biophotonics*. 2012. V. 5. N 7. P. 559–570. <https://doi.org/10.1002/jbio.201100122>
46. Haynes W.M. *CRC Handbook of Chemistry and Physics*. CRC press, 2016. 2670 p.
47. Thomas L.W. The chemical composition of adipose tissue of man and mice // *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*. 1962. V. 47. N 2. P. 179–188. <https://doi.org/10.1113/expphysiol.1962.sp001589>
48. Mujkić R., Mujkić D.Š., Ilić I., Rodak E., Šumanovac A., Grgić A., Divković D., Selthofer-Relatić K. Early childhood fat tissue changes—adipocyte morphometry, collagen deposition, and expression of CD163+ cells in subcutaneous and visceral adipose tissue of male children // *International Journal of Environmental Research and Public Health*. 2021. V. 18. N 7. P. 3627. <https://doi.org/10.3390/ijerph18073627>
- in healthy Chinese people by using high-frequency ultrasound. *Skin Research and Technology*, 2023, vol. 29, no. 1, pp. e13219. <https://doi.org/10.1111/srt.13219>
30. Bashkatov A.N., Genina E.A., Tuchin V.V. Optical properties of skin, subcutaneous, and muscle tissues: a review. *Journal of Innovative Optical Health Sciences*, 2011, vol. 4, no. 1, pp. 9–38. <https://doi.org/10.1142/S1793545811001319>
31. Ding H., Lu J.Q., Wooden W.A., Kragel P.J., Hu X.-H. Refractive indices of human skin tissues at eight wavelengths and estimated dispersion relations between 300 and 1600 nm. *Physics in Medicine & Biology*, 2006, vol. 51, no. 6, pp. 1479–1489. <https://doi.org/10.1088/0031-9155/51/6/008>
32. Müller G.J., Roggan A. *Laser-Induced Interstitial Thermotherapy*. SPIE Press, 1995, 549 p.
33. Wokalek H., Vanscheidt W., Martay K., Leder O. Morphology and localization of sunburst varicosities: an electron microscopic and morphometric study. *The Journal of Dermatologic Surgery and Oncology*, 1989, vol. 15, no. 2, pp. 149–154. <https://doi.org/10.1111/j.1524-4725.1989.tb03021.x>
34. Lee S.H., Jeong S.K., Ahn S.K. An update of the defensive barrier function of skin. *Yonsei Medical Journal*, 2006, vol. 47, no. 3, pp. 293–306. <https://doi.org/10.3349/ymj.2006.47.3.293>
35. Groen D., Gooris G.S., Barlow D.J., Lawrence M.J., Van Mechelen J.B., Demé B., Bouwstra J.A. Disposition of ceramide in model lipid membranes determined by neutron diffraction. *Biophysical Journal*, 2011, vol. 100, no. 6, pp. 1481–1489. <https://doi.org/10.1016/j.bpj.2011.02.001>
36. Seteikin A.Yu. Model for calculating the temperature fields that appear when laser radiation acts on multilayer biological tissue. *Journal of Optical Technology*, 2005, vol. 72, no. 7, pp. 535–539. <https://doi.org/10.1364/JOT.72.000535>
37. Labib R.S., Anhalt G.J., Patel H.P., Diaz L.A. Epidermal proteins. I. Differential extraction and quantitative polyacrylamide gel-electrophoretic analysis of basal spinous-cell proteins of neonatal mouse epidermis. *Archives of Dermatological Research*, 1985, vol. 277, no. 4, pp. 253–263. <https://doi.org/10.1007/BF00509077>
38. Chaudhary P., Kumar A., Saxena N., Biswal U.C. Hyperbilirubinemia as a predictor of gangrenous/perforated appendicitis: a prospective study. *Annals of Gastroenterology: Quarterly Publication of the Hellenic Society of Gastroenterology*, 2013, vol. 26, no. 4, pp. 325–331.
39. Malskies C.R., Eibenberger E., Angelopoulou E. The Recognition of ethnic groups based on histological skin properties. *Proc. of the Vision, Modeling, and Visualization Workshop*. Berlin, Germany, 2011, pp. 353–360. <https://doi.org/10.2312/PE/VMV/VMV11/353-360>
40. Nicolaides N. Skin lipids. II. Lipid class composition of samples from various species and anatomical sites. *Journal of the American Oil Chemists' Society*, 1965, vol. 42, no. 8, pp. 691–702. <https://doi.org/10.1007/BF02540042>
41. McMaster J.D., Jenkinson D.M., Noble R.C., Elder H.Y. The lipid composition of bovine sebum and dermis. *British Veterinary Journal*, 1985, vol. 141, no. 1, pp. 34–41. [https://doi.org/10.1016/0007-1935\(85\)90124-1](https://doi.org/10.1016/0007-1935(85)90124-1)
42. Waller J.M., Maibach H.I. Age and skin structure and function, a quantitative approach (II): protein, glycosaminoglycan, water, and lipid content and structure. *Skin Research and Technology*, 2006, vol. 12, no. 3, pp. 145–154. <https://doi.org/10.1111/j.0909-752X.2006.00146.x>
43. Yannas I.V., Burke J.F., Gordon P.L., Huang C., Rubenstein R.H. Design of an artificial skin. II. Control of chemical composition. *Journal of Biomedical Materials Research*, 1980, vol. 14, no. 2, pp. 107–132. <https://doi.org/10.1002/jbm.820140203>
44. Bonnett R., Davies J.E., Hursthouse M.B. Structure of bilirubin. *Nature*, 1976, vol. 262, no. 5566, pp. 326–328. <https://doi.org/10.1038/262326a0>
45. Ermakov I.V., Gellermann W. Dermal carotenoid measurements via pressure mediated reflection spectroscopy. *Journal of Biophotonics*, 2012, vol. 5, no. 7, pp. 559–570. <https://doi.org/10.1002/jbio.201100122>
46. Haynes W.M. *CRC Handbook of Chemistry and Physics*. CRC press, 2016, 2670 p.
47. Thomas L.W. The chemical composition of adipose tissue of man and mice. *Quarterly Journal of Experimental Physiology and Cognate Medical Sciences*, 1962, vol. 47, no. 2, pp. 179–188. <https://doi.org/10.1113/expphysiol.1962.sp001589>
48. Mujkić R., Mujkić D.Š., Ilić I., Rodak E., Šumanovac A., Grgić A., Divković D., Selthofer-Relatić K. Early childhood fat tissue

49. Arenander E., Grimby G., Hallberg L., Westling H., Carlsten A. The volume and distribution of blood in patients with varicose veins // *Scandinavian Journal of Clinical and Laboratory Investigation*. 1962. V. 14. N 2. P. 170–178. <https://doi.org/10.3109/00365516209079690>
50. Hale G.M., Query M.R. Optical constants of water in the 200-nm to 200- $\mu$ m wavelength region // *Applied Optics*. 1973. V. 12. N 3. P. 555–563. <https://doi.org/10.1364/AO.12.000555>
51. Calabro K. Modeling biological tissues in light tools [Электронный ресурс]. URL: <https://www.synopsys.com/content/dam/synopsys/optical-solutions/documents/modeling-tissues-lighttools-paper.pdf>, свободный. Яз. англ. (дата обращения: 09.05.2023)
52. Lister T., Wright P.A., Chappell P.H. Optical properties of human skin // *Journal of Biomedical Optics*. 2012. V. 17. N 9. P. 090901. <https://doi.org/10.1117/1.JBO.17.9.090901>
53. Tozar T., Andrei I.R., Costin R., Pirvulescu R., Pascu M.L. Case series about ex vivo identification of squamous cell carcinomas by laser-induced autofluorescence and Fourier transform infrared spectroscopy // *Lasers in Medical Science*. 2018. V. 33. N 4. P. 861–869. <https://doi.org/10.1007/s10103-018-2445-5>
54. Sekar S.K.V., Bargigia I., Mora A.D., Taroni P., Ruggeri A., Tosi A., Farina A. Diffuse optical characterization of collagen absorption from 500 to 1700 nm // *Journal of Biomedical Optics*. 2017. V. 22. N 1. P. 015006. <https://doi.org/10.1117/1.JBO.22.1.015006>
55. Du H., Fuh R.-C.A., Li J., Corkan L.A., Lindsey J.S. PhotochemCAD: a computer-aided design and research tool in photochemistry // *Photochemistry and Photobiology*. 1998. V. 68. N 2. P. 141–142. <https://doi.org/10.1111/j.1751-1097.1998.tb02480.x>
56. Bosschaart N., Edelman G.J., Aalders M.C.G., Van Leeuwen T.G., Faber D.J. A literature review and novel theoretical approach on the optical properties of whole blood // *Lasers in Medical Science*. 2014. V. 29. N 2. P. 453–479. <https://doi.org/10.1007/s10103-013-1446-7>
57. Sommer A., Van Mierlo P.L.H., Neumann H.A.M., Kessels A.G.H. Red and blue telangiectasias: Differences in oxygenation? // *Dermatologic Surgery*. 1997. V. 23. N 1. P. 55–59. <https://doi.org/10.1111/j.1524-4725.1997.tb00009.x>
58. Khatun F., Aizu Y., Nishidate I. In vivo transcutaneous monitoring of hemoglobin derivatives using a red-green-blue camera-based spectral imaging technique // *International Journal of Molecular Sciences*. 2021. V. 22. N 4. P. 1528. <https://doi.org/10.3390/ijms22041528>
59. Kuenstner J.T., Norris K.H. Spectrophotometry of human hemoglobin in the near infrared region from 1000 to 2500 nm // *Journal of Near Infrared Spectroscopy*. 1994. V. 2. N 2. P. 59–65. <https://doi.org/10.1255/jnirs.32>
60. Al'tshuler G.B., Smirnov M.Z., Pushkareva A.E. Modeling of the laser and lamp treatment of telangiectasia // *Optics and Spectroscopy*. 2004. V. 97. N 1. P. 141–144. <https://doi.org/10.1134/1.1781295>
- changes—adipocyte morphometry, collagen deposition, and expression of CD163+ cells in subcutaneous and visceral adipose tissue of male children. *International Journal of Environmental Research and Public Health*, 2021, vol. 18, no. 7, pp. 3627. <https://doi.org/10.3390/ijerph18073627>
49. Arenander E., Grimby G., Hallberg L., Westling H., Carlsten A. The volume and distribution of blood in patients with varicose veins. *Scandinavian Journal of Clinical and Laboratory Investigation*, 1962, vol. 14, no. 2, p. 170–178. <https://doi.org/10.3109/00365516209079690>
50. Hale G.M., Query M.R. Optical constants of water in the 200-nm to 200- $\mu$ m wavelength region. *Applied Optics*, 1973, vol. 12, no. 3, pp. 555–563. <https://doi.org/10.1364/AO.12.000555>
51. Calabro K. *Modeling biological tissues in light tools* Available at: <https://www.synopsys.com/content/dam/synopsys/optical-solutions/documents/modeling-tissues-lighttools-paper.pdf> (accessed: 09.05.2023)
52. Lister T., Wright P.A., Chappell P.H. Optical properties of human skin. *Journal of Biomedical Optics*, 2012, vol. 17, no. 9, pp. 090901. <https://doi.org/10.1117/1.JBO.17.9.090901>
53. Tozar T., Andrei I.R., Costin R., Pirvulescu R., Pascu M.L. Case series about ex vivo identification of squamous cell carcinomas by laser-induced autofluorescence and Fourier transform infrared spectroscopy. *Lasers in Medical Science*, 2018, vol. 33, no. 4, pp. 861–869. <https://doi.org/10.1007/s10103-018-2445-5>
54. Sekar S.K.V., Bargigia I., Mora A.D., Taroni P., Ruggeri A., Tosi A., Farina A. Diffuse optical characterization of collagen absorption from 500 to 1700 nm. *Journal of Biomedical Optics*, 2017, vol. 22, no. 1, pp. 015006. <https://doi.org/10.1117/1.JBO.22.1.015006>
55. Du H., Fuh R.-C.A., Li J., Corkan L.A., Lindsey J.S. PhotochemCAD: a computer-aided design and research tool in photochemistry. *Photochemistry and Photobiology*, 1998, vol. 68, no. 2, pp. 141–142. <https://doi.org/10.1111/j.1751-1097.1998.tb02480.x>
56. Bosschaart N., Edelman G.J., Aalders M.C.G., Van Leeuwen T.G., Faber D.J. A literature review and novel theoretical approach on the optical properties of whole blood. *Lasers in Medical Science*, 2014, vol. 29, no. 2, pp. 453–479. <https://doi.org/10.1007/s10103-013-1446-7>
57. Sommer A., Van Mierlo P.L.H., Neumann H.A.M., Kessels A.G.H. Red and blue telangiectasias: Differences in oxygenation? *Dermatologic Surgery*, 1997, vol. 23, no. 1, pp. 55–59. <https://doi.org/10.1111/j.1524-4725.1997.tb00009.x>
58. Khatun F., Aizu Y., Nishidate I. In vivo transcutaneous monitoring of hemoglobin derivatives using a red-green-blue camera-based spectral imaging technique. *International Journal of Molecular Sciences*, 2021, vol. 22, no. 4, pp. 1528. <https://doi.org/10.3390/ijms22041528>
59. Kuenstner J.T., Norris K.H. Spectrophotometry of human hemoglobin in the near infrared region from 1000 to 2500 nm. *Journal of Near Infrared Spectroscopy*, 1994, vol. 2, no. 2, pp. 59–65. <https://doi.org/10.1255/jnirs.32>
60. Al'tshuler G.B., Smirnov M.Z., Pushkareva A.E. Modeling of the laser and lamp treatment of telangiectasia. *Optics and Spectroscopy*, 2004, vol. 97, no. 1, pp. 141–144. <https://doi.org/10.1134/1.1781295>

#### Авторы

**Беликов Андрей Вячеславович** — доктор физико-математических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; научный сотрудник, Первый Санкт-Петербургский государственный медицинский университет имени академика И.П. Павлова, Санкт-Петербург, 197022, Российская Федерация, [sc 7005600018](https://orcid.org/0000-0001-8733-4860), <https://orcid.org/0000-0001-8733-4860>, [avbelikov@gmail.com](mailto:avbelikov@gmail.com)

**Чучин Виктор Юрьевич** — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; инженер-исследователь, ООО «Научно-производственное предприятие волоконно-оптическое и лазерного оборудования», Санкт-Петербург, 199034, Российская Федерация, <https://orcid.org/0009-0001-5062-1265>, [treasure\\_planet@mail.ru](mailto:treasure_planet@mail.ru)

#### Authors

**Andrey V. Belikov** — D.Sc. (Physics & Mathematics), Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation; Senior Researcher, Pavlov University, Saint Petersburg, 197022, Russian Federation, [sc 7005600018](https://orcid.org/0000-0001-8733-4860), <https://orcid.org/0000-0001-8733-4860>, [avbelikov@gmail.com](mailto:avbelikov@gmail.com)

**Viktor Yu. Chuchin** — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation; Research Engineer, Limited Liability Company “NPP VOLO”, Saint Petersburg, 199034, Russian Federation, <https://orcid.org/0009-0001-5062-1265>, [treasure\\_planet@mail.ru](mailto:treasure_planet@mail.ru)

Статья поступила в редакцию 10.05.2023  
Одобрена после рецензирования 14.06.2023  
Принята к печати 27.07.2023

Received 10.05.2023  
Approved after reviewing 14.06.2023  
Accepted 27.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-696-702

## Low-temperature cell for IR Fourier spectrometric investigation of hydrocarbon substances

Alisher A. Kenbay<sup>1</sup>, Oleg Yu. Golikov<sup>2</sup>, Abdurakhman U. Aldiyarov<sup>3</sup>, Darkhan E. Yerezhep<sup>4</sup>✉

<sup>1,2,3,4</sup> Al-Farabi Kazakh National University, Almaty, 050040, Kazakhstan

<sup>4</sup> Institute of Physics and Technology, Satbayev University, Almaty, 050032, Kazakhstan

<sup>4</sup> Satbayev University, Almaty, 050013, Kazakhstan

<sup>1</sup> [mr.kenbay@gmail.com](mailto:mr.kenbay@gmail.com), <https://orcid.org/0000-0001-6803-5873>

<sup>2</sup> [golikov@physics.kz](mailto:golikov@physics.kz), <https://orcid.org/0000-0002-6691-8346>

<sup>3</sup> [abdurakhman.aldiyarov@kaznu.kz](mailto:abdurakhman.aldiyarov@kaznu.kz), <https://orcid.org/0000-0002-5091-7699>

<sup>4</sup> [darhan\\_13@physics.kz](mailto:darhan_13@physics.kz)✉, <https://orcid.org/0000-0002-2232-2911>

### Abstract

A specialized low-temperature measuring cell with a cryogenic capillary system for infrared spectral analysis of ethanol developed by the authors is presented. The use of the created low-temperature cell is possible for further studies of the low-temperature properties of both pure ethanol and mixtures with its contents, which is currently an urgent task, and the data obtained with its help can be used for ice research. Two methods of ethanol research at low temperature are presented in comparison. In the first method proposed by the authors, a specially developed low-temperature measuring cell based on a diffuse reflection prefix of the Fourier spectrometer FSM 2203 with a cryogenic capillary system is used. This system allows you to achieve the required low-temperature regime at normal atmospheric pressure. The results of the experiment are compared with the traditional method of gas-phase condensation of the test sample under low temperature conditions at the pressure  $P = 1.0 \cdot 10^{-5}$  Torr. Infrared spectra of low molecular weight amorphous and crystalline ethanol were obtained at a temperature of 150 K at normal atmospheric pressure and in vacuum. Comparison of experimental results confirmed the operability of the new installation. In the experiments, peaks were observed in the absorption bands from 2850 to 3000  $\text{cm}^{-1}$  and from 2950 to 3100  $\text{cm}^{-1}$ , corresponding to the valence CH vibrations of ethanol as well as in the absorption bands from 3150 to 3400  $\text{cm}^{-1}$  and from 3300 to 3500  $\text{cm}^{-1}$ , which corresponds to the valence vibrations of OH. The results of the study showed the prospects of the proposed method and can be useful by researchers in the field of low-temperature spectroscopy at normal pressure.

### Keywords

low-temperature cell, cryogenic capillary system, ethanol, IR spectra, diffuse reflection

### Acknowledgements

This research was carried out under grant AP15473758 with the financial support of the Ministry of Education and Science of the Republic of Kazakhstan.

**For citation:** Kenbay A.A., Golikov O.Yu., Aldiyarov A.U., Yerezhep D.E. Low-temperature cell for IR Fourier spectrometric investigation of hydrocarbon substances. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 696–702. doi: 10.17586/2226-1494-2023-23-4-696-702

УДК 543.42

## Низкотемпературная ячейка для инфракрасных фурье-спектрометрических исследований углеводородных веществ

Алишер Асылбекулы Кенбай<sup>1</sup>, Олег Юрьевич Голиков<sup>2</sup>,

Абдурахман Уалиевич Алдияров<sup>3</sup>, Дархан Есейұлы Ережеп<sup>4</sup>✉

<sup>1,2,3,4</sup> Казахский Национальный Университет им. аль-Фараби, Алматы, 050040, Республика Казахстан

<sup>4</sup> Физико-технический институт, КазНИТУ им. К. И. Сатпаева, Алматы, 050032, Республика Казахстан

<sup>4</sup> КазНИТУ им. К. И. Сатпаева, Алматы, 050013, Республика Казахстан

© Kenbay A.A., Golikov O.Yu., Aldiyarov A.U., Yerezhep D.E., 2023

<sup>1</sup> mr.kenbay@gmail.com, <https://orcid.org/0000-0001-6803-5873>

<sup>2</sup> golikov@physics.kz, <https://orcid.org/0000-0002-6691-8346>

<sup>3</sup> Abdurakhman.Aldiyarov@kaznu.kz, <https://orcid.org/0000-0002-5091-7699>

<sup>4</sup> darhan\_13@physics.kz, <https://orcid.org/0000-0002-2232-2911>

#### Аннотация

**Введение.** Разработана специализированная низкотемпературная измерительная ячейка с криогенной капиллярной системой для инфракрасного спектрального анализа этанола. Созданная низкотемпературная ячейка может быть применена для исследований низкотемпературных свойств чистого этанола и смесей с его содержимым. Такое использование в настоящее время является актуальной задачей, а получаемые с ее помощью данные могут найти применение для исследования льдов. **Метод.** Выполнено сравнение двух методов исследования этанола при низкой температуре. В первом, предложенном методе применена специально разработанная низкотемпературная измерительная ячейка на базе приставки диффузного отражения Фурье-спектрометра ФСМ 2203 с криогенной капиллярной системой. Использование системы позволило получить требуемый низкотемпературный режим при нормальном атмосферном давлении. Результаты эксперимента сопоставлены с традиционным методом газофазной конденсации исследуемого образца в условиях низкой температуры при давлении  $P = 1,0 \cdot 10^{-5}$  торр. **Результаты.** Получены инфракрасные спектры низкомолекулярного аморфного и кристаллического этанола при температуре 150 К, нормальном атмосферном давлении и в вакууме. Сравнение экспериментальных результатов подтвердило работоспособность новой установки. В результате экспериментов наблюдались пики в полосах поглощения от 2850 до 3000  $\text{см}^{-1}$  и от 2950 до 3100  $\text{см}^{-1}$ , соответствующие валентным СН-колебаниям этанола, а также в полосах поглощения от 3150 до 3400  $\text{см}^{-1}$  и от 3300 до 3500  $\text{см}^{-1}$ , что соответствует валентным колебаниям ОН. **Обсуждение.** Полученные результаты показали перспективность предложенного метода и могут быть полезны исследователями в области низкотемпературной спектроскопии при нормальном давлении.

#### Ключевые слова

низкотемпературная ячейка, криогенная капиллярная система, этанол, ИК спектры, диффузное отражение

#### Благодарности

Исследование выполнено при финансовой поддержке Министерства образования и науки Республики Казахстан, грант AP15473758.

**Ссылка для цитирования:** Кенбай А.А., Голиков О.Ю., Алдияров А.У., Ережеп Д.Е. Низкотемпературная ячейка для инфракрасных фурье-спектрометрических исследований углеводородных веществ // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 696–702 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-4-696-702

## Introduction

Amorphous ices are the dominant form of water and alcohol in the universe, but they do not naturally form on Earth's lithosphere. Amorphous ices mostly form on interstellar dust, in comets, and many other astrophysical environments including the Saturnian rings. For example, crystalline benzene has been identified in the atmosphere of Titan [1] within the Solar System, while benzene ice has been detected in the direction of a proto-planetary nebula beyond the Solar System [2]. There are plenty of conditions of the amorphous ice formation in nature. Crystalline ice may amorphize under the influence of UV- or ion-irradiation [3]. A significant amount of laboratory researches has been conducted on extraterrestrial ices at low temperatures, primarily focusing on analyzing the infrared (IR) spectra and other characteristics of amorphous molecular solids that consist of one or more components [4]. Unfortunately, this emphasis on amorphous solids has often resulted in the neglect of studying crystalline compounds, despite their presence in various environments within the Solar System.

Among the organic matters that have been found in ices by astronomers were also methanol ( $\text{CH}_3\text{OH}$ ) which has received the most detailed and widespread attention [5–8], and solid ethanol ( $\text{C}_2\text{H}_5\text{OH}$ ) which has almost been ignored by astrochemists. Taking into account the importance of ethanol for such areas as energy [9–12], medicine [13, 14], industry [15, 16], study of ethanol in its various states

and conditions is very relevant. Insufficient information about amorphous  $\text{C}_2\text{H}_5\text{OH}$ , in which phase it is expected in many extraterrestrial environments, makes it impossible to understand and evaluate ethanol-ice by IR astronomers [17].

One of the methods that can help us to get more information about amorphous ethanol is IR spectroscopy that has been invaluable in the study of extraterrestrial ices, both through astronomical observations and laboratory studies. In both cases, the two reference spectroscopic quantities of greatest need are spectral positions and IR intensities, such as band strengths or absorption coefficients for IR peaks. However, although the IR literature contains extensive results on spectral positions, peak assignments, and vibrational modes, quantitative results on intensities for many compounds in the solid state has not been studied in depth [18–21]. The centrality of such reference IR intensities for calculating abundances in astrochemical problems is well illustrated by discussions in both a recent and an older review of icy solids in extraterrestrial environments [22]. IR spectra are crucial in identifying and characterizing extraterrestrial ices, and the quantitative measurements of intensity such as band strengths and absorption coefficients are used to measure abundances. Additionally, optical constants have applications in computational models of solid-phase compositions [23]. IR spectroscopy is the most reliable method for remote identification of molecules and ions in extraterrestrial solids, and it remains possible to not only identify but also

quantify IR detections to obtain molecular abundances by comparing astronomical data to reference measurements from terrestrial laboratories [24]. The IR spectra of ethanol can be used to study the low-temperature chemistry of interstellar clouds and of extraterrestrial objects in our solar system, such as ice-covered moons and trans-Neptunian objects in connection with upcoming NASA and ESA space missions involving IR instrumentation [25].

The present work is aimed at confirming the measurement technique using the developed by us specialized low-temperature cell through conducting fundamental research of ice (amorphous/solid) and liquid ethanol properties in the temperature range of 90–200 K under normal conditions (atmospheric pressure) which are provided by the developed cell.

### Methodology

To research the thermophysical properties of ethanol within the framework of this research, we conducted two types of fundamentally different experiments of studying ethanol at low temperature to compare them later. The first one was conducted at low temperatures and normal atmospheric pressure, the second one — at low temperatures in vacuum chamber.

To conduct the first experiment, it was necessary to create such an installation that will allow researching the properties of ethanol in the IR range at low temperatures and at normal pressure. Currently available diffuse reflection attachments for spectrometers do not have the ability to record the spectral characteristics of substances at temperatures different from the ambient temperature. Nevertheless, the forced choice of the diffuse reflection mode of IR spectroscopy is due to the possibility of creating a modified attachment that would allow to record spectra at different temperatures with the simultaneous possibility of cooling the sample.

The basis on which the modernization was carried out within the framework of this research was the IR Fourier spectrometer FSM 2203 of the INFRASPEK company (St. Petersburg, Russia) with a special attachment for diffuse reflection. This IR Fourier spectrometer allows obtaining high-resolution IR spectra for the qualitative and quantitative analysis of the samples under research. Fourier spectrometer FSM 2203 (Fig. 1) operates in the

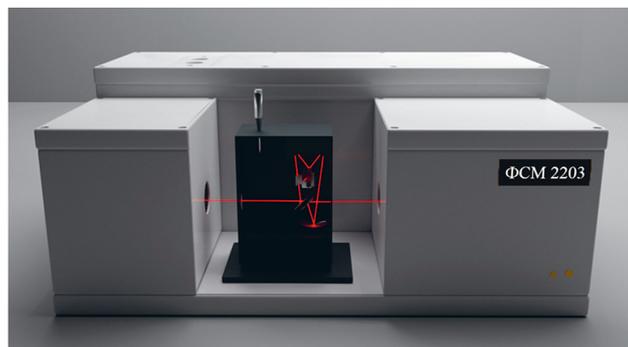


Fig. 1. Infrared Fourier spectrometer “FSM 2203” with a diffuse reflection attachment

Mid-Infrared Region (MIR) with a spectral resolution of  $0.1 \text{ cm}^{-1}$  in the range of  $370\text{--}7800 \text{ cm}^{-1}$ .

Since this research implies an experimental research of ethanol thermophysical properties at low temperatures and at normal pressure, the question of creating the possibility of the test sample temperature controlling inside the diffuse reflection attachment arises (Fig. 2), the sample is placed in the cuvette compartment of the FSM 2203 spectrometer. The greatest importance is the possibility of lowering the temperature to 80–90 K. The diffuse reflection method has minimal requirements for preparing a sample of the test material. It is possible to research irregular surfaces or coatings including polymer ones. In addition, it is possible to obtain spectra in a wide range, which gives it an advantage over the classic transmission measurement method.

The main elements of this attachment are the system of mirrors and a retractable sample holder (Fig. 2, pos. 4) in the cells of which the test samples are placed. A micrometer screw (Fig. 2, pos. 1) is used to adjust the vertical position of the sample holder by moving the directive (Fig. 2, pos. 5) in the direction of the vertical axis. The radiation passing through the inlet of the attachment hits the surface of a flat double-sided mirror (Fig. 2, pos. 2), being reflected from it, the beam falls on an elliptical mirror (Fig. 2, pos. 3) which focuses it on the surface of the test sample located in the cell retractable holder. The beam diffusely reflected from the sample surface again hits the elliptical mirror which redirects it to the spherical mirror (Fig. 2, pos. 6), then the beam hits the other side of the two-sided flat mirror exiting into the detector chamber of the FSM 2203 spectrometer. Primary processing of the received IR spectra is carried out in the FSpec software, the analysis is performed using the Origin software.

To be able to lower the test sample temperature inside the cell of the retractable holder of this attachment, a specially developed modification based on the standard

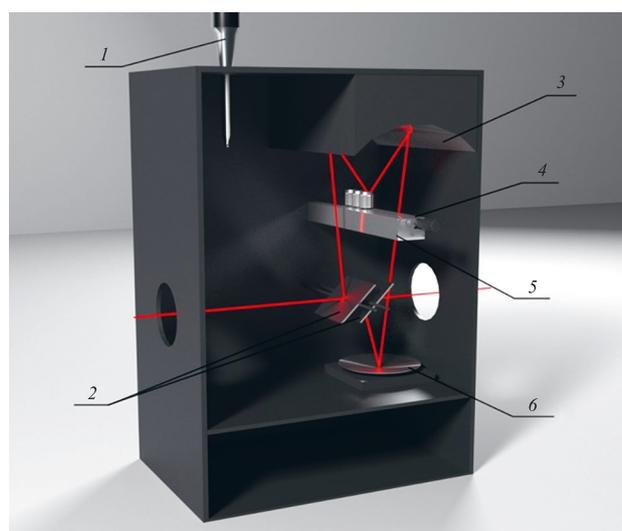


Fig. 2. Schematic representation of the DRA diffuse reflection attachment: 1 — micrometer screw; 2 — double-sided flat mirror; 3 — elliptical mirror; 4 — retractable sample holder; 5 — directive; 6 — spherical mirror

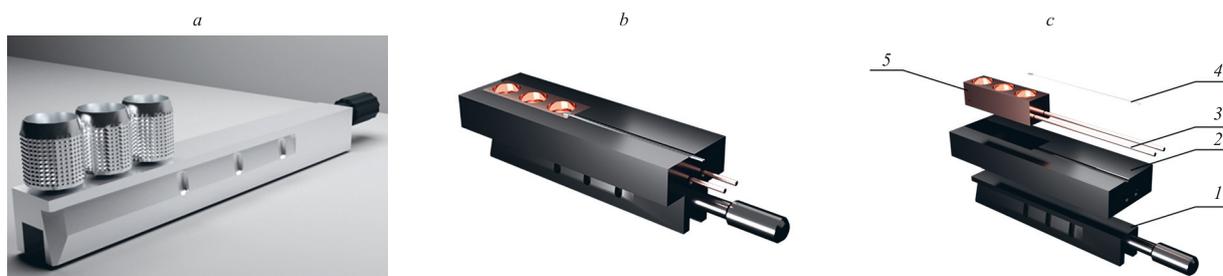


Fig. 3. Three-dimensional model of a retractable holder: standard retractable sample holder (a); modernized specialized retractable sample holder, assembled (b); modernized specialized retractable sample holder, disassembled (c).

1 — holder handle; 2 — heat insulating cover of the holder; 3 — cryogenic capillary system; 4 — temperature sensor; 5 — container

sample holder was used (Fig. 3). The design of the upgraded holder is made up of two main materials: polylactide and copper. The choice of these materials for specific parts of the holder is determined by their thermal conductivity. Thus, the choice of polylactide, which has a thermal conductivity of about 0.11–0.19 W/(m·K), is made with the purpose to prevent the negative effect of the temperature of the external atmosphere on the temperatures of the copper bar and samples placed in the measuring cells. The bar with cylindrical holes for placing the test samples (Fig. 3, pos. 5) is made of copper, a material with high thermal conductivity, to enable rapid and uniform distribution of the temperature supplied through the cryogenic capillary system. The cryogenic capillary system (Fig. 3, pos. 3) is also made of copper for the same reasons.

Cooling of the studied samples to the temperature of 80 K is carried out on account of the continuous circulation of liquid and gaseous nitrogen inside the copper tube of the cryogenic capillary system. The direct contact of the copper tube and copper bar with the cells, as well as the high thermal conductivity of copper, make it possible to achieve rapid cooling of the test substance (in our case, ethanol) to the required temperature. The current temperature of the copper bar is measured using the LakeShore PID controller thermocouple (Fig. 3, pos. 4) mounted on the surface of the copper bar. Heating of the test samples is achieved by stopping nitrogen circulation through the cryogenic capillary system and restoring the thermodynamic equilibrium between the test substance and the environment. The nitrogen circulation through the copper tubes of the cryogenic capillary system is achieved by establishing a high pressure by heating in a special Dewar vessel. As a result, nitrogen vapor is “squeezed out” from the vessel into the cryogenic capillary system with the help of which cooling of the samples is achieved.

For the sake of the experiment purity, a blowing system was also added to the diffuse reflection attachment. Gaseous nitrogen supplied inside the attachment creates the necessary inert environment and also prevents frost formation on the surface of the samples and the copper bar. Ethanol produced by SeccoSolv was chosen as the test substance (purity  $\geq 99.9\%$ , dried, Darmstadt, Germany).

The second experiment (Physical Vapor Deposition (PVD)) was conducted with using experimental setup for cryovacuum condensation [26] shown in Fig. 4.

A substrate (Fig. 4, pos. 6) is thermally connected to a closed-cycle helium Gifford-McMahon refrigerator and

placed within the vacuum chamber (Fig. 4, pos. 1). The cooling system has two stages, first is capable of cooling the substrate to a temperature of 7.5 K and the second one — cooling to 12 K. A resistor heater is attached to the second stage. The PID temperature controller LakeShore allows for temperature adjustments between 12–200 K [27]. To obtain the condensed ethanol sample, the gas phase was deposited onto the substrate at a pressure of  $P = 1.0 \cdot 10^{-5}$  Torr while maintaining the substrate temperature at 150 K. IR radiation supplied from an external source reflects from sample, passes through a system of mirrors and goes into FSM 2203 where it is being processed and transformed into spectra we receive.

## Results and discussion

In our researches, the main attention was paid to the IR spectra of low molecular weight amorphous and crystalline ethanol in the range from 2700 to 4000  $\text{cm}^{-1}$ , obtained under normal conditions at a temperature of 90–200 K, with an emphasis on amorphous ices at temperature of

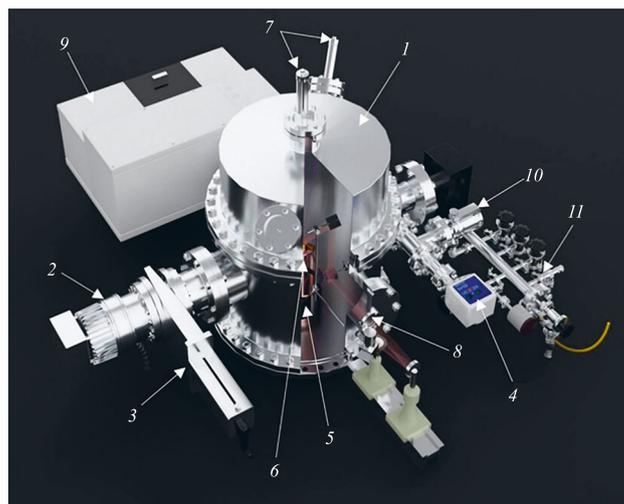


Fig. 4. Experimental setup for cryovacuum condensation: 1 — vacuum chamber; 2 — vacuum pump Turbo-V-301; 3 — vacuum gate valve CFF-100; 4 — pressure detector FRG-700; 5 — Gifford-McMahon refrigerator; 6 — substrate; 7 — photo multiplier, laser interferometer; 8 — light source, optical channel; 9 — IR spectrometer; 10 — high-precision gas supply leak into the chamber; and 11 — gas leak into the mixture production system

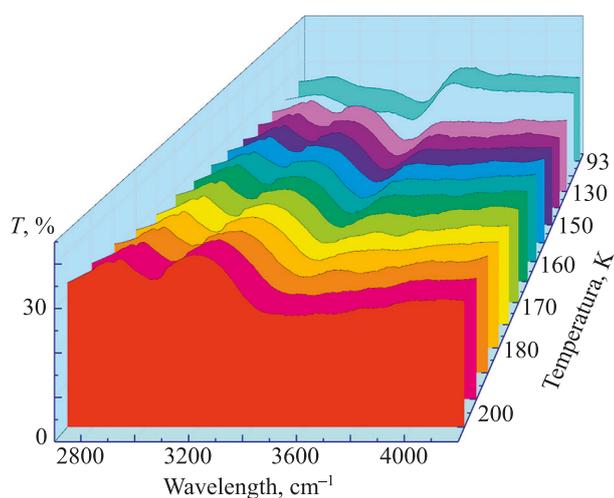


Fig. 5. A family of transmittance ( $T$ ) of the IR spectra of ethanol in the temperature range of 93–200 K

150 K. The family of obtained spectra and their temperature evolution are shown on Fig. 5. IR spectra were recorded from a sample placed in a measuring low-temperature cell at room temperature and subsequently cooled with cryogenic capillary system in a special attachment to a temperature of 93 K. Direct recording was carried out by heating the sample from the minimum obtained temperature to room temperature. The temperature interval from 200 K to 300 K was cut out because of the absence of any important structural-phase changes in the sample in this interval. The inner part of the diffuse reflection attachment, as mentioned earlier, was continuously blowing with nitrogen to create the most inert environment and exclude the possibility of foreign substances getting inside the sample under research.

Main result is shown on the Fig. 6. How you can see on the figure, we conducted two type of experiments and now present comparing and analysis of them. First experiment was conducted under normal atmospheric pressure using developed low-temperature cell with cryogenic capillary system which provides cooling of ethanol, and second one in vacuum uses cryovacuum chamber with PVD method. In the experiment with designed low-temperature cell, we can observe hypsochromic shift at 150 K relative to the experiment in vacuum. This is due to the method of obtaining a solid sample, since the vacuum sample is obtained using PVD, and the atmospheric sample is obtained by cooling the liquid phase ethanol. Despite the shift, we observe the peaks in absorption bands from  $2850\text{ cm}^{-1}$  to  $3000\text{ cm}^{-1}$  and from  $2950\text{ cm}^{-1}$  to  $3100\text{ cm}^{-1}$ , which correspond to the C–H stretching vibrations of ethanol in experiment with PVD and with Cell, respectively, from  $3150\text{ cm}^{-1}$  to  $3400\text{ cm}^{-1}$  and from

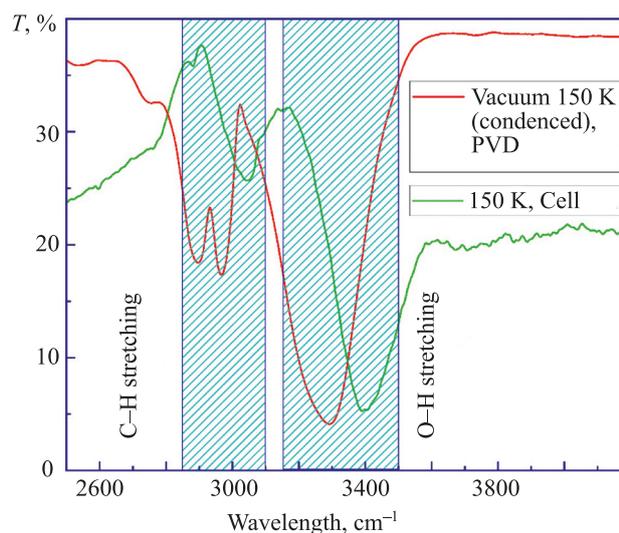


Fig. 6. Transmittance ( $T$ ) of the IR spectrum of ethanol at 150 K from two experiments

$3300\text{ cm}^{-1}$  to  $3500\text{ cm}^{-1}$ , which correspond to the O–H stretching vibrations in the experiments with PVD and Cell, respectively.

Presence of ethanol functional groups characteristic peaks confirms working capability and efficiency of measurement technique with our cryogenic capillary system and particularly the low-temperature cell.

## Conclusion

The initial motivation for this article arose from desire to modernize existing methods in the study of hydrocarbons with habitual equipment and create fundamentally new installation to research them at low temperatures but under normal atmospheric pressure. The present work contains useful results about successful modernizing of diffuse reflection attachment for IR Fourier spectrometer. The developed at Al-Farabi University cryogenic capillary system with low-temperature cell is completely workable. These results can be used, for example, for a new researching works in the sphere of low-temperature spectroscopy at normal pressure. This cell and measurement technique made it possible to obtain low-temperature spectra of ethanol under atmospheric conditions. This makes it possible to state the operability of this method as well as the possibility of conducting a number of further researches aimed at studying the thermophysical properties of many hydrocarbons including monohydric alcohols at low temperatures. Using the developed low-temperature cell will help to understand how many types of hydrocarbons behave at low temperatures and atmospheric pressure.

## References

- Vinatier S., Schmitt B., Bézard B., Rannou P., Dauphin C., de Kok R., Jennings D.E., Flasar F.M. Study of Titan's fall southern stratospheric polar cloud composition with Cassini/CIRS: Detection of benzene ice. *Icarus*, 2018, vol. 310, pp. 89–104. <https://doi.org/10.1016/j.icarus.2017.12.040>
- Cernicharo J., Heras A.M., Tielens A.G.G.M., Pardo J.R., Herpin F., Guélin M., Waters L.B.F.M. Infrared space observatory's discovery of C<sub>4</sub>H<sub>2</sub>, C<sub>6</sub>H<sub>2</sub>, and benzene in CRL 618. *Astrophysical Journal*, 2001, vol. 546, no. 2, pp. L123–L126. <https://doi.org/10.1086/318871>
- Loerting T., Fuentes-Landete V., Handle P.H., Seidl M., Amann-Winkel K., Gainaru C., Böhmer R. The glass transition in high-density amorphous ice. *Journal of Non-Crystalline Solids*, 2015, vol. 407, pp. 423–430. <https://doi.org/10.1016/j.jnoncrysol.2014.09.003>
- Yarnall Y.Y., Hudson R.L. Crystalline ices — Densities and comparisons for planetary and interstellar applications. *Icarus*, 2022, vol. 373, pp. 114799. <https://doi.org/10.1016/j.icarus.2021.114799>
- Yarnall Y.Y., Hudson R.L. Infrared intensities of methyl acetate, an interstellar compound — comparisons of three organic esters. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2022, vol. 283, pp. 121738. <https://doi.org/10.1016/j.saa.2022.121738>
- Gibb E.L., Whittet D.C.B., Boogert A.C.A., Tielens A.G.G.M. Interstellar ice: The Infrared Space Observatory legacy. *Astrophysical Journal Supplement Series*, 2004, vol. 151, no. 1, pp. 35–73. <https://doi.org/10.1086/381182>
- Hudson R.L., Mullikin E.F. Infrared band strengths for amorphous and crystalline methyl propionate, a candidate interstellar molecule. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, vol. 207, pp. 216–221. <https://doi.org/10.1016/j.saa.2018.09.032>
- Allamandola L.J., Sandford S.A., Tielens A.G.G.M., Herbst T.M. Infrared spectroscopy of dense clouds in the C-H stretch region - Methanol and “diamonds”. *Astrophysical Journal*, 1992, vol. 399, pp. 134. <https://doi.org/10.1086/171909>
- Shelar M.N., Matsagar V.K., Patil V.S., Barahate S.D. Net energy analysis of sugarcane based ethanol production. *Cleaner Energy Systems*, 2023, vol. 4, pp. 100059. <https://doi.org/10.1016/j.cles.2023.100059>
- Li X., Dong Y., Chang L., Chen L., Wang G., Zhuang Y., Yan X. Dynamic hybrid modeling of fuel ethanol fermentation process by integrating biomass concentration XGBoost model and kinetic parameter artificial neural network model into mechanism model. *Renewable Energy*, 2023, vol. 205, pp. 574–582. <https://doi.org/10.1016/j.renene.2023.01.113>
- Kumar S., Cho J.H., Park J., Moon I. Advances in diesel–alcohol blends and their effects on the performance and emissions of diesel engines. *Renewable and Sustainable Energy Reviews*, 2013, vol. 22, pp. 46–72. <https://doi.org/10.1016/j.rser.2013.01.017>
- Wei L., Cheung C.S., Ning Z. Effects of biodiesel-ethanol and biodiesel-butanol blends on the combustion, performance and emissions of a diesel engine. *Energy*, 2018, vol. 155, pp. 957–970. <https://doi.org/10.1016/j.energy.2018.05.049>
- Miao W.G., Tang C., Ye Y., Quinn R.J., Feng Y. Traditional Chinese medicine extraction method by ethanol delivers drug-like molecules. *Chinese Journal of Natural Medicines*, 2019, vol. 17, no. 9, pp. 713–720. [https://doi.org/10.1016/s1875-5364\(19\)30086-x](https://doi.org/10.1016/s1875-5364(19)30086-x)
- Day S.M., Gironde S.C., Clarke C.W., Snipes J.A., Nicol N.I., Kamran H., Vaughan W., Weiner J.L., Macauley S.L. Ethanol exposure alters Alzheimer's-related pathology, behavior, and metabolism in APP/PS1 mice. *Neurobiology of Disease*, 2023, vol. 177, pp. 105967. <https://doi.org/10.1016/j.nbd.2022.105967>
- Zhang L., Shen Q., Pang C.H., Chao W., Tong S., Kow K.W., Lester E., Wu T., Shang L., Song X., Sun N., Wei W. Life cycle assessment of bio-fermentation ethanol production and its influence in China's steeling industry. *Journal of Cleaner Production*, 2023, vol. 397, pp. 136492. <https://doi.org/10.1016/j.jclepro.2023.136492>
- Lui M.Y., Masters A.F., Maschmeyer T., Yuen A.K.L. Molybdenum carbide, supercritical ethanol and base: Keys for unlocking renewable BTEX from lignin. *Applied Catalysis B: Environmental*, 2023, vol. 325, pp. 122351. <https://doi.org/10.1016/j.apcatb.2022.122351>
- Hudson R.L. An IR investigation of solid amorphous ethanol – Spectra, properties, and phase changes. *Spectrochimica Acta Part A:*

## Литература

- Vinatier S., Schmitt B., Bézard B., Rannou P., Dauphin C., de Kok R., Jennings D.E., Flasar F.M. Study of Titan's fall southern stratospheric polar cloud composition with Cassini/CIRS: Detection of benzene ice // *Icarus*. 2018. V. 310. P. 89–104. <https://doi.org/10.1016/j.icarus.2017.12.040>
- Cernicharo J., Heras A.M., Tielens A.G.G.M., Pardo J.R., Herpin F., Guélin M., Waters L.B.F.M. Infrared space observatory's discovery of C<sub>4</sub>H<sub>2</sub>, C<sub>6</sub>H<sub>2</sub>, and benzene in CRL 618 // *Astrophysical Journal*. 2001. V. 546. N 2. P. L123–L126. <https://doi.org/10.1086/318871>
- Loerting T., Fuentes-Landete V., Handle P.H., Seidl M., Amann-Winkel K., Gainaru C., Böhmer R. The glass transition in high-density amorphous ice // *Journal of Non-Crystalline Solids*. 2015. V. 407. P. 423–430. <https://doi.org/10.1016/j.jnoncrysol.2014.09.003>
- Yarnall Y.Y., Hudson R.L. Crystalline ices – Densities and comparisons for planetary and interstellar applications // *Icarus*. 2022. V. 373. P. 114799. <https://doi.org/10.1016/j.icarus.2021.114799>
- Yarnall Y.Y., Hudson R.L. Infrared intensities of methyl acetate, an interstellar compound — comparisons of three organic esters // *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2022. V. 283. P. 121738. <https://doi.org/10.1016/j.saa.2022.121738>
- Gibb E.L., Whittet D.C.B., Boogert A.C.A., Tielens A.G.G.M. Interstellar ice: The Infrared Space Observatory legacy // *Astrophysical Journal Supplement Series*. 2004. V. 151. N 1. P. 35–73. <https://doi.org/10.1086/381182>
- Hudson R.L., Mullikin E.F. Infrared band strengths for amorphous and crystalline methyl propionate, a candidate interstellar molecule // *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2019. V. 207. P. 216–221. <https://doi.org/10.1016/j.saa.2018.09.032>
- Allamandola L.J., Sandford S.A., Tielens A.G.G.M., Herbst T.M. Infrared spectroscopy of dense clouds in the C-H stretch region - Methanol and “diamonds” // *Astrophysical Journal*. 1992. V. 399. P. 134. <https://doi.org/10.1086/171909>
- Shelar M.N., Matsagar V.K., Patil V.S., Barahate S.D. Net energy analysis of sugarcane based ethanol production // *Cleaner Energy Systems*. 2023. V. 4. P. 100059. <https://doi.org/10.1016/j.cles.2023.100059>
- Li X., Dong Y., Chang L., Chen L., Wang G., Zhuang Y., Yan X. Dynamic hybrid modeling of fuel ethanol fermentation process by integrating biomass concentration XGBoost model and kinetic parameter artificial neural network model into mechanism model // *Renewable Energy*. 2023. V. 205. P. 574–582. <https://doi.org/10.1016/j.renene.2023.01.113>
- Kumar S., Cho J.H., Park J., Moon I. Advances in diesel–alcohol blends and their effects on the performance and emissions of diesel engines // *Renewable and Sustainable Energy Reviews*. 2013. V. 22. P. 46–72. <https://doi.org/10.1016/j.rser.2013.01.017>
- Wei L., Cheung C.S., Ning Z. Effects of biodiesel-ethanol and biodiesel-butanol blends on the combustion, performance and emissions of a diesel engine // *Energy*. 2018. V. 155. P. 957–970. <https://doi.org/10.1016/j.energy.2018.05.049>
- Miao W.G., Tang C., Ye Y., Quinn R.J., Feng Y. Traditional Chinese medicine extraction method by ethanol delivers drug-like molecules // *Chinese Journal of Natural Medicines*. 2019. V. 17. N 9. P. 713–720. [https://doi.org/10.1016/s1875-5364\(19\)30086-x](https://doi.org/10.1016/s1875-5364(19)30086-x)
- Day S.M., Gironde S.C., Clarke C.W., Snipes J.A., Nicol N.I., Kamran H., Vaughan W., Weiner J.L., Macauley S.L. Ethanol exposure alters Alzheimer's-related pathology, behavior, and metabolism in APP/PS1 mice // *Neurobiology of Disease*. 2023. V. 177. P. 105967. <https://doi.org/10.1016/j.nbd.2022.105967>
- Zhang L., Shen Q., Pang C.H., Chao W., Tong S., Kow K.W., Lester E., Wu T., Shang L., Song X., Sun N., Wei W. Life cycle assessment of bio-fermentation ethanol production and its influence in China's steeling industry // *Journal of Cleaner Production*. 2023. V. 397. P. 136492. <https://doi.org/10.1016/j.jclepro.2023.136492>
- Lui M.Y., Masters A.F., Maschmeyer T., Yuen A.K.L. Molybdenum carbide, supercritical ethanol and base: Keys for unlocking renewable BTEX from lignin // *Applied Catalysis B: Environmental*. 2023. V. 325. P. 122351. <https://doi.org/10.1016/j.apcatb.2022.122351>
- Hudson R.L. An IR investigation of solid amorphous ethanol – Spectra, properties, and phase changes // *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 2017. V. 187. P. 82–86. <https://doi.org/10.1016/j.saa.2017.06.027>

- Molecular and Biomolecular Spectroscopy*, 2017, vol. 187, pp. 82–86. <https://doi.org/10.1016/j.saa.2017.06.027>
18. Materese C.K., Gerakines P.A., Hudson R.L. Laboratory studies of astronomical ices: Reaction chemistry and spectroscopy. *Accounts of Chemical Research*, 2021, vol. 54, no. 2, pp. 280–290. <https://doi.org/10.1021/acs.accounts.0c00637>
  19. Hudgins D.M., Sandford S.A., Allamandola L.J., Tielens A.G.G.M. Mid- and far-infrared spectroscopy of ices - Optical constants and integrated absorbances. *Astrophysical Journal Supplement Series*, 1993, vol. 86, pp. 713. <https://doi.org/10.1086/191796>
  20. Drobyshev A., Aldiyarov A., Sokolov D., Shinbaeva A., Nurmukan A. IR Spectrometry studies of methanol cryovacuum condensates. *Low Temperature Physics*, 2019, vol. 45, no. 4, pp. 441–451. <https://doi.org/10.1063/1.5093525>
  21. Drobyshev A., Aldiyarov A., Sokolov D., Shinbayeva A., Tokmoldin N. Refractive indices vs deposition temperature of thin films of ethanol, methane and nitrous oxide in the vicinity of their phase transition temperatures. *Low Temperature Physics*, 2017, vol. 43, no. 10, pp. 1214–1216. <https://doi.org/10.1063/1.5008415>
  22. Boogert A.C.A., Gerakines P.A., Whittet D.C.B. Observations of the icy universe. *Annual Review of Astronomy and Astrophysics*, 2015, vol. 53, no. 1, pp. 541–581. <https://doi.org/10.1146/annurev-astro-082214-122348>
  23. Hudson R.L. Infrared spectra of benzene ices: Reexamination and comparison of two recent papers and the literature. *Icarus*, 2022, vol. 384, pp. 115091. <https://doi.org/10.1016/j.icarus.2022.115091>
  24. Hudson R.L., Gerakines P.A., Yarnall Y.Y. Ammonia ices revisited: New IR intensities and optical constants for solid NH<sub>3</sub>. *Astrophysical Journal*, 2022, vol. 925, no. 2, pp. 156. <https://doi.org/10.3847/1538-4357/ac3e74>
  25. *The Science of Solar System Ices*. Ed. by M.S. Gudipati, J. Castillo-Rogez. New York, NY, Springer New York, 2013. Astrophysics and Space Science Library; V. 356. <https://doi.org/10.1007/978-1-4614-3076-6>
  26. Sokolov D.Y., Yerezhap D., Vorobyova O., Golikov O., Aldiyarov A.U. Infrared analysis and effect of nitrogen and nitrous oxide on the glass transition of methanol cryofilms. *ACS Omega*, 2022, vol. 7, no. 50, pp. 46402–46410. <https://doi.org/10.1021/acsomega.2c05090>
  27. Sokolov D.Y., Yerezhap D., Vorobyova O., Ramos M.A., Shinbayeva A. Optical studies of thin films of cryocondensed mixtures of water and admixture of nitrogen and argon. *Materials (Basel)*, 2022, vol. 15, no. 21, pp. 441. <https://doi.org/10.3390/ma15217441>
  18. Materese C.K., Gerakines P.A., Hudson R.L. Laboratory studies of astronomical ices: Reaction chemistry and spectroscopy // *Accounts of Chemical Research*. 2021. V. 54. N 2. P. 280–290. <https://doi.org/10.1021/acs.accounts.0c00637>
  19. Hudgins D.M., Sandford S.A., Allamandola L.J., Tielens A.G.G.M. Mid- and far-infrared spectroscopy of ices - Optical constants and integrated absorbances // *Astrophysical Journal Supplement Series*. 1993. V. 86. P. 713. <https://doi.org/10.1086/191796>
  20. Drobyshev A., Aldiyarov A., Sokolov D., Shinbaeva A., Nurmukan A. IR Spectrometry studies of methanol cryovacuum condensates // *Low Temperature Physics*. 2019. V. 45. N 4. P. 441–451. <https://doi.org/10.1063/1.5093525>
  21. Drobyshev A., Aldiyarov A., Sokolov D., Shinbayeva A., Tokmoldin N. Refractive indices vs deposition temperature of thin films of ethanol, methane and nitrous oxide in the vicinity of their phase transition temperatures // *Low Temperature Physics*. 2017. V. 43. N 10. P. 1214–1216. <https://doi.org/10.1063/1.5008415>
  22. Boogert A.C.A., Gerakines P.A., Whittet D.C.B. Observations of the icy universe // *Annual Review of Astronomy and Astrophysics*. 2015. V. 53. N 1. P. 541–581. <https://doi.org/10.1146/annurev-astro-082214-122348>
  23. Hudson R.L. Infrared spectra of benzene ices: Reexamination and comparison of two recent papers and the literature // *Icarus*. 2022. V. 384. P. 115091. <https://doi.org/10.1016/j.icarus.2022.115091>
  24. Hudson R.L., Gerakines P.A., Yarnall Y.Y. Ammonia ices revisited: New IR intensities and optical constants for solid NH<sub>3</sub> // *Astrophysical Journal*. 2022. V. 925. N 2. P. 156. <https://doi.org/10.3847/1538-4357/ac3e74>
  25. *The Science of Solar System Ices* / ed. by M.S. Gudipati, J. Castillo-Rogez. New York, NY: Springer New York, 2013. (Astrophysics and Space Science Library; V. 356). <https://doi.org/10.1007/978-1-4614-3076-6>
  26. Sokolov D.Y., Yerezhap D., Vorobyova O., Golikov O., Aldiyarov A.U. Infrared analysis and effect of nitrogen and nitrous oxide on the glass transition of methanol cryofilms // *ACS Omega*. 2022. V. 7. N 50. P. 46402–46410. <https://doi.org/10.1021/acsomega.2c05090>
  27. Sokolov D.Y., Yerezhap D., Vorobyova O., Ramos M.A., Shinbayeva A. Optical studies of thin films of cryocondensed mixtures of water and admixture of nitrogen and argon // *Materials (Basel)*. 2022. V. 15. N 21. P. 441. <https://doi.org/10.3390/ma15217441>

## Authors

**Alisher A. Kenbay** — Doctoral Student, Al-Farabi Kazakh National University, Almaty, 050040, Kazakhstan, <https://orcid.org/0000-0001-6803-5873>, [mr.kenbay@gmail.com](mailto:mr.kenbay@gmail.com)

**Oleg Yu. Golikov** — Doctoral Student, Junior Researcher, Al-Farabi Kazakh National University, Almaty, 050040, Kazakhstan, [sc 57843805800](https://orcid.org/0000-0002-6691-8346), <https://orcid.org/0000-0002-6691-8346>, [golikov@physics.kz](mailto:golikov@physics.kz)

**Abdurakhman U. Aldiyarov** — PhD (Physics & Mathematics), Associate Professor, Acting Professor, Al-Farabi Kazakh National University, Almaty, 050040, Kazakhstan, [sc 57843805800](https://orcid.org/0000-0002-6691-8346), <https://orcid.org/0000-0002-6691-8346>, [golikov@physics.kz](mailto:golikov@physics.kz)

**Darkhan E. Yerezhap** — PhD, Senior Lecturer, Al-Farabi Kazakh National University, Almaty, 050040, Kazakhstan; Leading Researcher, Institute of Physics and Technology, Satbayev University, Almaty, 050032, Kazakhstan; Associate Professor, Satbayev University, Almaty, 050013, Kazakhstan, [sc 57194012596](https://orcid.org/0000-0002-2232-2911), <https://orcid.org/0000-0002-2232-2911>, [darhan\\_13@physics.kz](mailto:darhan_13@physics.kz)

## Авторы

**Кенбай Алишер Асылбекулы** — докторант, Казахский Национальный Университет им. аль-Фараби, Алматы, 050040, Республика Казахстан, <https://orcid.org/0000-0001-6803-5873>, [mr.kenbay@gmail.com](mailto:mr.kenbay@gmail.com)

**Голиков Олег Юрьевич** — младший научный сотрудник, докторант, Казахский Национальный Университет им. аль-Фараби, Алматы, 050040, Республика Казахстан, [sc 57843805800](https://orcid.org/0000-0002-6691-8346), <https://orcid.org/0000-0002-6691-8346>, [golikov@physics.kz](mailto:golikov@physics.kz)

**Алдияров Абдурахман Уалиевич** — кандидат физико-математических наук, доцент, и.о. профессора, Казахский Национальный Университет им. аль-Фараби, Алматы, 050040, Республика Казахстан, [sc 16201950600](https://orcid.org/0000-0002-5091-7699), <https://orcid.org/0000-0002-5091-7699>, [Abdiyarov@kaznu.kz](mailto:Abdiyarov@kaznu.kz)

**Ережеп Дархан Есейұлы** — кандидат технических наук, старший преподаватель, старший научный сотрудник, Казахский национальный университет им. аль-Фараби, Алматы, 050040, Республика Казахстан; ведущий научный сотрудник, Физико-технический институт, КазНУ им. К. И. Сатпаева, Алматы, 050032, Республика Казахстан; доцент, КазНУ им. К. И. Сатпаева, Алматы, 050013, Республика Казахстан, [sc 57194012596](https://orcid.org/0000-0002-2232-2911), <https://orcid.org/0000-0002-2232-2911>, [darhan\\_13@physics.kz](mailto:darhan_13@physics.kz)

Received 22.06.2022

Approved after reviewing 09.06.2023

Accepted 18.07.2023

Статья поступила в редакцию 22.06.2022

Одобрена после рецензирования 09.06.2023

Принята к печати 18.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

НОВЫЕ МАТЕРИАЛЫ И НАНОТЕХНОЛОГИИ  
MATERIAL SCIENCE AND NANOTECHNOLOGIES

doi: 10.17586/2226-1494-2023-23-4-703-710

УДК 539.23

Особенности выращивания твердых растворов  $Ga_{1-x}In_xAs$  на подложках GaAs в поле температурного градиента через тонкую газовую зонуОлег Васильевич Девицкий<sup>1</sup>✉, Леонид Сергеевич Лунин<sup>2</sup>,  
Даниил Вячеславович Митрофанов<sup>3</sup>, Игорь Александрович Сысоев<sup>4</sup>,  
Дмитрий Александрович Никулин<sup>5</sup>, Олег Михайлович Чапура<sup>6</sup><sup>1,2,5</sup> Федеральный исследовательский центр Южный научный центр Российской академии наук, Ростов-на-Дону, 344006, Российская Федерация<sup>1,2,3,4,5,6</sup> Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация<sup>2,3,5</sup> Южно-Российский государственный политехнический университет (НПИ) им. Платова, Новочеркасск, 346428, Российская Федерация<sup>1</sup> [v2517@rambler.ru](mailto:v2517@rambler.ru)✉, <https://orcid.org/0000-0003-3153-696X><sup>2</sup> [lunin\\_ls@mail.ru](mailto:lunin_ls@mail.ru), <https://orcid.org/0000-0002-5534-9694><sup>3</sup> [danik-20@yandex.ru](mailto:danik-20@yandex.ru), <https://orcid.org/0009-0009-2137-4312><sup>4</sup> [eianpisia@yandex.ru](mailto:eianpisia@yandex.ru), <https://orcid.org/0000-0001-5415-0782><sup>5</sup> [dmitnikul@gmail.com](mailto:dmitnikul@gmail.com), <https://orcid.org/0000-0003-0755-4068><sup>6</sup> [chapuro1-7@mail.ru](mailto:chapuro1-7@mail.ru), <https://orcid.org/0000-0002-6691-0010>

## Аннотация

**Введение.** Твердый раствор  $Ga_{1-x}In_xAs$  имеет широкое применение в современной оптоэлектронике в качестве материала для  $p-i-n$  фотодетекторов и лазеров, излучающих в спектральном диапазоне 1,3–1,55 мкм. Исследованы особенности получения пленок  $Ga_{1-x}In_xAs$  методом зонной перекристаллизации градиентом температур. Смысл метода заключается в последовательной перекристаллизации частей расплава источника, движущегося под действием температурного градиента. **Метод.** В поле температурного градиента 30 К/см через тонкую газовую зону в специально разработанной графитовой cassette получены пленки  $Ga_{1-x}In_xAs$  на подложках GaAs при температуре 1123 К. В качестве газа носителя использована смесь азота и водорода в соотношении 1:1. Толщина газовой зоны между источником и подложкой составила 1 мм, время осаждения для всех пленок — 10 мин. **Основные результаты.** Исследована кинетика роста, морфология и структура химических связей полученных пленок. По результатам теоретического расчета установлено, что увеличение концентрации индия приводит к понижению скорости роста пленок до 0,3137 мкм/мин. Выполнено сравнение теоретического расчета и экспериментальных данных, которое показало расхождение значительной скорости роста для пленок с концентрацией индия в ростовом источнике более 20 %, что связано с сегрегацией индия на поверхность пленки. Среднеквадратическая шероховатость пленок составила от 9,1 до 24,2 нм. Подтверждено, что содержание индия в ростовом источнике существенно влияет на свойства выращенных пленок и приводит к уменьшению скорости роста, увеличению упругих напряжений в слое и нестехиометрическому составу пленки. Установлено, что с увеличением концентрации индия в пленке наблюдается существенное смещение частоты LO- и TO-фононных мод GaAs влево на 13 и 16  $cm^{-1}$  соответственно из-за влияния упругих механических напряжений. **Обсуждение.** Представленные результаты продемонстрировали, что методом зонной перекристаллизации в градиенте температур получены пленки твердого раствора  $Ga_{1-x}In_xAs$ , которые имеют ближний порядок химических связей.

## Ключевые слова

поле температурного градиента, тонкая газовая зона, соединения III–V,  $Ga_{1-x}In_xAs$ , рамановская спектроскопия, атомно-силовая микроскопия

## Благодарности

Работа выполнена в рамках государственного задания ФИЦ ЮНЦ РАН № 122020100254-3 (исследования морфологии поверхности пленок) и № 122020100326-7 (рамановские исследования и определение химического состава), а также с использованием ресурсов центра коллективного пользования Северо-Кавказского федерального университета и при финансовой поддержке Минобрнауки России, уникальный идентификатор

© Девицкий О.В., Лунин Л.С., Митрофанов Д.В., Сысоев И.А., Никулин Д.А., Чапура О.М., 2023

проекта RF-2296.61321X0029 (соглашение № 075-15-2021-687) (получение экспериментальных образцов пленок). Авторы выражают благодарность Северо-Кавказскому федеральному университету за помощь в рамках конкурса поддержки проектов научных групп и отдельных ученых университета.

**Ссылка для цитирования:** Девицкий О.В., Лунин Л.С., Митрофанов Д.В., Сысоев И.А., Никулин Д.А., Чапура О.М. Особенности выращивания твердых растворов Ga<sub>1-x</sub>In<sub>x</sub>As на подложках GaAs в поле температурного градиента через тонкую газовую зону // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 703–710. doi: 10.17586/2226-1494-2023-23-4-703-710

## Peculiarities of growing Ga<sub>1-x</sub>In<sub>x</sub>As solid solutions on GaAs substrates in the field of a temperature gradient through a thin gas zone

Oleg V. Devitsky<sup>1,2</sup>, Leonid S. Lunin<sup>2</sup>, Daniil V. Mitrofanov<sup>3</sup>, Igor A. Sysoev<sup>4</sup>,  
Dmitry A. Nikulin<sup>5</sup>, Oleg M. Chapura<sup>6</sup>

<sup>1,2,5</sup> Federal Research Center the Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don, 344006, Russian Federation

<sup>1,2,3,4,5,6</sup> North-Caucasus Federal University, Stavropol, 355017, Russian Federation

<sup>2,3,5</sup> Platov South-Russian State Polytechnic University (NPI), Novocherkassk, 346428, Russian Federation

<sup>1</sup> v2517@rambler.ru, <https://orcid.org/0000-0003-3153-696X>

<sup>2</sup> lunin\_ls@mail.ru, <https://orcid.org/0000-0002-5534-9694>

<sup>3</sup> danik-20@yandex.ru, <https://orcid.org/0009-0009-2137-4312>

<sup>4</sup> cianpisia@yandex.ru, <https://orcid.org/0000-0001-5415-0782>

<sup>5</sup> dmitnikul@gmail.com, <https://orcid.org/0000-0003-0755-4068>

<sup>6</sup> chapuro1-7@mail.ru, <https://orcid.org/0000-0002-6691-0010>

### Abstract

Solid solution Ga<sub>1-x</sub>In<sub>x</sub>As is widely used in modern optoelectronics as a material for *p-i-n* photodetectors, lasers emitting in the spectral range 1.3–1.55 μm. In this paper, the features of obtaining Ga<sub>1-x</sub>In<sub>x</sub>As films by the method of zone recrystallization with a temperature gradient, the essence of which is the sequential recrystallization of parts of the source melt moving under the action of a temperature gradient, are studied. Ga<sub>1-x</sub>In<sub>x</sub>As films on GaAs substrates were obtained in a temperature gradient field through a thin gas zone in a specially designed graphite cassette. The films were prepared at a temperature of 1123 K with a temperature gradient of 30 K/cm. A 1:1 mixture of nitrogen and hydrogen was used as the carrier gas. The thickness of the gas zone between the source and the substrate was 1 mm. The deposition time for all films was 10 min. The growth kinetics, morphology, and structure of the chemical bonds of the obtained films have been studied. Based on the results of theoretical calculations, it was found that an increase in the concentration of indium leads to a decrease in the film growth rate to 0.3137 μm/min. A comparison of the results of theoretical calculations with experimental results showed a discrepancy between the growth rates for films with an indium concentration in the growth source of more than 20 %, which is primarily due to the segregation of indium on the film surface. The films have an RMS roughness from 9.1 to 24.2 nm. It is shown that the content of indium in the growth source significantly affects the properties of the grown films and leads to a decrease in the growth rate, an increase in the elastic stresses in the layer, and a nonstoichiometric composition of the film. It has been established that with an increase in the indium concentration in the film, a significant shift in the frequency of the LO and TO phonon modes of GaAs to the left by 13 and 16 cm<sup>-1</sup>, respectively, is observed due to the influence of elastic mechanical stresses. The presented results show that Ga<sub>1-x</sub>In<sub>x</sub>As solid solution films with short-range order of chemical bonds were obtained by the method of zone recrystallization in a temperature gradient.

### Keywords

temperature gradient field, thin gas band, III–V compounds, Ga<sub>1-x</sub>In<sub>x</sub>As, Raman spectroscopy, atomic force microscopy

### Acknowledgments

The work was carried out within the framework of the State Assignment no. 122020100254-3 (study of the surface morphology of films) and no. 122020100326-7 (Raman studies and determination of the chemical composition) of the SSC RAS as well as using the resources of the Center for Collective Use of the North Caucasus Federal University and with financial support of the Ministry of Education and Science of the Russian Federation, unique identifier of the project RF-2296.61321X0029 (agreement no. 075-15-2021-687) (obtaining experimental film samples). The authors express their gratitude to the North Caucasus Federal University for their help in the framework of the competition to support projects of scientific groups and individual scientists.

**For citation:** Devitsky O.V., Lunin L.S., Mitrofanov D.V., Sysoev I.A., Nikulin D.A., Chapura O.M. Peculiarities of growing Ga<sub>1-x</sub>In<sub>x</sub>As solid solutions on GaAs substrates in the field of a temperature gradient through a thin gas zone. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 703–710 (in Russian). doi: 10.17586/2226-1494-2023-23-4-703-710

### Введение

Гетероструктуры Ga<sub>1-x</sub>In<sub>x</sub>As/GaAs представляют большой интерес в оптоэлектронике для оптоволоконной техники [1–5]. В настоящее время для получения

данных гетероструктур используются методы: молекулярно-лучевой эпитаксии [6], химическое осаждение из паровой фазы металлоорганических соединений [7, 8], магнетронного напыления [9], импульсного лазерного напыления [10]. Для выращивания твердого раство-

ра  $Ga_{1-x}In_xAs$  на подложках GaAs применяется метод жидкофазной эпитаксии [11, 12]. Известно об использовании одной из модификаций метода жидкофазной эпитаксии — зонной перекристаллизации градиентом температур (ЗПГТ) [13], суть которого заключается в последовательной перекристаллизации частей полупроводника жидкой зоной, движущейся под действием температурного градиента. В настоящей работе использован метод близкого переноса через тонкую газовую зону в поле температурного градиента — модификация метода ЗПГТ [14]. Особенность данного метода — наличие твердого источника заданного состава, с помощью которого осуществляется постоянная подпитка жидкой фазы в процессе роста эпитаксиального слоя. Управление процессом роста и составом пленок выполнено: градиентом температуры, температурой подложки, толщиной газовой зазора и составом источника ростовых компонентов. Квазиравновесные условия роста и изотермичность метода ЗПГТ позволяют получать совершенные полупроводниковые кристаллы, твердые растворы, и как следствие, материалы с высокими оптическими и электрофизическими свойствами.

При получении многокомпонентных твердых растворов III-V соединений применяются только те компоненты, которые при взаимодействии с парами водорода и воды не дают устойчивых оксидов — для  $Ga_{1-x}In_xAs$  (GaAs и InAs) [15]. Процесс градиентной эпитаксии предполагает протекание реакции в обоих направлениях при незначительных изменениях температуры. В зоне подложки при температуре ниже температуры источника происходит процесс синтеза соединений, сопровождающийся освобождением воды транспортера. Возникающий градиент концентраций стимулирует диффузионный перенос газообразных продуктов реакций InAs и GaAs к подложке. Скорость осаждения трехкомпонентного соединения  $Ga_{1-x}In_xAs$  может быть рассчитана по эмпирической формуле [16]:

$$v_{Ga_{1-x}In_xAs} = [(C_{Ga} + C_{As})v_{GaAs} + (C_{In} + C_{As})v_{InAs}]lG \cdot 10^{-5}, \quad (1)$$

где  $l$  — толщина газовой зоны, мкм;  $G$  — температурный градиент;  $C_{Ga}$ ,  $C_{As}$ ,  $C_{In}$  — концентрации компонент в источнике;  $v_{GaAs}$ ,  $v_{InAs}$  — скорости роста бинарных компонентов  $Ga_{1-x}In_xAs$ .

Скорости роста бинарных соединений GaAs и InAs рассчитаны по эмпирическим формулам:

$$v_{GaAs} = \frac{A_{GaAs}T^3 + B_{GaAs}T^2 + C_{GaAs}T + D_{GaAs}}{60}, \quad (2)$$

$$v_{InAs} = \frac{A_{InAs}T^3 + B_{InAs}T^2 + C_{InAs}T + D_{InAs}}{60}, \quad (3)$$

где  $T$  — температура подложки; коэффициенты имеют значения:  $A_{GaAs} = 1,333 \cdot 10^{-6}$ ,  $B_{GaAs} = -4,5714 \cdot 10^{-3}$ ,  $C_{GaAs} = 4,928$ ,  $D_{GaAs} = -1679$ ,  $A_{InAs} = 5,705 \cdot 10^{-6}$ ,  $B_{InAs} = -14,835 \cdot 10^{-3}$ ,  $C_{InAs} = 12,75$ ,  $D_{InAs} = -3608$ .

Расчеты по формулам (1)–(3) показали, что при температуре подложки 1123 К, толщине газовой зоны 1000 мкм, температурном градиенте 30 К/см и значениях концентраций индия  $x_{In}$  в источнике 10 %, 20 % и

30 %, значения скоростей осаждения пленок  $Ga_{1-x}In_xAs$  составили 0,3245, 0,3191 и 0,3137 мкм/мин. Результаты расчетов показали, что повышение концентрации индия приводит к понижению скорости роста, а увеличение температурного градиента — к его повышению.

Цель работы — изучение влияния состава источника при заданной величине температурного градиента на кинетические особенности роста пленок твердого раствора  $Ga_{1-x}In_xAs$  на подложке GaAs в поле температурного градиента через тонкую газовую зону.

## Материалы и методы

Для проведения процесса массопереноса использована специально разработанная графитовая кассета (рис. 1), состоящая из двух графитовых дисков. В углубление нижнего диска помещался прессованный порошок с необходимым составом  $Ga_{1-x}In_xAs$ . Подложка GaAs размещалась на держатель, установленный в отверстии верхнего диска. Далее оба диска скреплялись с помощью зажима. Затем кассета помещалась в вакуумную камеру в подготовленный тепловой узел. Откачка воздуха производилась с помощью форвакуумного насоса до  $5 \cdot 10^{-2}$  Па. После чего в камеру напускалась смесь газов — азота ( $N_2$ ) и водорода ( $H_2$ ) в пропорции 1:1. Пары воды, которые всегда имеются в водороде в малых количествах, являются переносчиком атомов  $Ga_{1-x}In_xAs$  из источника на подложку [15].

Выращивание слоев выполнено на подложке GaAs при температуре 1123 К и температурном градиенте  $G = 30$  К/см. Контроль значения температурного градиента проводился при помощи высокоточного регулятора температуры с цифро-аналоговым преобразователем с точностью поддержания температуры до  $\pm 0,1$  °С. Термодпары марки ТХА-0196-01 располагались между верхней и нижней частями графитовой кассеты и в определенном месте эквивалентном температурным условиям, что и подложка с источником в графитовой кассете. Между верхней и нижней термодпарами расстояние составило 10 мм. В качестве подложек использованы пластины GaAs АГЧО (111), толщина газовой зоны между источником и подложкой — 1 мм. Источником

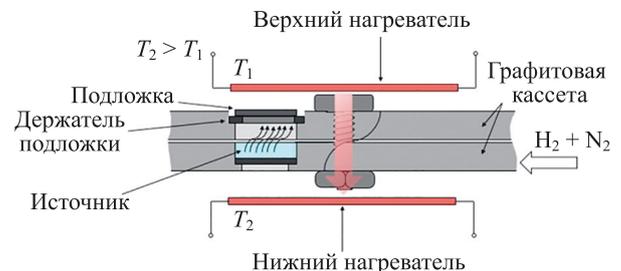


Рис. 1. Схематическое изображение процесса массопереноса в методе близкого переноса через тонкую газовую зону в поле температурного градиента: где  $T_1$  и  $T_2$  — температура верхнего и нижнего нагревателя  
Fig. 1. Schematic representation of the mass transfer process in the method of close transfer through a thin gas zone in the temperature gradient field: where  $T_1$  and  $T_2$  are the temperatures of the upper and lower heaters

ростовых компонентов являлась прессованная гомогенизированная шихта, состоящая из порошков GaAs и InAs с чистотой 99,99 % с различной концентрацией индия:  $x_{\text{In}}$  — 10 %, 20 % и 30 %. Время осаждения слоев составило 10 мин и являлось одинаковым для всех образцов.

Изучение распределения компонент твердого раствора в слое проведено системой энергодисперсионного рентгеноструктурного анализа (EDX) JEOL. Профиль концентрации по глубине вдоль скола пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$  с различной концентрацией  $x_{\text{In}}$  в источнике получен при помощи сканирующего электронного микроскопа MIRA 3 LMH с системой определения элементного состава AZtecEnergy Standard/Xmax 20 (Tescan).

Исследование морфологии поверхностей пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$  выполнено с помощью атомно-силовой микроскопии (АСМ) на сканах  $5 \times 5$  мкм и растровой микроскопии. Измерение толщины пленок на сколах гетероструктур проведено на растровом микроскопе Jeol JSM-6010 LA. Изучение структуры химических связей в твердом растворе  $\text{Ga}_{1-x}\text{In}_x\text{As}$  произведено методом рамановской спектроскопии. Рамановские спектры получены с помощью спектрометра inVia Raman Microscope (Renishaw) гелий-неоновым лазером с длиной волны 514 нм при комнатной температуре.

### Результаты и обсуждение

Изучение кинетики роста проведено путем определения толщины пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$  на подложке GaAs по микрофотографиям сколов, полученных в режиме вторичных электронов (SEI) (рис. 2) и анализа профиля концентрации по глубине вдоль скола пленок (рис. 3).

По микрофотографиям сколов получены следующие значения толщин пленок: 2,98 мкм, 2,61 мкм и 1,31 мкм для концентраций индия  $x_{\text{In}}$  10 %, 20 % и 30 % соответственно. Анализ профилей распределения Ga, As и In по глубине вдоль скола пленок продемонстрировал, что все выращенные пленки имеют неоднородный состав по глубине. По интервалу глубины от поверхности до значения концентрации индия менее 1 ат.% можно косвенно сделать вывод о толщине пленки. Исходя

из этого, толщины пленок при  $x_{\text{In}}$  10 %, 20 % и 30 %, определенные по профилям концентраций, составили: 3,07 мкм, 2,66 мкм и 1,29 мкм. Полученные результаты соответствуют результатам микрофотографии сколов пленок. Метод EDX-анализа носит полуколичественный характер, однако в настоящей работе важно показать, как меняется локальная концентрация индия по толщине пленки.

Концентрация индия для пленок с  $x_{\text{In}}$  — 10 % и 20 % относительно равномерна по всей толщине, а для  $x_{\text{In}}$  равной 30 %, она монотонно убывает с увеличением глубины скола, что свидетельствует о сегрегации высокоподвижных атомов индия в приповерхностную область пленки. Можно предположить, что сегрегация идет по кинетической модели, потому что равновесный предел достигается только при низких скоростях роста пленок и высоких температурах подложки [17]. Индий выходит на поверхность только при обмене с атомами галлия, и в этом случае наблюдается флуктуация галлия и мышьяка по толщине растущей пленки.

Сравнение результатов измерения толщин пленок с результатами расчета их скорости роста показало значительное расхождение экспериментальных значений скорости. Данный результат связан с влиянием концентрации паров воды, содержащейся в водороде, которая достигала до 7 ppm и не была технологически устранена.

Проведено исследования морфологии поверхности. На рис. 4 приведены результаты АСМ-сканирования.

Повышение концентрации индия в составе твердого раствора  $\text{Ga}_{1-x}\text{In}_x\text{As}$  приводит к существенному изменению морфологии поверхности. Чем выше концентрация индия, тем выше перепад высот поверхности и более явно проявляется зеренная структура пленок. Выполнен анализ среднеквадратической шероховатости ( $Rq$ ) для выращенных пленок. В качестве сравнения была измерена  $Rq$  подложки GaAs, которая составила 10,11 нм. Для пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$  получены следующие результаты: для концентрации индия:  $x_{\text{In}}$  10 %, 20 %, 30 %:  $Rq = 9,1; 17,9; 24,2$  нм. Таким образом, наиболее однородные пленки получены в условиях низкой концентрации индия. Сопоставление результатов анализа профиля концентрации по глубине вдоль скола пленок

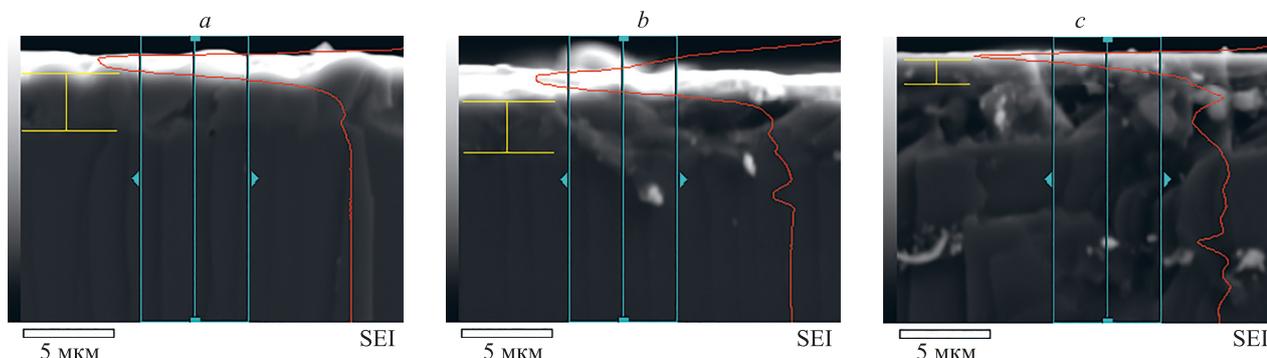


Рис. 2. Микрофотографии сколов пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$ , выращенных при  $T = 1123$  К и  $G = 30$  К/см для концентраций индия  $x_{\text{In}}$ : 10 % (a), 20 % (b) и 30 % (c) (толщина пленок ограничена желтыми линиями)

Fig. 2. Micrographs of the cleavage of  $\text{Ga}_{1-x}\text{In}_x\text{As}$  films grown at  $T = 1123$  K and  $G = 30$  K/cm for indium concentrations  $x_{\text{In}}$ : 10 % (a), 20 % (b) и 30 % (c) (film thickness is limited by yellow lines)

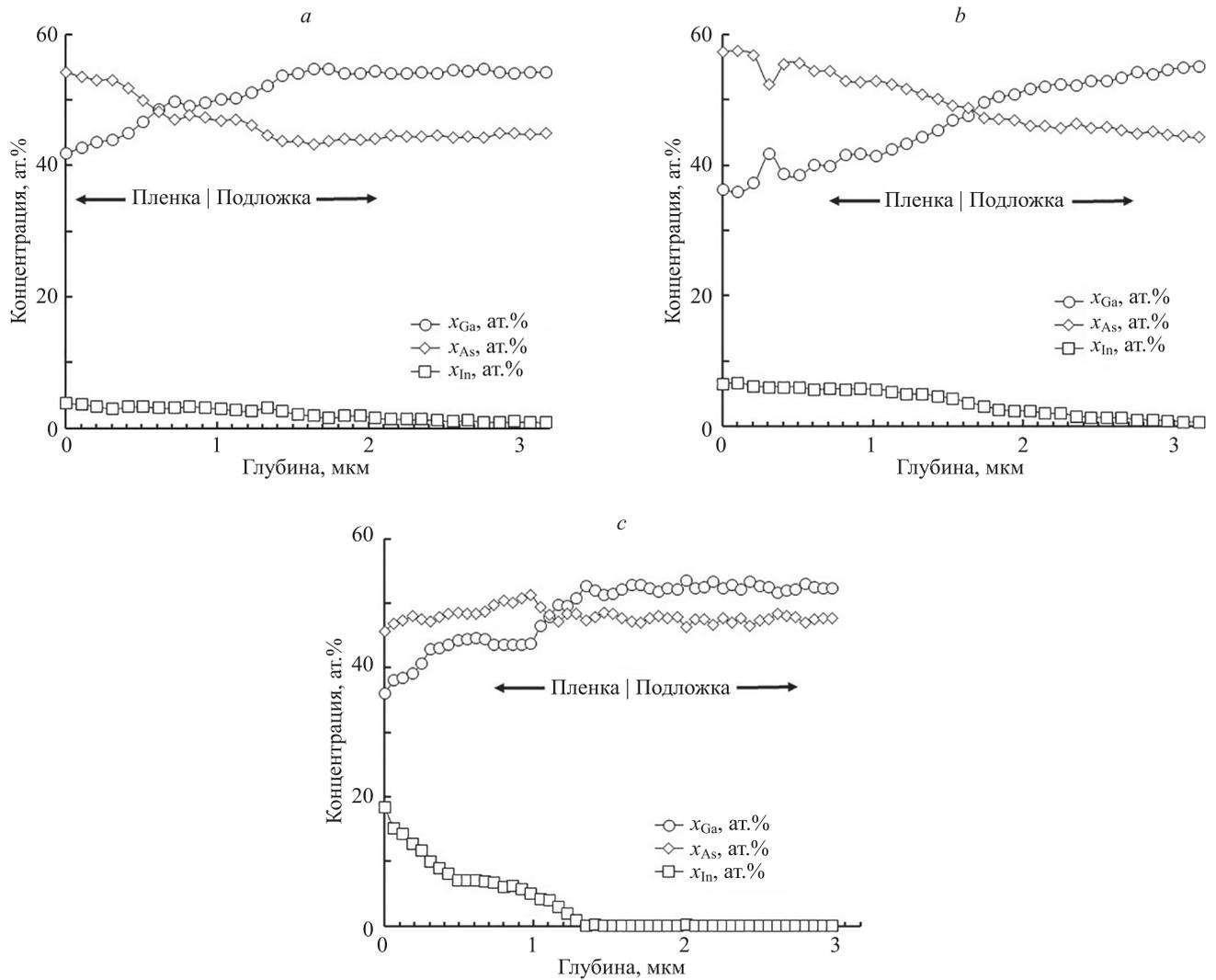


Рис. 3. Профиль концентрации по глубине вдоль скола пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$ , выращенных при  $T = 1123 \text{ K}$  и  $G = 30 \text{ K/cm}$  для концентраций индия  $x_{\text{In}}$ : 10 % (a), 20 % (b) и 30 % (c)

Fig. 3. Concentration depth profile along the film cleavage of  $\text{Ga}_{1-x}\text{In}_x\text{As}$  films grown at  $T = 1123 \text{ K}$  and  $G = 30 \text{ K/cm}$ : for indium concentrations  $x_{\text{In}}$ : 10 % (a), 20 % (b), 30 % (c)

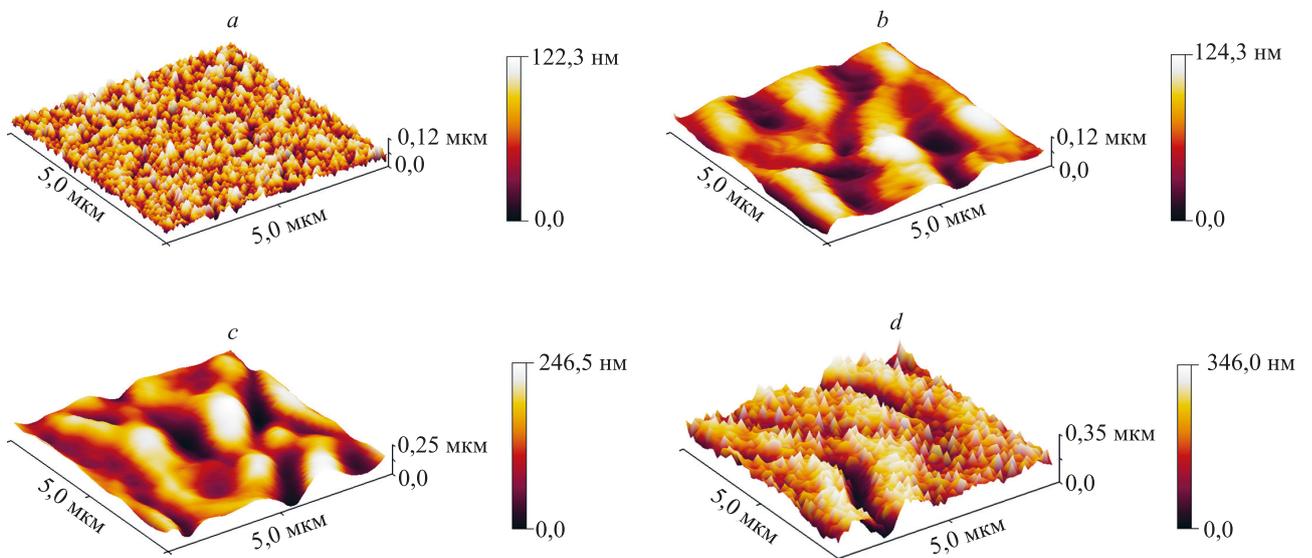


Рис. 4. АСМ-изображения морфологии поверхности  $\text{Ga}_{1-x}\text{In}_x\text{As}$ :  $x_{\text{In}} = 0 \%$  (a),  $x_{\text{In}} = 10 \%$  (b),  $x_{\text{In}} = 20 \%$  (c),  $x_{\text{In}} = 30 \%$  (d)

Fig. 4. AFM images of  $\text{Ga}_{1-x}\text{In}_x\text{As}$  surface morphology:  $x_{\text{In}} = 0 \%$  (a),  $x_{\text{In}} = 10 \%$  (b),  $x_{\text{In}} = 20 \%$  (c),  $x_{\text{In}} = 30 \%$  (d)

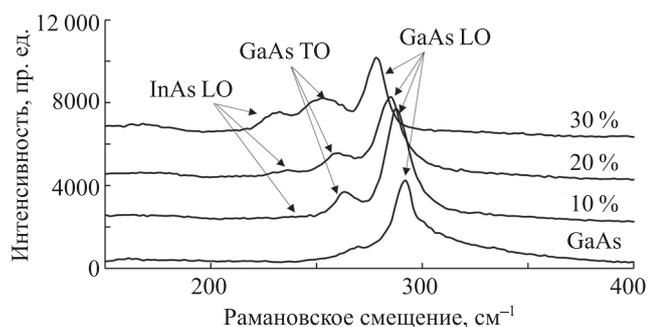


Рис. 5. Рамановские спектры пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$  при  $x_{\text{In}}$ : 10 %, 20 %, 30 % и подложки GaAs

Fig. 5. Raman spectra of  $\text{Ga}_{1-x}\text{In}_x\text{As}$  films at  $x_{\text{In}}$ : 10 %, 20 %, 30 % and GaAs substrates

и атомно-силовой спектроскопии свидетельствует о сильном влиянии индия в составе источника на их скорость роста при  $G = 30$  К/см и  $T = 1123$  К.

Результаты исследования пленок методом комбинационного рассеяния света представлены на рис. 5. В соответствии с правилом отбора для кристаллической решетки цинковой обманки можно индентифицировать продольные (LO) и поперечные (TO) фононные моды GaAs и LO-моду InAs. Из спектров видно, что с увеличением концентрации In в пленке наблюдается существенное смещение частоты LO- и TO-фононных мод GaAs влево на 13 и 16  $\text{cm}^{-1}$ . Это свидетельствует прежде всего о сильном влиянии упругих механических напряжений,

вызванных рассогласованием постоянных решеток пленки  $\text{Ga}_{1-x}\text{In}_x\text{As}$  и подложки GaAs [18].

Сегрегация индия (рис. 2) приводит к значительному смещению мод GaAs и InAs в выращенных пленках. Тем не менее, из рамановских исследований можно сделать вывод, что твердый раствор  $\text{Ga}_{1-x}\text{In}_x\text{As}$  синтезировался, но имеет ближний порядок химических связей. Результаты рамановской спектроскопии хорошо согласуются с результатами EDX-картирования и АСМ.

## Заключение

Изучены особенности выращивания твердых растворов  $\text{Ga}_{1-x}\text{In}_x\text{As}$  на подложке GaAs в поле температурного градиента через тонкую газовую зону. Выращены пленки с толщинами от 2 до 3,2 мкм и исследована их кинетика роста, морфология поверхности и структура химических связей. Показано существенное влияние индия в составе ростового источника на свойства пленок  $\text{Ga}_{1-x}\text{In}_x\text{As}$ . Обнаружено, что при концентрации индия  $x_{\text{In}}$  более 20 % проявляется существенная сегрегация индия в слое  $\text{Ga}_{1-x}\text{In}_x\text{As}$  в направлении роста. Этот фактор также оказывает влияние на кинетику, снижая скорость роста, и приводит к увеличению шероховатости с 9,1 до 24,2 нм для  $x_{\text{In}}$  в диапазоне от 10 % до 30 % соответственно. Выращенные пленки обладают нестехиометрическим составом и ближним порядком химических связей элементов твердого раствора  $\text{Ga}_{1-x}\text{In}_x\text{As}$ .

## Литература

- Zimmer M., Trachtmann A., Jetter M., Michler P. MOVPE grown InGaAs quantum dots with emission near 1.3  $\mu\text{m}$  for electrically driven single-photon sources // *Journal of Crystal Growth*. 2023. V. 605. P. 127081. <https://doi.org/10.1016/j.jcrysgro.2023.127081>
- Ozturk O., Ozturk E., Elagoz S. Linear and nonlinear optical absorption coefficient and electronic features of triple GaAlAs/GaAs and GaInAs/GaAs quantum wells depending on barrier widths // *Optik*. 2019. V. 180. P. 394–405. <https://doi.org/10.1016/j.ijleo.2018.11.091>
- Paulauskas T., Pačebutas V., Geižutis A., Kamarauskas M., Drazdys M., Rudzikas M., Kondrotas R., Naujokaitis A., Nevinskas I., Šebeka B., Strazdienė V., Krotkus A. Performance analysis of GaAsBi/InGaAs heterostructure for III-V multi-junction solar cells // *Solar Energy Materials and Solar Cells*. 2022. V. 248. P. 112013. <https://doi.org/10.1016/j.solmat.2022.112013>
- Nishiyama S., Takenaka C., Kusunoki T., Fujii T., Nakajima K. LPE growth of InP and InGaAs on MQW layers below 500°C // *Journal of Crystal Growth*. 1990. V. 104. N 4. P. 809–814. [https://doi.org/10.1016/0022-0248\(90\)90106-U](https://doi.org/10.1016/0022-0248(90)90106-U)
- Bahari A., Salieneh M.G., Biranvand N. Structural nonlinear effects in  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{GaAs}$  heterostructure bipolar transistor lasers // *Optik*. 2015. V. 126. N 24. P. 5249–5252. <https://doi.org/10.1016/j.ijleo.2015.08.269>
- Li Y., Yan X., Zhang X., Wu Ch., Zheng J., Zha Ch., Fu T., Gong L., Ren X. Low-threshold miniaturized core-shell GaAs/InGaAs nanowire/quantum-dot hybrid structure nanolasers // *Optics and Laser Technology*. 2022. V. 152. P. 108150. <https://doi.org/10.1016/j.optlastec.2022.108150>
- Chen X., Xiao Y., Cheng Y., Zhang Z., Gou Y., Wang J. MOCVD growth and thermal stability analysis of 1.2  $\mu\text{m}$  InGaAs/GaAs multi quantum well structure // *Journal of Alloys and Compounds*. 2022. V. 922. P. 166173 <https://doi.org/10.1016/j.jallcom.2022.166173>
- Mintairov S.A., Emelyanov V.M., Rybalchenko D.V., Saliu R.A., Timoshina N.K., Shvarts M.Z., Kalyuzhnyy N.A. Heterostructures of

## References

- Zimmer M., Trachtmann A., Jetter M., Michler P. MOVPE grown InGaAs quantum dots with emission near 1.3  $\mu\text{m}$  for electrically driven single-photon sources. *Journal of Crystal Growth*, 2023, vol. 605, pp. 127081. <https://doi.org/10.1016/j.jcrysgro.2023.127081>
- Ozturk O., Ozturk E., Elagoz S. Linear and nonlinear optical absorption coefficient and electronic features of triple GaAlAs/GaAs and GaInAs/GaAs quantum wells depending on barrier widths. *Optik*, 2019, vol. 180, pp. 394–405. <https://doi.org/10.1016/j.ijleo.2018.11.091>
- Paulauskas T., Pačebutas V., Geižutis A., Kamarauskas M., Drazdys M., Rudzikas M., Kondrotas R., Naujokaitis A., Nevinskas I., Šebeka B., Strazdienė V., Krotkus A. Performance analysis of GaAsBi/InGaAs heterostructure for III-V multi-junction solar cells. *Solar Energy Materials and Solar Cells*, 2022, vol. 248, pp. 112013. <https://doi.org/10.1016/j.solmat.2022.112013>
- Nishiyama S., Takenaka C., Kusunoki T., Fujii T., Nakajima K. LPE growth of InP and InGaAs on MQW layers below 500°C. *Journal of Crystal Growth*, 1990, vol. 104, no. 4, pp. 809–814. [https://doi.org/10.1016/0022-0248\(90\)90106-U](https://doi.org/10.1016/0022-0248(90)90106-U)
- Bahari A., Salieneh M.G., Biranvand N. Structural nonlinear effects in  $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{GaAs}$  heterostructure bipolar transistor lasers. *Optik*, 2015, vol. 126, no. 24, pp. 5249–5252. <https://doi.org/10.1016/j.ijleo.2015.08.269>
- Li Y., Yan X., Zhang X., Wu Ch., Zheng J., Zha Ch., Fu T., Gong L., Ren X. Low-threshold miniaturized core-shell GaAs/InGaAs nanowire/quantum-dot hybrid structure nanolasers. *Optics and Laser Technology*, 2022, vol. 152, pp. 108150. <https://doi.org/10.1016/j.optlastec.2022.108150>
- Chen X., Xiao Y., Cheng Y., Zhang Z., Gou Y., Wang J. MOCVD growth and thermal stability analysis of 1.2  $\mu\text{m}$  InGaAs/GaAs multi quantum well structure. *Journal of Alloys and Compounds*, 2022, vol. 922, pp. 166173 <https://doi.org/10.1016/j.jallcom.2022.166173>
- Mintairov S.A., Emelyanov V.M., Rybalchenko D.V., Saliu R.A., Timoshina N.K., Shvarts M.Z., Kalyuzhnyy N.A. Heterostructures of

- metamorphic GaInAs photovoltaic converters fabricated by MOCVD on GaAs substrates // *Semiconductors*. 2016. V. 50. N 4. P. 517–522. <https://doi.org/10.1134/S1063782616040163>
9. Devitsky O.V., Zakharov A.A., Lunin L.S., Sysoev I.A., Pashchenko A.S., Vakalov D.S., Chapura O.M. Influence of magnetron sputtering conditions on the structure and surface morphology of  $\text{In}_x\text{Ga}_{1-x}\text{As}$  thin films on a GaAs (100) substrate // *Condensed Matter and Interphases*. 2022. V. 24. N 3. P. 300–305. <https://doi.org/10.17308/kcmf.2022.24/9851>
  10. Lunin L.S., Lunina M.L., Devitsky O.V., Sysoev I.A. Pulsed laser deposition of  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  and GaP thin films onto Si substrates for photoelectric converters // *Semiconductors*. 2017. V. 51. N 3. P. 387–391. <https://doi.org/10.1134/S1063782617030174>
  11. Сысоев И.А. Эпитаксия твердых растворов соединений АІІВVI с микро- и наноструктурой в поле температурного градиента: диссертация на соискание ученой степени доктора технических наук. Ставрополь, 2010.
  12. Gorbachev A., De Anda Salazar F., Galván Montalvo J.A., Michourmyi V. Peculiarities of the thermodynamic conditions to grow InGaAs epitaxial layers by LPE on GaAs substrate at low temperatures // *MRS Advances*. 2021. V. 6. N 46. P. 1005–1009. <https://doi.org/10.1557/s43580-021-00198-8>
  13. Лозовский В.Н., Лунин Л.С., Попов В.П. Зонная перекристаллизация градиентом температуры полупроводниковых материалов. М.: Металлургия, 1987. 232 с.
  14. Сысоев И.А., Лунин Л.С. Градиентная эпитаксия для получения микро- и наноструктур твердых растворов АІІВV через тонкую газовую фазу. Ставрополь: Изд-во СКФУ, 2015. 94 с.
  15. Куликов И.С. Термодинамика оксидов. М.: Металлургия, 1986. 344 с.
  16. Dehaese O., Wallart X., Mollot F. Kinetic model of element III segregation during molecular beam epitaxy of III-III'-V semiconductor compounds // *Applied Physics Letters*. 1995. V. 66. N 1. P. 52–54. <https://doi.org/10.1063/1.114180>
  17. Zheng Y.-J., Lam A.M., Engstrom J.R. Modeling of Ge surface segregation in vapor-phase deposited  $\text{Si}_{1-x}\text{Ge}_x$  thin films // *Applied Physics Letters*. 1999. V. 75. N 6. P. 817–819. <https://doi.org/10.1063/1.124523>
  18. Pashchenko A.S., Devitsky O.V., Lunin L.S., Kasyanov I.V., Nikulin D.A., Pashchenko O.S. Structure and morphology of GaInAsP solid solutions on GaAs substrates grown by pulsed laser deposition // *Thin Solid Films*. 2022. V. 743. P. 139064. <https://doi.org/10.1016/j.tsf.2021.139064>
  - metamorphic GaInAs photovoltaic converters fabricated by MOCVD on GaAs substrates. *Semiconductors*, 2016, vol. 50, no. 4, pp. 517–522. <https://doi.org/10.1134/S1063782616040163>
  9. Devitsky O.V., Zakharov A.A., Lunin L.S., Sysoev I.A., Pashchenko A.S., Vakalov D.S., Chapura O.M. Influence of magnetron sputtering conditions on the structure and surface morphology of  $\text{In}_x\text{Ga}_{1-x}\text{As}$  thin films on a GaAs (100) substrate. *Condensed Matter and Interphases*, 2022, vol. 24, no. 3, pp. 300–305. <https://doi.org/10.17308/kcmf.2022.24/9851>
  10. Lunin L.S., Lunina M.L., Devitsky O.V., Sysoev I.A. Pulsed laser deposition of  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  and GaP thin films onto Si substrates for photoelectric converters. *Semiconductors*, 2017, vol. 51, no. 3, pp. 387–391. <https://doi.org/10.1134/S1063782617030174>
  11. Sysoev I.A. *Epitaxy of Solid Solutions of A3B5 Compounds with Micro and Nanostructures in the Field of a Temperature Gradient*. Thesis for the degree of Doctor of Technical Sciences. Stavropol, 2010. (in Russian)
  12. Gorbachev A., De Anda Salazar F., Galván Montalvo J.A., Michourmyi V. Peculiarities of the thermodynamic conditions to grow InGaAs epitaxial layers by LPE on GaAs substrate at low temperatures. *MRS Advances*, 2021, vol. 6, no. 46, pp. 1005–1009. <https://doi.org/10.1557/s43580-021-00198-8>
  13. Lozovsky V.N., Lunin L.S., Popov V.P. *Zone Recrystallization of Semiconductor Materials with a Temperature Gradient*. Moscow, Metallurgy Publ., 1987, 232 p. (in Russian)
  14. Sysoev I.A., Lunin L.S. *Gradient Epitaxy for Obtaining Micro- and Nanostructures of III–V Solid Solutions Through a thin Gas Phase*. Stavropol, NCFU Publ., 2015, 94 p. (in Russian)
  15. Kulikov I.S. *Thermodynamics of Oxides*. Moscow, Metallurgy Publ., 1986, 344 p. (in Russian)
  16. Dehaese O., Wallart X., Mollot F. Kinetic model of element III segregation during molecular beam epitaxy of III-III'-V semiconductor compounds. *Applied Physics Letters*, 1995, vol. 66, no. 1, pp. 52–54. <https://doi.org/10.1063/1.114180>
  17. Zheng Y.-J., Lam A.M., Engstrom J.R. Modeling of Ge surface segregation in vapor-phase deposited  $\text{Si}_{1-x}\text{Ge}_x$  thin films. *Applied Physics Letters*, 1999, vol. 75, no. 6, pp. 817–819. <https://doi.org/10.1063/1.124523>
  18. Pashchenko A.S., Devitsky O.V., Lunin L.S., Kasyanov I.V., Nikulin D.A., Pashchenko O.S. Structure and morphology of GaInAsP solid solutions on GaAs substrates grown by pulsed laser deposition. *Thin Solid Films*, 2022, vol. 743, pp. 139064. <https://doi.org/10.1016/j.tsf.2021.139064>

#### Авторы

**Девицкий Олег Васильевич** — кандидат технических наук, ведущий научный сотрудник, Федеральный исследовательский центр Южный научный центр Российской академии наук, Ростов-на-Дону, 344006, Российская Федерация; старший научный сотрудник, Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация; [sc 57193670678](https://orcid.org/0000-0003-3153-696X), <https://orcid.org/0000-0003-3153-696X>, [v2517@rambler.ru](mailto:v2517@rambler.ru)

**Лунин Леонид Сергеевич** — доктор физико-математических наук, главный научный сотрудник лаборатории, Федеральный исследовательский центр Южный научный центр Российской академии наук, Ростов-на-Дону, 344006, Российская Федерация; главный научный сотрудник, Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация; профессор, Южно-Российский государственный политехнический университет (НПИ) им. Платова, Новочеркасск, 346428, Российская Федерация, [sc 7004378221](https://orcid.org/0000-0002-5534-9694), <https://orcid.org/0000-0002-5534-9694>, [lunin\\_ls@mail.ru](mailto:lunin_ls@mail.ru)

**Митрофанов Даниил Вячеславович** — младший научный сотрудник, Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация; аспирант, Южно-Российский государственный политехнический университет (НПИ) им. Платова, Новочеркасск, 346428, Российская Федерация, [sc 57226807601](https://orcid.org/0009-0009-2137-4312), <https://orcid.org/0009-0009-2137-4312>, [danik-20@yandex.ru](mailto:danik-20@yandex.ru)

**Сысоев Игорь Александрович** — доктор технических наук, доцент, директор научно-образовательного центра фотовольтаики и нанотехнологии, Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация; [sc 32467535800](https://orcid.org/0000-0001-5415-0782), <https://orcid.org/0000-0001-5415-0782>, [elianpisia@yandex.ru](mailto:elianpisia@yandex.ru)

#### Authors

**Oleg V. Devitsky** — PhD, Leading Researcher, Federal Research Center the Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don, 344006, Russian Federation; Senior Researcher, North-Caucasus Federal University, Stavropol, 355017, Russian Federation, [sc 57193670678](https://orcid.org/0000-0003-3153-696X), <https://orcid.org/0000-0003-3153-696X>, [v2517@rambler.ru](mailto:v2517@rambler.ru)

**Leonid S. Lunin** — D.Sc. (Physics & Mathematics), Chief Researcher, Federal Research Center the Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don, 344006, Russian Federation; Chief Researcher, North-Caucasus Federal University, Stavropol, 355017, Russian Federation; Professor, Platov South-Russian State Polytechnic University (NPI), Novocherkassk, 346428, Russian Federation, [sc 7004378221](https://orcid.org/0000-0002-5534-9694), <https://orcid.org/0000-0002-5534-9694>, [lunin\\_ls@mail.ru](mailto:lunin_ls@mail.ru)

**Daniil V. Mitrofanov** — Junior Researcher, North-Caucasus Federal University, Stavropol, 355017, Russian Federation; PhD Student, Platov South-Russian State Polytechnic University (NPI), Novocherkassk, 346428, Russian Federation, [sc 57226807601](https://orcid.org/0009-0009-2137-4312), <https://orcid.org/0009-0009-2137-4312>, [danik-20@yandex.ru](mailto:danik-20@yandex.ru)

**Igor A. Sysoev** — D.Sc., Associate Professor, Director of the Scientific and Educational Center for Photovoltaics and Nanotechnology, North-Caucasus Federal University, Stavropol, 355017, Russian Federation, [sc 32467535800](https://orcid.org/0000-0001-5415-0782), <https://orcid.org/0000-0001-5415-0782>, [elianpisia@yandex.ru](mailto:elianpisia@yandex.ru)

**Никулин Дмитрий Александрович** — младший научный сотрудник, Федеральный исследовательский центр Южный научный центр Российской академии наук, Ростов-на-Дону, 344006, Российская Федерация; младший научный сотрудник, Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация; аспирант, Южно-Российский государственный политехнический университет (НПИ) им. Платова, Новочеркасск, 346428, Российская Федерация, [sc 57213150940](https://orcid.org/0000-0003-0755-4068), <https://orcid.org/0000-0003-0755-4068>, [dmitnikul@gmail.com](mailto:dmitnikul@gmail.com)

**Чапура Олег Михайлович** — инженер-исследователь, Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация; [sc 57209565002](https://orcid.org/0000-0002-6691-0010), <https://orcid.org/0000-0002-6691-0010>, [chapuro1-7@mail.ru](mailto:chapuro1-7@mail.ru)

**Dmitry A. Nikulin** — Junior Researcher, Federal Research Center the Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don, 344006, Russian Federation; Junior Researcher, North-Caucasus Federal University, Stavropol, 355017, Russian Federation; PhD Student, Platov South-Russian State Polytechnic University (NPI), Novocherkassk, 346428, Russian Federation, [sc 57213150940](https://orcid.org/0000-0003-0755-4068), <https://orcid.org/0000-0003-0755-4068>, [dmitnikul@gmail.com](mailto:dmitnikul@gmail.com)

**Oleg M. Chapura** — Research Engineer, North-Caucasus Federal University, Stavropol, 355017, Russian Federation, [sc 57209565002](https://orcid.org/0000-0002-6691-0010), <https://orcid.org/0000-0002-6691-0010>, [chapuro1-7@mail.ru](mailto:chapuro1-7@mail.ru)

*Статья поступила в редакцию 15.02.2023*

*Одобрена после рецензирования 11.05.2023*

*Принята к печати 16.07.2023*

*Received 15.02.2023*

*Approved after reviewing 11.05.2023*

*Accepted 16.07.2023*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ  
COMPUTER SCIENCE

doi: 10.17586/2226-1494-2023-23-4-711-719

**An enhanced AES-GCM based security protocol for securing the IoT communication****A.B. Feroz Khan<sup>1</sup>✉, S. Kalpana Devi<sup>2</sup>, K. Rama Devi<sup>3</sup>**<sup>1</sup> Syed Hameedha Arts and Science College, Kilakarai, Ramanathapuram, 623806, India<sup>2</sup> RV University, Bengaluru, 560059, India<sup>3</sup> Panimalar Engineering College, Chennai, 600069, India<sup>1</sup> [abferozkhan@gmail.com](mailto:abferozkhan@gmail.com)✉, <https://orcid.org/0000-0002-9395-9493><sup>2</sup> [kalpanasubramaniyan2010@gmail.com](mailto:kalpanasubramaniyan2010@gmail.com), <https://orcid.org/0000-0001-7049-3144><sup>3</sup> [ramadevi.sarav@gmail.com](mailto:ramadevi.sarav@gmail.com), <https://orcid.org/0000-0002-6431-8363>**Abstract**

In the recent years, the devices in Internet of Things (IoT) are growing exponentially due to the emergence of many sophisticated applications. This tremendous growth leads to serious security challenges and the devices of Wireless Sensor Networks should be protected from various attacks. IoT can be configured dynamically without fixed infrastructure and the devices are communicated with one another in an Ad-hoc manner. The work presents the classification of various DDoS attacks in the IoT environment and provides a solution for replay attack. All variations of DDoS attacks are modeled using UML based activity modeling. This clearly understands the behavior of each version of attacks and their performance in the environment. The modeling also helps to construct a solution to prevent this attack from its execution. The work also proposed a trust based protocol for replay attacks which allows the attack inside the network and blocks it after identifying the attack based on its specific behavior. The network performance is improved after implementing this proposed protocol inside the network with help of simulation under realistic conditions. The performance metrics considered in the work are energy, packet loss, computational time and throughput. The paper compares the performance with the state-of-the-art schemes such as Efficient Distributed Deterministic Key and Hash-based Message Authentication Code. The experimental analysis proved that the proposed scheme outperforms the other state-of-the-works in terms of computational cost, throughput, and delay.

**Keywords**

DDoS, security, trusted metrics, IoT

**For citation:** Feroz Khan A.B., Kalpana Devi S., Rama Devi K. An enhanced AES-GCM based security protocol for securing the IoT communication. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 711–719. doi: 10.17586/2226-1494-2023-23-4-711-719

УДК 004.75

**Усовершенствованный протокол безопасности на основе AES-GCM для защиты связи в интернете вещей****А.Б. Фероз Хан<sup>1</sup>✉, Девн С. Калпана<sup>2</sup>, К. Рама Девн<sup>3</sup>**<sup>1</sup> Колледж искусств и науки Сайеда Хамидхи, Раманатхапурам, 623806, Индия<sup>2</sup> Университет RV, Бангалор, 560059, Индия<sup>3</sup> Панимальский инженерный колледж, Ченнаи, 600069, Индия<sup>1</sup> [abferozkhan@gmail.com](mailto:abferozkhan@gmail.com)✉, <https://orcid.org/0000-0002-9395-9493><sup>2</sup> [kalpanasubramaniyan2010@gmail.com](mailto:kalpanasubramaniyan2010@gmail.com), <https://orcid.org/0000-0001-7049-3144><sup>3</sup> [ramadevi.sarav@gmail.com](mailto:ramadevi.sarav@gmail.com), <https://orcid.org/0000-0002-6431-8363>**Аннотация**

В последние годы количество устройств в интернете вещей (IoT) растет в геометрической прогрессии из-за появления множества сложных приложений. Это приводит к серьезным проблемам безопасности. Устройства

© Feroz Khan A.B., Kalpana Devi S., Rama Devi K., 2023

Wireless Sensor Network (WSN) должны быть защищены от различных атак. Интернет вещей можно настраивать динамически без фиксированной инфраструктуры, а устройства WSN взаимодействуют друг с другом в режиме прямого беспроводного динамического соединения (Ad-hoc). Представлена классификация различных Distributed Denial of Service (DDoS)-атак в среде интернета вещей и предложено решение для предотвращения повторной атаки. Выполнено моделирование DDoS-атак с использованием UML-диаграмм активности, что дает четкое понимание поведения каждой версии атаки и их производительности в среде. В результате моделирования построено решение, предотвращающее выполнение атак. Предложен протокол, основанный на доверии, для анализа поведения повторных атак, который допускает атаку внутри сети и блокирует ее после идентификации. Проведено моделирование в реальных условиях для улучшения производительности сети. Рассмотрены показатели производительности сети: энергия, потеря пакетов, время вычислений и пропускная способность. Проведено сравнение производительности сети предложенного решения с современными схемами, такими как EDDK и HMAC. Экспериментальный анализ показал, что предложенный протокол превосходит схемы EDDK и HMAC по параметрам вычислительных затрат, пропускной способности и задержки.

**Ключевые слова**

DDoS, безопасность, надежные показатели, интернет вещей

**Ссылка для цитирования:** Фероз Хан А.Б., Калпана Деви С., Рама Деви К. Усовершенствованный протокол безопасности на основе AES-GCM для защиты связи в интернете вещей // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 711–719 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-4-711-719

**Introduction**

The Internet of Things (IoT) is a vast network of interconnected smart devices that can be found all over the world. These devices are equipped with sensors that allow them to measure their surroundings and communicate with other devices over the internet [1, 2]. The IoT is a network of connected smart devices that can transfer data to one another. These devices include sensors, actuators, Radio Frequency IDentification and Wireless Sensor Network (WSN) components, and other technologies that make life easier in various fields, such as healthcare, agriculture, education, transportation, and more. However, the increase in IoT applications has also led to security challenges that need to be addressed.

Protecting the privacy of users is crucial since sensitive and important data is often transmitted through these smart applications. While there are currently available security solutions, they fall short in fully meeting all the security

requirements of smart applications. These solutions tend to focus only on functionality requirements, and do not integrate well with the entire system. Therefore, there is a need for integrated security solutions that comply with different standards and technologies to ensure the security of smart applications [3, 4]. Fig. 1 depicts the architecture of IoT. The protection of communication channels is a primary security requirement in safeguarding of IoT devices, given that physical attacks pose critical threats to sensor nodes deployed in various environments. To address the significant harm that Distributed Denial of Service (DDoS) attacks can cause in IoT networks, this study presents a proposed solution. Multiple works have emerged to prevent DDoS attacks in IoT, which typically involve an attacker turning a normal node into a “zombie” and spreading the attack to other nodes in the network. The zombie node then interacts with other nodes in the cluster as if they were normal, even allowing it to share confidential information without the network knowledge [5].

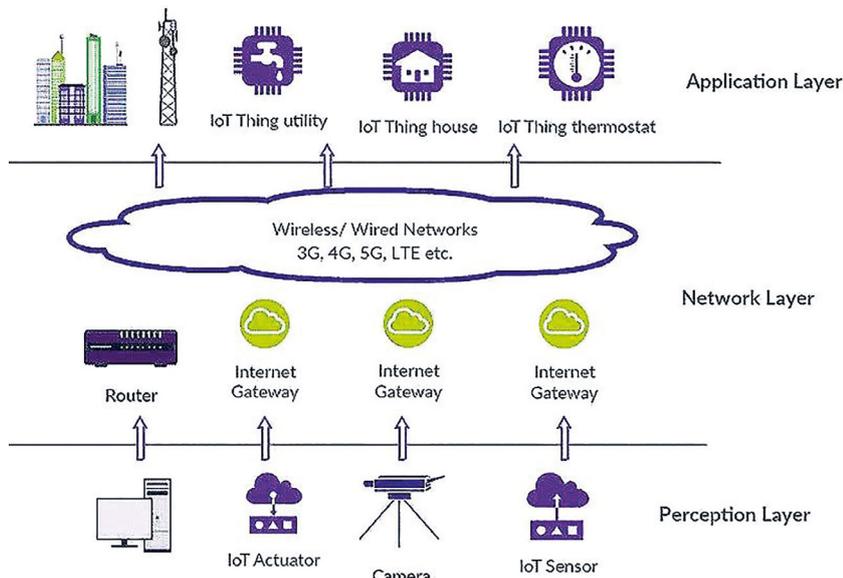


Fig. 1. Overview of IoT communication

With the continued increase of IoT devices, strong security policies must be implemented to ensure confidentiality, integrity, and availability — the three core principles of security. Collaboration between governments and industry players is crucial to establish these policies which should prioritize end-to-end security to combat automated and distributed attacks such as DDoS. The security of IoT components as a whole must be taken into account, rather than individual components. Advanced standards that support global interoperability and enhance IoT security must also be developed. The policies must be established with government and industrial collaboration based on the following considerations [6].

- To combat automated and distributed attacks like DDoS, end-to-end security measures must be implemented in IoT systems.
- The security of IoT systems must be viewed holistically, rather than solely focusing on individual components, to ensure that potential vulnerabilities in the overall network are identified and addressed.
- Advanced standards that establish a core baseline are required to support worldwide interoperability and strengthen IoT security.

The rest of the paper is divided into four parts: the literature review of various related works for the detection of node capturing attack, the introduction of the proposed work and its novelty compared to other works, the discussion of the results of the proposed work, and the conclusion.

### Related works

After analyzing various countermeasures available against DDoS attacks, it is observed that reactive, proactive, and mobile agent countermeasures are the most common techniques used for protection. Reactive technique starts its protection mechanism only after encountering an attack, whereas proactive techniques set up a secure channel before the attack is executed to prevent it from happening. Mobile agent techniques are used in sensor nodes to act as a defensive mechanism and save the network from DDoS attacks. Several research works have proposed different methods to detect and prevent jamming attacks in wireless networks.

Verma et al. [7] proposed a detection technique using timing information and signal strength to detect jamming attacks. However, their proposed scheme works well in combination with other countermeasures. Fadele et al. [8] proposed a Countermeasure Detection and Consistency Algorithm technique that uses signal strength and location information to detect reactive jamming attacks. However, their work increases the computational cost and consumes more energy. Jia et al. [9] proposed a stakelberg game approach to address the anti-jamming problem. Their proposed method is well-suited to overcome jamming issues but it increases energy and implementation costs.

In addition to the aforementioned works, Korzhuk V. et al. [10] proposed the “Identification of Attacks against Wireless Sensor Networks Based on Behaviour Analysis”. This paper presents a method for detecting attacks against WSNs by analyzing their behavior. Their approach focuses

on identifying patterns and anomalies in network behavior to identify potential attacks. By leveraging behavior analysis techniques, their method offers a promising approach to detecting attacks in WSN.

Yao [11] proposed a multi-agent method that uses a Bernaola Galvan Segmentation Algorithm (BGSA) algorithm to reduce the effect of external jamming attacks. Gezici et al. [12] proposed an optimum jammer placement problem to detect the presence of jammer nodes, using timing, and location information. Sun et al. [13] proposed a multi-data flow topology scheme for mobile jamming attacks, which effectively prevents jamming in the network using routing information in the secret path. Muraleedharan and Osadcw [14] proposed a defensive method that mitigates jamming attacks using receiver operating characteristics. However, their method increases energy and computational costs. Strasser et al. [15] proposed a reactive technique for detecting jamming attacks using different approaches, but their proposed method works best with more than one detection technique. Sasikala and Rengarajan [16] proposed using an artificial bee colony to detect jamming attacks on wireless networks. Their work uses different performance metrics, such as packet loss rate, energy, and Received Signal Strength, to provide an efficient anti-jamming technique. Existing works in the field of DDoS attack detection and mitigation for securing communication in IoT networks have made significant contributions to enhancing the security of IoT devices. These works have focused on developing techniques and protocols to identify and prevent various types of attacks, such as jamming attacks, replay attacks, and other forms of DDoS attacks. Table 1 summarizes the details of the main parameters used in the existing works with their advantages and disadvantages.

From Table 1, we observed that the contributions of existing works have significantly advanced the field of DDoS attack detection and mitigation in IoT networks. These works have improved the security, reliability, and performance of IoT devices, paving the way for the secure deployment and utilization of IoT technologies in various domains. Therefore, all the related works discussed in the literature section perform well in detecting and preventing jamming attacks. However, most of these proposed schemes work well in combination with other detection schemes, and this may result in more overhead on resource-constrained devices in IoT. In the evaluation of our proposed work, it is observed that our scheme efficiently detects and prevents jamming attacks, improving the network performance in terms of energy, delay, computational cost, and the variable number of malicious nodes.

### The proposed protocol

The proposed solution presents a trust-based security protocol that can detect and isolate DDoS attacks effectively while aiming to achieve strong security at a reduced time and cost. The computational process used in the solution is based on a few parameters to improve efficiency. Since the network layer is a critical component of the network responsible for addressing and packet delivery, attackers often target this layer to redirect packets to their desired

Table 1. Summary of related works

Related Work	Main Parameters	Advantages	Disadvantages
Verma et al. [7]	Timing information, signal strength	Effective detection of jamming attacks	Works best in combination with other countermeasures
Fadele et al. [8]	Signal strength, location information	Reactive jamming attack detection	Increased computational cost and energy consumption
Jia et al. [9]	Stakelberg game approach	Anti-jamming solution	Increased energy and implementation costs
Korzhuik et al. [10]	Behavior analysis	Effective technique for identifying attacks based on machine learning algorithms, Study of the dependence of accuracy on confidence level and a priori probability of the normal state	Reliance on a complex machine learning algorithm (Random Forest) may require significant computational resources and training data, potentially limiting its practical deployment in resource-constrained IoT devices or networks.
Yao [11]	BGSA algorithm	Reduction of external jamming attacks	Not suitable for all types of jamming scenarios
Gezici et al. [12]	Timing, location information	Optimum jammer placement detection	Limited to specific scenarios and network conditions
Sun et al. [13]	Multi-data flow topology	Effective prevention of mobile jamming attacks	Increased complexity and configuration overhead
Muraleedharan and Osadciw [14]	Receiver operating characteristics	Mitigation of jamming attacks	Higher energy and computational costs
Strasser et al. [15]	Reactive jamming detection	Wide range of detection approaches	Works best with multiple detection techniques
Sasikala and Rengarajan [16]	Artificial bee colony algorithm	Efficient anti-jamming technique	Performance impact on resource-constrained devices

destination. The proposed solution addresses this issue with a trust-based detection protocol designed to protect IoT devices from various types of DDoS attacks.

**Overview of the proposed protocol**

The proposed protocol aims to authenticate devices before allowing them to exchange data with other devices in the network as in Fig. 2.

In Fig. 2, the variables ID<sub>1</sub>, ID<sub>2</sub> are identifier of the device, Rights<sub>1</sub>, Rights<sub>2</sub> are the access rights of the devices, E<sub>K</sub> refers to encryption key, r refers to random number and H refers to Hash function. The proposed solution introduces a trust-based security protocol designed to effectively detect and isolate DDoS attacks while striving for strong security with reduced time and cost. For better understanding of the Fig. 2, the work is divided into following logical parts:

- The proposed protocol aims to authenticate devices before allowing them to exchange data with other devices in the network. It achieves this by generating a trusted value that serves as a ticket for device communication. The trust value is calculated based on the device ID, a random key, and the access rights. The protocol consists of two phases. In the first phase, the trust value is calculated using the device ID, random number, and access rights. The second phase involves applying an AES\_GCM-based encryption technique to encrypt the trusted metrics. The random number generated in the process helps protect the network from forgery.
- Once the trust value is established, it is hashed and encrypted with a public key. The proposed protocol

also introduces the concept of trustworthiness score T<sub>i</sub>, which is calculated based on the number of successful and failed interactions of a node i with other nodes. A successful interaction increases the trustworthiness score, while a failed interaction decreases it. Direct trust DT<sub>ij</sub> is the trustworthiness score of node j as perceived by node i based on their direct interactions. Indirect trust IT<sub>ij</sub> is the trustworthiness score of node j as perceived by node i based on the recommendations of other nodes in the network. A successful interaction increases the trustworthiness score, while a failed interaction decreases it. The trustworthiness score can be calculated as follows:

$$T_i = \frac{S_i + F_i}{S_i - F_i}$$

where S<sub>i</sub> is the number of successful interactions of node i, and F<sub>i</sub> is the number of failed interactions.

Direct trust DT<sub>ij</sub>: It is the trustworthiness score of node j as perceived by node i based on their direct interactions. The direct trust can be calculated using the following equation:

$$DT_{ij} = T_j$$

Indirect trust IT<sub>ij</sub>: It is the trustworthiness score of node j as perceived by node i based on the recommendations of other nodes in the network. The indirect trust can be calculated using the following equation:

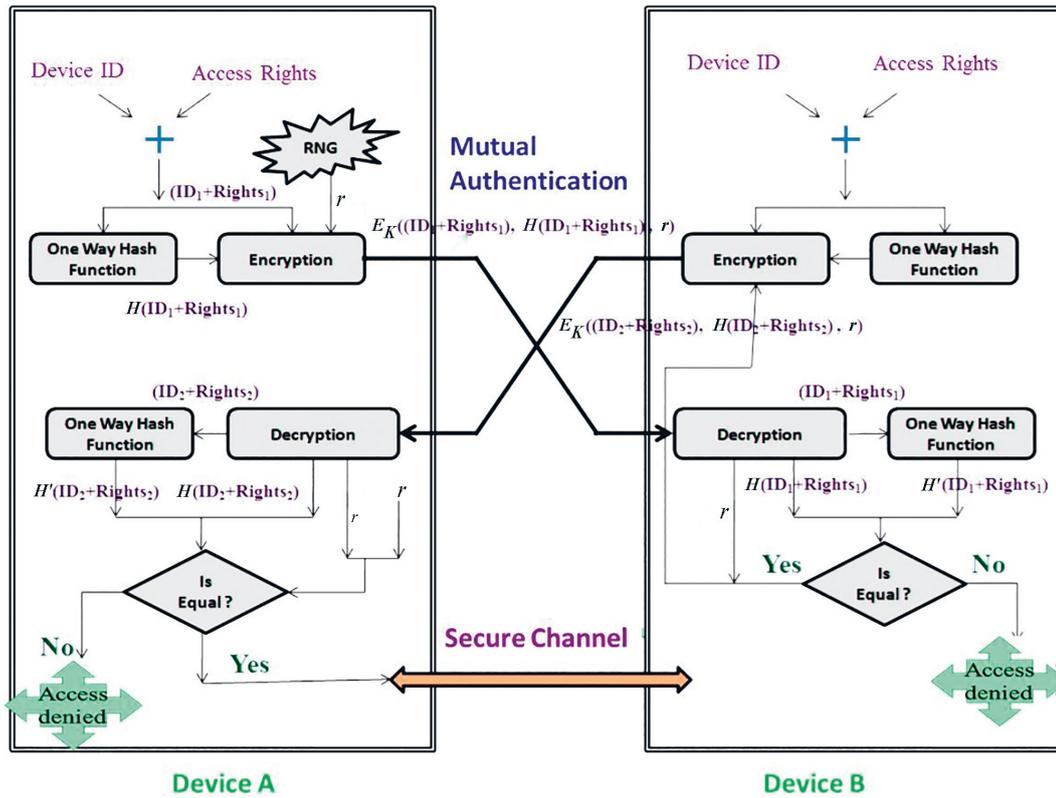


Fig. 2. The Proposed Protocol Architecture

$$IT_{ij} = (1 - w)T_j + w(SUM_k(T_k R_{kj}) / SUM_k(R_{kj})),$$

where  $w$  is a weighting factor that determines the importance of the recommendations,  $R_{kj}$  is the recommendation of node  $k$  for node  $j$ , and the  $SUM$  is taken over all nodes  $k$  that have recommended node  $j$ .

The above equation helps to find a malicious node by calculating its Trusted Value (TV). The TV is based on several metrics, such as packet forwarding ratio, residual energy, and hop count. If a node TV falls below a certain threshold, it is considered untrusted, and its packets are not forwarded in the network. When a node TV falls below the threshold, it is moved to the “suspicious” state. In this state, the node is monitored to see if it continues to behave maliciously. If the node behavior improves and its TV rises above the threshold, it is moved back to the “normal” state. However, if the node behavior remains suspicious or worsens, it is moved to the “attack” state and is removed from the network.

The presented work proposes a trust-based protocol for device authentication in IoT networks. The protocol generates a trust value based on device ID, access rights, and a random key, which is used as a ticket for device communication. The protocol contains two phases: trust value calculation and AES\_GCM-based encryption. The encryption method involved in the protocol ensures authentication and confidentiality of the ticket, and both devices are mutually authenticated before communication. The paper analyzes the various types of attacks possible in each layer of IoT communication and identifies DDoS attacks as the most catastrophic. UML activity modeling is used to understand the behavior of each attack and propose

a solution. Replay attack is identified as the most damaging DDoS attack as it minimizes throughput and maximizes delay, leading to increased energy consumption. The attack is executed using a vulnerable node which becomes a jammer node if successful. The jammer node senses the channel for availability and sends data continuously to jam the network, making detection challenging and leading to performance degradation. While many security solutions have been proposed, they are not efficient for the IoT environment, reducing network performance in terms of energy, delay, and computational cost. The proposed protocol addresses these issues by reducing computational cost, energy consumption, and delay. To combat the various types of DDoS attacks, different countermeasures have been proposed in the literature. Some of the commonly used countermeasures include filtering, rate-limiting, and packet-marking techniques. However, these techniques have certain limitations, such as high overhead, complexity, and reduced network performance. In recent years, trust-based security protocols have gained popularity as a promising approach for protecting IoT devices from DDoS attacks. These protocols ensure mutual authentication of devices before allowing data transmission between them. By establishing a trust value for each device based on factors, such as device ID, random key, and access rights, the protocol ensures that only trusted devices can access the network. The proposed trust-based protocol discussed earlier uses AES\_GCM-based encryption technique to protect the trust value from attackers. This protocol provides both authentication and confidentiality, making it an efficient solution for securing IoT networks. In addition to trust-based security protocols, UML modeling has also

been used to analyze the different variations of DDoS attacks and develop new solutions to prevent them. By identifying the behavior of different attacks, researchers have been able to propose more effective countermeasures to protect IoT networks. One of the most catastrophic attacks in IoT communication is the DDoS attack [17] which can bring the entire network down. Replay attack, which minimizes throughput and maximizes time delay by rapidly increasing energy consumption, is considered the most important attack among all the variations of DDoS attacks [18, 19]. This attack can severely damage the performance of the network and should be prevented at all costs. To detect replay attacks, various techniques, such as watchdog timers, sequence numbers, and hash chains have been proposed. However, these techniques have limitations, such as high overhead and complexity. In recent years, machine learning techniques, such as clustering and decision trees, have also been explored for detecting and preventing DDoS attacks in IoT networks [20]. Overall, protecting IoT networks from DDoS attacks requires a multi-faceted approach that involves using efficient security protocols, analyzing attack behavior, and implementing effective countermeasures. By adopting these approaches, we can ensure the security and reliability of IoT networks while minimizing the cost and energy consumption.

**Results discussion**

The proposed work was implemented and evaluated through simulations in a real-time scenario. The performance of the system was analyzed in terms of energy consumption, delay, and throughput across different time intervals. The simulations were conducted in two sets; the first set evaluated the system performance under varying traffic intervals, while the second set considered the presence of malicious nodes in the network. To analyze the system performance under different traffic conditions, intervals ranging from 1 s to 10 s were used, where 1 s was considered as fast and 10 s as slow. The simulation parameters used in the evaluation are presented in Table 2.

Fig. 3, *a, b, c* depict the results obtained from the proposed TBC approach, which measures energy, delay, and throughput based on traffic intervals ranging from 1 s

Table 2. Simulation parameters

Parameter	Value
Network Area	1000 m × 1000 m
Number of nodes	100
Radio Range, m	250
Transmission Range, m	100
Bandwidth, Mbps	2
MAC Protocol	IEEE 802.11
Routing Protocol	AODV
Traffic Type	CBR (Constant Bit Rate)
Packet Size, B	512
Simulation Time, s	100
Mobility Model	Random Waypoint
Pause Time, s	30
Traffic Intervals	1 s to 10 s
Attacker Nodes	2 to 20
Performance Metrics	Energy, Delay, and Throughput

to 10 s. The outcomes clearly demonstrate that Threshold Based Countermeasure (TBC) outperforms energy Efficient Distributed Deterministic Key (EDDK) and Hash-based Message Authentication Code (HMAC) in terms of energy consumption during DDoS attacks. This is due to the fact that TBC identifies and eliminates malicious nodes, thereby saving energy that could otherwise be wasted by those nodes. Additionally, TBC yields the lowest delay, as it effectively detects and isolates attacks, enabling other nodes to continue their operations without any unnecessary delays. A single node execution of an attack can result in its spread to multiple nodes in the network causing other nodes to wait longer to access the channel. After implementing TBC, the waiting time for nodes in the network is reduced, resulting in an increase in throughput under non-attack conditions.

Fig. 3, *d, e* depict the results of simulations conducted to evaluate the performance of the proposed solution under realistic conditions with respect to energy and delay considerations. The simulation was conducted by

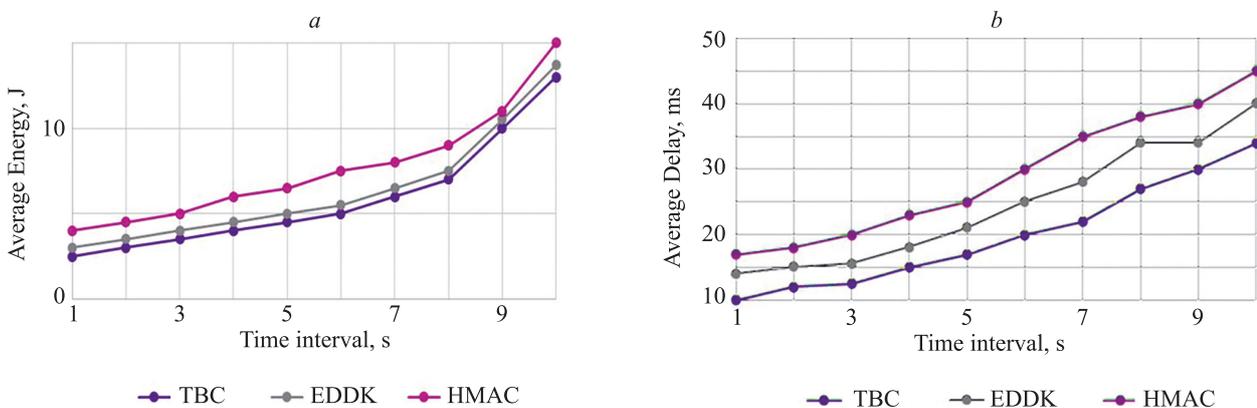


Fig. 3. Performance analysis of the proposed work. Energy Comparative Analyzes: power consumption (a), average delay (b) and throughput (c) of EDDK, HMAC and TBC. Average power consumption by malicious nodes (d), analysis of average latency by malicious nodes (e)

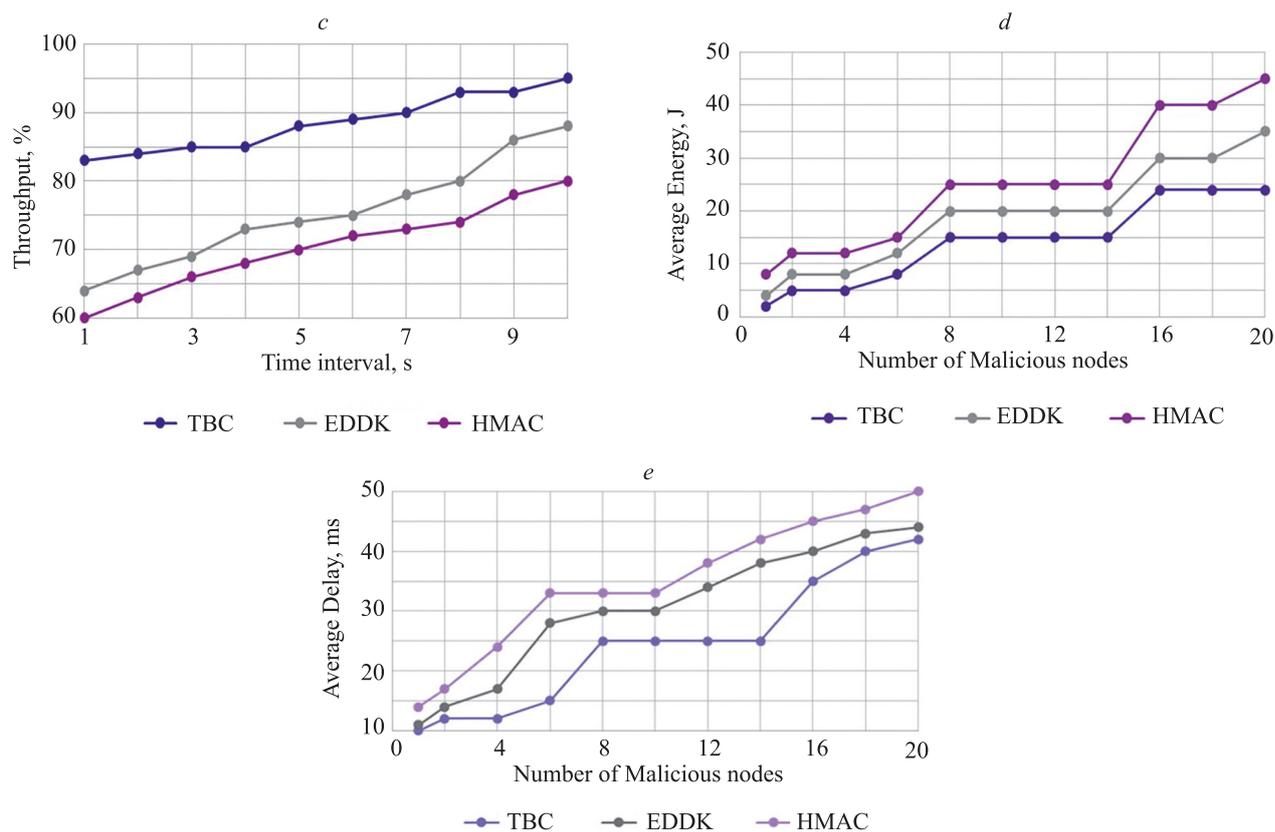


Fig. 3. Performance analysis of the proposed work. Energy Comparative Analyzes: power consumption (a), average delay (b) and throughput (c) of EDDK, HMAC and TBC. Average power consumption by malicious nodes (d), analysis of average latency by malicious nodes (e)

introducing varying numbers of malicious nodes in the network while keeping the total number of nodes constant at 100. The number of attacker nodes ranged from 2 to 20. The improved performance of the proposed solution can be attributed to its ability to effectively detect and remove attacks from the network through its unique features including cross-layer capabilities, reduced false detections, and node cooperation. It was observed that introducing a trade-off between the severity of the attack and its detectability reduces the performance of the optimal strategy, and the performance of the considered optimal model deteriorates as the number of attacks in the network increases. Moreover, the proposed algorithm increases energy consumption only under node mobility conditions. The performance of the work is evaluated in NS-2 which is a widely used discrete event network simulator that allows researchers to simulate and evaluate various network protocols and algorithms. In the context of the proposed work, NS-2 was used to simulate a WSN and evaluate the performance of the proposed TBC algorithm in terms of energy, delay, and throughput under various traffic intervals and different numbers of malicious nodes. The network topology, traffic patterns, and simulation parameters are defined using the NS-2 script. They then ran the simulation to collect performance metrics, such as energy consumption, delay, and throughput, which were used to evaluate the effectiveness of the proposed TBC algorithm. The simulation was conducted in two sets. In the first set, the researchers evaluated the performance of the proposed

TBC algorithm under various traffic intervals ranging from 1 s to 10 s. The simulation parameters were specified in a table, and the results were plotted in Fig. 3, a, b, c. In the second set, the researchers evaluated the performance of the proposed TBC algorithm by varying the number of malicious nodes in the network. The simulation was conducted with 100 nodes, and the number of attacker nodes added in the network ranged from 2 to 20. The simulation results were plotted in Fig. 3, d, e. Overall, NS-2 was used to simulate and evaluate the performance of the proposed TBC algorithm in a controlled environment, which allowed the researchers to analyze the effectiveness of the algorithm under different scenarios and conditions.

## Conclusion

The proposed work introduces a trusted-based protocol to tackle DDoS attacks in the network. The protocol mandates that each node should possess a trusted value before accessing any device in the network. Only trusted nodes are allowed to participate in the communication process. Nodes can be in one of three states: normal, suspicious, or attack. Suspicious and attack states are used to identify potential attacks. If any node is confirmed as an attacker based on trusted metrics, it is immediately removed from the network. The multi-hop path analysis process, which is initiated once a suspicious node is confirmed, is a key feature of the proposed work. The evaluation results demonstrate that the proposed work

outperforms existing methods in terms of energy, delay, and throughput. The proposed protocol has shown a reduction in energy consumption by approximately 15 % compared to existing solutions. The work also revealed a decrease in average delay by approximately 20 % and the increased throughput of 10 % when employing the proposed protocol. By promptly identifying and isolating

DDoS attacks, unnecessary delays in packet delivery are minimized enabling smoother communication among nodes in the network. As future work, we aim to investigate our approach effectiveness against different types of jamming attacks. This would help to expand the scope of the research and demonstrate the protocol potential in a wider range of scenarios.

## References

- Rao N.S.V., Poole S.W., Ma C.Y.T., He F., Zhuang J., Yau D.K.Y. Defense of cyber infrastructures against cyber-physical attacks using game-theoretic models. *Risk Analysis*, 2016, vol. 35, no. 4, pp. 602–616. <https://doi.org/10.1111/risa.12362>
- Wang J., Yang G., Sun Y., Chen S. Defending against Sybil attacks based on received signal strength in wireless sensor networks. *Chinese Journal of Electronics*, 2008, vol. 17, no. 4, pp. 611–614.
- Khan G.A., Anandharaj G. A Cognitive energy efficient and trusted routing model for the security of wireless sensor networks: CEMT. *Wireless Personal Communications*, 2021, vol. 119, no. 4, pp. 3149–3159. <https://doi.org/10.1007/s11277-021-08391-6>
- Meena U., Sharma A. Secure key agreement with rekeying using FLSSO routing protocol in wireless sensor network. *Wireless Personal Communications*, 2018, vol. 101, no. 2, pp. 1177–1199. <https://doi.org/10.1007/s11277-018-5755-9>
- Renold A.P., Ganesh A.B. Energy efficient secure data collection with path-constrained mobile sink in duty-cycled unattended wireless sensor network. *Pervasive and Mobile Computing*, 2019, vol. 55, pp. 1–12. <https://doi.org/10.1016/j.pmcj.2019.02.002>
- Vasudeva A., Sood M. Survey on sybil attack defense mechanisms in wireless ad hoc networks. *Journal of Network and Computer Applications*, 2018, vol. 120, pp. 78–118. <https://doi.org/10.1016/j.jnca.2018.07.006>
- Verma R., Darak S.J., Tikkiwal V., Joshi H., Kumar R. Countermeasures against jamming attack in sensor networks with timing and power constraints. *Proc. of the 11<sup>th</sup> International Conference on Communication Systems & Networks (COMSNETS)*, 2019, pp. 485–488. <https://doi.org/10.1109/comsnets.2019.8711437>
- Fadele A.A., Othman M., Hashem I.A.T., Yaqoob I., Imran M., Shoaib M. A novel countermeasure technique for reactive jamming attack in internet of things. *Multimedia Tools and Applications*, 2019, vol. 78, no. 21, pp. 29899–29920. <https://doi.org/10.1007/s11042-018-6684-z>
- Jia L., Xu Y., Sun Y., Feng S., Anpalagan A. Stackelberg game approaches for anti-jamming defence in wireless networks. *IEEE Wireless Communications*, 2018, vol. 25, no. 6, pp. 120–128. <https://doi.org/10.1109/mwc.2017.1700363>
- Korzhuik V., Groznykh A., Menshikov A., Strecker M. Identification of attacks against wireless sensor networks based on behaviour analysis. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*, 2019, vol. 10, no. 2, pp. 1–21. <https://doi.org/10.22667/JoWUA.2019.06.30.001>
- Yao Y., Xiao B., Wu G., Liu X., Yu Z., Zhang K., Zhou X. Multi-channel based sybil attack detection in vehicular ad hoc networks using RSSI. *IEEE Transactions on Mobile Computing*, 2019, vol. 18, no. 2, pp. 362–375. <https://doi.org/10.1109/tmc.2018.2833849>
- Gezici S., Bayram S., Kurt M.N., Gholami M.R. Optimal jammer placement in wireless localization systems. *IEEE Transactions on Signal Processing*, 2016, vol. 64, no. 17, pp. 4534–4549. <https://doi.org/10.1109/tsp.2016.2552503>
- Sun H., Chen C., Hsu S. Mobile jamming attack and its countermeasure in wireless sensor networks. *Proc. of the 21<sup>st</sup> International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, 2017, pp. 457–462. <https://doi.org/10.1109/ainaw.2007.255>
- Muraleedharan R., Osadciw L.A. Jamming attack detection and countermeasures in wireless sensor network using ant system. *Proceedings of SPIE*, 2006, vol. 6248, pp. 62480G. <https://doi.org/10.1117/12.666330>
- Strasser M., Danev B., Čapkun S. Detection of reactive jamming in sensor networks. *ACM Transactions on Sensor Networks*, 2010, vol. 7, no. 2, pp. 1–29. <https://doi.org/10.1145/1824766.1824772>

## Литература

- Rao N.S.V., Poole S.W., Ma C.Y.T., He F., Zhuang J., Yau D.K.Y. Defense of cyber infrastructures against cyber-physical attacks using game-theoretic models // *Risk Analysis*. 2016. V. 35. N 4. P. 602–616. <https://doi.org/10.1111/risa.12362>
- Wang J., Yang G., Sun Y., Chen S. Defending against Sybil attacks based on received signal strength in wireless sensor networks // *Chinese Journal of Electronics*. 2008. V. 17. N 4. P. 611–614.
- Khan G.A., Anandharaj G. A Cognitive energy efficient and trusted routing model for the security of wireless sensor networks: CEMT // *Wireless Personal Communications*. 2021. V. 119. N 4. P. 3149–3159. <https://doi.org/10.1007/s11277-021-08391-6>
- Meena U., Sharma A. Secure key agreement with rekeying using FLSSO routing protocol in wireless sensor network // *Wireless Personal Communications*. 2018. V. 101. N 2. P. 1177–1199. <https://doi.org/10.1007/s11277-018-5755-9>
- Renold A.P., Ganesh A.B. Energy efficient secure data collection with path-constrained mobile sink in duty-cycled unattended wireless sensor network // *Pervasive and Mobile Computing*. 2019. V. 55. P. 1–12. <https://doi.org/10.1016/j.pmcj.2019.02.002>
- Vasudeva A., Sood M. Survey on sybil attack defense mechanisms in wireless ad hoc networks // *Journal of Network and Computer Applications*. 2018. V. 120. P. 78–118. <https://doi.org/10.1016/j.jnca.2018.07.006>
- Verma R., Darak S.J., Tikkiwal V., Joshi H., Kumar R. Countermeasures against jamming attack in sensor networks with timing and power constraints // *Proc. of the 11<sup>th</sup> International Conference on Communication Systems & Networks (COMSNETS)*. 2019. P. 485–488. <https://doi.org/10.1109/comsnets.2019.8711437>
- Fadele A.A., Othman M., Hashem I.A.T., Yaqoob I., Imran M., Shoaib M. A novel countermeasure technique for reactive jamming attack in internet of things // *Multimedia Tools and Applications*. 2019. V. 78. N 21. P. 29899–29920. <https://doi.org/10.1007/s11042-018-6684-z>
- Jia L., Xu Y., Sun Y., Feng S., Anpalagan A. Stackelberg game approaches for anti-jamming defence in wireless networks // *IEEE Wireless Communications*. 2018. V. 25. N 6. P. 120–128. <https://doi.org/10.1109/mwc.2017.1700363>
- Korzhuik V., Groznykh A., Menshikov A., Strecker M. Identification of attacks against wireless sensor networks based on behaviour analysis // *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA)*. 2019. V. 10. N 2. P. 1–21. <https://doi.org/10.22667/JoWUA.2019.06.30.001>
- Yao Y., Xiao B., Wu G., Liu X., Yu Z., Zhang K., Zhou X. Multi-channel based sybil attack detection in vehicular ad hoc networks using RSSI // *IEEE Transactions on Mobile Computing*. 2019. V. 18. N 2. P. 362–375. <https://doi.org/10.1109/tmc.2018.2833849>
- Gezici S., Bayram S., Kurt M.N., Gholami M.R. Optimal jammer placement in wireless localization systems // *IEEE Transactions on Signal Processing*. 2016. V. 64. N 17. P. 4534–4549. <https://doi.org/10.1109/tsp.2016.2552503>
- Sun H., Chen C., Hsu S. Mobile jamming attack and its countermeasure in wireless sensor networks // *Proc. of the 21<sup>st</sup> International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*. 2017. P. 457–462. <https://doi.org/10.1109/ainaw.2007.255>
- Muraleedharan R., Osadciw L.A. Jamming attack detection and countermeasures in wireless sensor network using ant system // *Proceedings of SPIE*. 2006. V. 6248. P. 62480G. <https://doi.org/10.1117/12.666330>
- Strasser M., Danev B., Čapkun S. Detection of reactive jamming in sensor networks // *ACM Transactions on Sensor Networks*. 2010. V. 7. N 2. P. 1–29. <https://doi.org/10.1145/1824766.1824772>

16. Sasikala E., Rengarajan N. An intelligent technique to detect jamming attack in wireless sensor networks (WSNs). *International Journal of Fuzzy Systems*, 2015, vol. 17, no. 1, pp. 76–83. <https://doi.org/10.1007/s40815-015-0009-4>
  17. Alaba F.A., Awodele O., Afolabi I. Security challenges in internet of things (IoT) enabled healthcare systems. *Proc. of the 2017 International Conference on Computing Networking and Informatics*. IEEE, 2017, pp. 1–7.
  18. Kumar P.H., AnandhaMala G.S. HMAC-R: Hash-based message authentication code and Rijndael-based multilevel security model for data storage in cloud environment. *Journal of Supercomputing*, 2023, vol. 79, no. 3, pp. 3181–3209. <https://doi.org/10.1007/s11227-022-04714-x>
  19. Zhang X., He J., Wei Q. EDDK: Energy-efficient distributed deterministic key management for wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking*, 2011, no. 1, pp. 765143. <https://doi.org/10.1155/2011/765143>
  20. Chaganti R., Bhushan B., Ravi V. A survey on Blockchain solutions in DDoS attacks mitigation: Techniques, open challenges and future directions. *Computer Communications*, 2023, vol. 197, pp. 96–112. <https://doi.org/10.1016/j.comcom.2022.10.026>
16. Sasikala E., Rengarajan N. An intelligent technique to detect jamming attack in wireless sensor networks (WSNs) // *International Journal of Fuzzy Systems*. 2015. V. 17. N 1. P. 76–83. <https://doi.org/10.1007/s40815-015-0009-4>
  17. Alaba F.A., Awodele O., Afolabi I. Security challenges in internet of things (IoT) enabled healthcare systems // *Proc. of the 2017 International Conference on Computing Networking and Informatics*. IEEE, 2017. P. 1–7.
  18. Kumar P.H., AnandhaMala G.S. HMAC-R: Hash-based message authentication code and Rijndael-based multilevel security model for data storage in cloud environment // *Journal of Supercomputing*. 2023. V. 79. N 3. P. 3181–3209. <https://doi.org/10.1007/s11227-022-04714-x>
  19. Zhang X., He J., Wei Q. EDDK: Energy-efficient distributed deterministic key management for wireless sensor networks // *EURASIP Journal on Wireless Communications and Networking*. 2011. N 1. P. 765143. <https://doi.org/10.1155/2011/765143>
  20. Chaganti R., Bhushan B., Ravi V. A survey on Blockchain solutions in DDoS attacks mitigation: Techniques, open challenges and future directions // *Computer Communications*. 2023. V. 197. P. 96–112. <https://doi.org/10.1016/j.comcom.2022.10.026>

#### Authors

**A.B. Feroz Khan** — PhD, Assistant Professor, Syed Hameedha Arts and Science College, Kilakarai, Ramanathapuram, 623806, India, [sc 57209466258](https://orcid.org/0000-0002-9395-9493), <https://orcid.org/0000-0002-9395-9493>, [abferozkhan@gmail.com](mailto:abferozkhan@gmail.com)

**S. Kalpana Devi** — Assistant Professor, RV University, Bangaluru, 560059, India, [sc 57211952643](https://orcid.org/0000-0001-7049-3144), <https://orcid.org/0000-0001-7049-3144>, [Kalpanasubramaniyan2010@gmail.com](mailto:Kalpanasubramaniyan2010@gmail.com)

**K. Rama Devi** — PhD, Associate Professor, Panimalar Engineering College, Chennai, 600069, India, [sc 56708714800](https://orcid.org/0000-0002-6431-8363), <https://orcid.org/0000-0002-6431-8363>, [ramadevi.sarav@gmail.com](mailto:ramadevi.sarav@gmail.com)

Received 15.11.2022

Approved after reviewing 21.05.2023

Accepted 26.07.2023

#### Авторы

**Фероз Хан А.Б.** — PhD, доцент, Колледж искусств и науки Сайеда Хамидхи, Раманатхапурам, 623806, Индия, [sc 57209466258](https://orcid.org/0000-0002-9395-9493), <https://orcid.org/0000-0002-9395-9493>, [abferozkhan@gmail.com](mailto:abferozkhan@gmail.com)

**Калпана Девии С.** — доцент, Университет RV, Бангалор, 560059, Индия, [sc 57211952643](https://orcid.org/0000-0001-7049-3144), <https://orcid.org/0000-0001-7049-3144>, [Kalpanasubramaniyan2010@gmail.com](mailto:Kalpanasubramaniyan2010@gmail.com)

**Рама Девии К.** — PhD, доцент, Панимальарский инженерный колледж, Ченнаи, 600069, Индия, [sc 56708714800](https://orcid.org/0000-0002-6431-8363), <https://orcid.org/0000-0002-6431-8363>, [ramadevi.sarav@gmail.com](mailto:ramadevi.sarav@gmail.com)

Статья поступила в редакцию 15.11.2022

Одобрена после рецензирования 21.05.2023

Принята к печати 26.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-720-733

УДК 004.056

## Атаки на основе вредоносных возмущений на системы обработки изображений и методы защиты от них

Дмитрий Андреевич Есипов<sup>1</sup>✉, Абдулхамид Яхьяевич Бучаев<sup>2</sup>, Акылжан Керимбай<sup>3</sup>, Яна Владиславовна Пузикова<sup>4</sup>, Семен Кириллович Сайдумаров<sup>5</sup>, Никита Сергеевич Сулименко<sup>6</sup>, Илья Юрьевич Попов<sup>7</sup>, Николай Сергеевич Кармановский<sup>8</sup>

1,2,3,4,5,6,7,8 Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>1</sup> some1else.d.ma@gmail.com✉, <https://orcid.org/0000-0003-4467-5117>

<sup>2</sup> abdulhamid0055@yandex.ru, <https://orcid.org/0009-0001-1058-9125>

<sup>3</sup> akerimbai@itmo.ru, <https://orcid.org/0009-0009-9945-9906>

<sup>4</sup> Yanapuzikova19@ya.ru, <https://orcid.org/0009-0007-7604-3022>

<sup>5</sup> semen.say@ya.ru, <https://orcid.org/0009-0008-0774-9803>

<sup>6</sup> n.s.sulimenko@mail.ru, <https://orcid.org/0009-0007-3218-9249>

<sup>7</sup> ilyapopov27@gmail.com, <https://orcid.org/0000-0002-6407-7934>

<sup>8</sup> karmanov50@mail.ru, <https://orcid.org/0000-0002-0533-9893>

### Аннотация

Системы, реализующие технологии искусственного интеллекта, получили широкое распространение благодаря их эффективности в решении прикладных задач, включая компьютерное зрение. Обработка изображений посредством нейронных сетей применяется в критически важных для безопасности системах. В то же время использование искусственного интеллекта сопряжено с характерными угрозами, к которым относится и нарушение работы моделей машинного обучения. Феномен провокации некорректного отклика нейронной сети посредством внесения визуально незаметных человеку искажений впервые описан и привлёк внимание исследователей в 2013 году. Методы атак на нейронные сети на основе вредоносных возмущений непрерывно совершенствовались, были предложены способы нарушения работы нейронных сетей при обработке различных типов данных и задач целевой модели. Угрозы нарушения функционирования нейронных сетей посредством указанных атак стала значимой проблемой для систем, реализующих технологии искусственного интеллекта. Таким образом, исследование в области противодействия атакам на основе вредоносных возмущений являются весьма актуальными. В данной статье представлено описание актуальных атак, приведен обзор и сравнительный анализ таких атак на системы обработки изображений с использованием искусственного интеллекта. Сформулированы подходы к классификации атак на основе вредоносных возмущений. Рассмотрены методы защиты от подобных атак, выявлены их недостатки. Показаны ограничения применяемых методов защиты, снижающие эффективность противодействия атакам. Предложены подходы по обнаружению и устранению вредоносных возмущений.

### Ключевые слова

искусственный интеллект, искусственная нейронная сеть, обработка изображений, состязательная атака, встраивание бэкдора, вредоносное возмущение, состязательное обучение, защитная дистилляция, сжатие параметров, сертифицированная защита, предобработка данных

**Ссылка для цитирования:** Есипов Д.А., Бучаев А.Я., Керимбай А., Пузикова Я.В., Сайдумаров С.К., Сулименко Н.С., Попов И.Ю., Кармановский Н.С. Атаки на основе вредоносных возмущений на системы обработки изображений и методы защиты от них // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 720–733. doi: 10.17586/2226-1494-2023-23-4-720-733

## Attacks based on malicious perturbations on image processing systems and defense methods against them

Dmitry A. Esipov<sup>1</sup>, Abdulhamid Y. Buchaev<sup>2</sup>, Akylzhan Kerimbay<sup>3</sup>, Yana V. Puzikova<sup>4</sup>, Semen K. Saidumarov<sup>5</sup>, Nikita S. Sulimenko<sup>6</sup>, Ilya Yu. Popov<sup>7</sup>, Nikolay S. Karmanovskiy<sup>8</sup>

<sup>1,2,3,4,5,6,7,8</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>1</sup> some1else.d.ma@gmail.com, <https://orcid.org/0000-0003-4467-5117>

<sup>2</sup> abdulhamid0055@yandex.ru, <https://orcid.org/0009-0001-1058-9125>

<sup>3</sup> akerimbai@itmo.ru, <https://orcid.org/0009-0009-9945-9906>

<sup>4</sup> Yanapuzikova19@ya.ru, <https://orcid.org/0009-0007-7604-3022>

<sup>5</sup> semen.say@ya.ru, <https://orcid.org/0009-0008-0774-9803>

<sup>6</sup> n.s.sulimenko@mail.ru, <https://orcid.org/0009-0007-3218-9249>

<sup>7</sup> ilyapopov27@gmail.com, <https://orcid.org/0000-0002-6407-7934>

<sup>8</sup> karmanov50@mail.ru, <https://orcid.org/0000-0002-0533-9893>

### Abstract

Systems implementing artificial intelligence technologies have become widespread due to their effectiveness in solving various applied tasks including computer vision. Image processing through neural networks is also used in security-critical systems. At the same time, the use of artificial intelligence is associated with characteristic threats including disruption of machine learning models. The phenomenon of triggering an incorrect neural network response by introducing perturbations that are visually imperceptible to a person was first described and attracted the attention of researchers in 2013. Methods of attacks on neural networks based on malicious perturbations have been continuously improved, ways of disrupting the operation of neural networks in processing various types of data and tasks of the target model have been proposed. The threat of disrupting the functioning of neural networks through these attacks has become a significant problem for systems implementing artificial intelligence technologies. Thus, research in the field of countering attacks based on malicious perturbations is very relevant. This article describes current attacks, provides an overview and comparative analysis of such attacks on image processing systems based on artificial intelligence. Approaches to the classification of attacks based on malicious perturbations are formulated. Defense methods against such attacks are considered, their shortcomings are revealed. The limitations of the applied defense methods that reduce the effectiveness of counteraction to attacks are shown. Approaches and practical measures to detect and eliminate harmful disturbances are proposed.

### Keywords

artificial intelligence, artificial neural network, image processing, adversarial attack, backdoor embedding, adversarial perturbation, adversarial learning, defense distillation, feature squeezing, certified defense, data preprocessing

**For citation:** Esipov D.A., Buchaev A.Y., Kerimbay A., Puzikova Y.V., Saidumarov S.K., Sulimenko N.S., Popov I.Yu., Karmanovskiy N.S. Attacks based on malicious perturbations on image processing systems and defense methods against them. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 720–733 (in Russian). doi: 10.17586/2226-1494-2023-23-4-720-733

### Введение

Системы, реализующие технологии искусственного интеллекта, получили широкое распространение за счет значительной эффективности по сравнению с другими методами при решении множества прикладных задач, в том числе в обработке естественного языка [1], распознавании речи [2], биометрической аутентификации [3–5], медицинской диагностике [6, 7], видеонаблюдении [8], беспилотном транспорте [9–11] и др. Указанные системы нередко используются для обработки изображений в критической информационной инфраструктуре. По прогнозам компании Gartner, количество беспилотных транспортных средств увеличится на 740 тысяч единиц по всему миру<sup>1</sup>.

В то же время использование систем, реализующих технологии искусственного интеллекта, и нейронных

сетей сопряжено с характерными угрозами безопасности<sup>2,3</sup>. Значительную опасность представляют такие угрозы, как нарушение функционирования или обход средств, реализующих технологии искусственного интеллекта, и модификации модели машинного обучения при помощи искажения данных. За 2022 год 30 % кибератак были направлены на обход или кражу моделей нейронных сетей [12]. Такие угрозы могут быть реализованы в виде атак посредством внесения вредоносных возмущений [13–69]. Отметим, что атаки могут быть реализованы вне зависимости от типа обрабатываемых данных [13]. Наибольшее распространение указанные атаки получили для систем обработки изображений и видеопотоков.

Состязательные атаки приводят к некорректному отклику нейронных сетей и, как следствие, нарушению работы системы, реализующей технологии искус-

<sup>1</sup> Rimol M. Gartner Forecasts More Than 740,000 Autonomous-Ready Vehicles to Be Added to Global Market in 2023 [Электронный ресурс]. 2019. Режим доступа: <https://www.gartner.com/en/newsroom/press-releases/2019-11-14-gartner-forecasts-more-than-740000-autonomous-ready-vehicles-to-be-added-to-global-market-in-2023>, свободный (дата обращения: 21.01.2023).

<sup>2</sup> ФСТЭК. Банк данных угроз безопасности информации [Электронный ресурс]. Режим доступа: <https://bdu.fstec.ru/threat>, свободный (дата обращения: 03.03.2023).

<sup>3</sup> MITRE. Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS) [Электронный ресурс]. Режим доступа: <https://atlas.mitre.org/>, свободный (дата обращения: 03.03.2023).

ственного интеллекта, что может повлечь критические последствия в зависимости от области применения: некорректная постановка диагноза, игнорирование дорожных знаков беспилотным автомобилем и т. д. Кроме того, модификация модели машинного обучения может приводить к манипулированию ее откликами, провоцируя желаемое для злоумышленника поведение [14].

В настоящей работе рассмотрены атаки на основе вредоносных возмущений, их формальное описание, подходы к классификации и конкретные методы атаки, а также методы защиты от указанных атак, их достоинства и недостатки. Кроме того, предложены подходы к обнаружению и устранению вредоносных искажений, встраиваемых в изображения при атаке.

**Атаки на основе вредоносных возмущений**

Атаки на основе вредоносных возмущений предполагают внесение искажений во входные данные, приводящих к нарушению функционирования целевой системы. Нормой таких возмущений являются метрики расстояния [13], характеризующие вредоносное возмущение как разницу между оригинальным и искаженным изображениями. Для них характерны следующие аксиомы:

- неотрицательность:  $\|x\| \geq 0$ ;
- тождество неразличимых:  $\|x - y\| = 0 \leftrightarrow x = y$ ;
- симметричность:  $\|x - y\| = \|y - x\|$ ;
- неравенство треугольника:  $\|x - y\| + \|y - z\| > \|x - z\|$ .

На практике активно используются следующие метрики (рис. 1):

- $L_1$  (манхэттенское расстояние):

$$\|x\|_1 = \sum_{i=1}^n |x_i|,$$

где  $x$  — вредоносное возмущение;  $n$  — число пикселей;  $x_i$  —  $i$ -й пиксел;

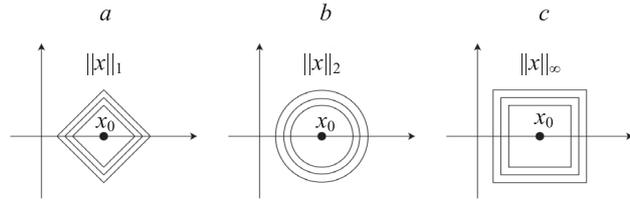


Рис. 1. Метрики расстояния: манхэттенское расстояние  $L_1$  (a); евклидова метрика  $L_2$  (b); расстояние Чебышева  $L_\infty$  (c)

Fig. 1. Distance metrics: Manhattan distance  $L_1$  (a); Euclid distance  $L_2$  (b); Chebyshev distance  $L_\infty$  (c)

- $L_2$  (евклидова метрика):

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2};$$

- $L_\infty$  (расстояние Чебышева):

$$\|x\|_\infty = \max_i |x_i|.$$

В дополнение к перечисленным метрикам также используется метрика  $L_0$ :  $\|x\|_0 = |\{x_i | x_i \neq 0, i = 1, \dots, n\}|$ , показывающая число ненулевых элементов (в случае изображений — модифицированных пикселей). Указанная метрика не соответствует аксиомам, потому формально не является метрикой расстояния и нормой возмущения.  $L_0$  также называют псевдонормой.

**Классификации атак**

Атаки на основе внесения вредоносных возмущений на модели машинного обучения могут быть классифицированы по следующим признакам: преследуемая цель; тип обрабатываемых данных; задача целевой модели; тип искажения; знание атакующего о целевой системе; направленность; число преобразований (рис. 2).

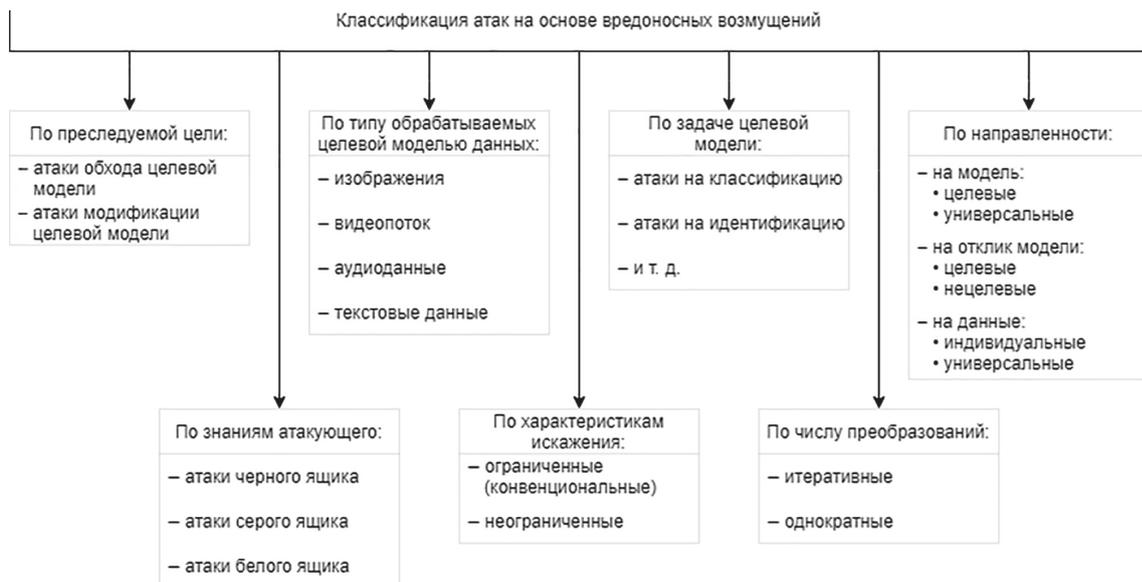


Рис. 2. Классификация атак на основе вредоносных возмущений

Fig. 2. Classification of attacks based on malicious perturbation

Важно отметить, что атаки на основе вредоносных возмущений характерны в первую очередь для искусственных глубоких нейронных сетей (Deep Neural Network, DNN), в то же время, они также могут быть направлены и на другие технологии машинного обучения. В текущей работе будут рассмотрены атаки на нейронные сети.

**Классификация по преследуемой цели.** По преследуемой цели можно выделить следующие типы атак: — атака обхода целевой модели; — атака модификации целевой модели.

Атака обхода (уклонения, evasion) соответствует реализации угрозы УБИ.220 «Угроза нарушения функционирования («обхода») средств, реализующих технологии искусственного интеллекта» согласно банку данных угроз ФСТЭК, тогда как атака модификации целевой модели — УБИ.221 «Угроза модификации модели машинного обучения путем искажения («отравления») обучающих данных»<sup>1</sup>. Аналогичные техники обозначены также и в базе знаний MITRE ATLAS: Evade ML Model и Backdoor ML Model<sup>2</sup> соответственно.

$$N(x) = y; N(x + p) = y'; y \neq y', \quad (1)$$

где  $N$  — нейронная сеть;  $x$  — входные данные;  $y$  — легитимный отклик модели;  $p$  — вредоносное искажение,  $y'$  — нелегитимный отклик модели.

Атаки обхода целевой модели, также известные как состязательные атаки (adversarial attack) [15–65], предполагают модификацию входных данных для изменения отклика модели и поведения системы, реализующей технологии искусственного интеллекта (1). Такие атаки не оказывают влияния на целостность целевой модели.

$$S': \forall x' \in S', x' = x + p_t; L(N_0, S + S') = N', \quad (2)$$

где  $S$  — обучающая выборка без искаженных данных;  $S'$  — набор искаженных данных;  $x'$  — искаженный экземпляр;  $p_t$  — встраиваемый триггер;  $L(\cdot)$  — процедура обучения модели;  $N_0$  — подготовленная к обучению модель;  $N'$  — модифицированная нейронная сеть.

$$N'(x) = y; N'(x + p_t) = y'_t; y \neq y'_t, \quad (3)$$

где  $y'_t$  — некорректный отклик модели, провоцируемый триггером  $p_t$ .

Атаки модификации целевой модели [66–69], основанные на вредоносных возмущениях, предполагают в первую очередь модификацию обучающей выборки целевой нейронной сети посредством модификации или добавления искаженных элементов и обучения модели на этих данных (2). В случае встраивания бэкдора,

вредоносное возмущение представляет собой триггер, провоцирующий конкретный отклик модели.

Дальнейшее использование бэкдора предполагает наложение возмущения, идентичного триггеру, на входные данные (3).

В текущей работе преимущественно будут рассмотрены состязательные атаки.

**Классификация по типу обрабатываемых данных.** Атаки также могут быть классифицированы в зависимости от типа обрабатываемых целевой моделью данных:

- изображения [15–59, 66–70];
- видеопоток [60, 61];
- аудиоданные [62, 63];
- текстовые данные [64, 65].

Следует отметить, что к текстовым данным также относится и сетевой трафик [64]. В текущей работе будут рассмотрены атаки на системы обработки изображений.

**Классификация по задаче целевой модели.** По задаче целевой модели можно выделить следующие типы атак:

- атаки на системы классификации данных [15–56, 60–69];
- сегментация изображения [57, 58];
- повышение разрешающей способности [59];
- и др.

Потенциально, к атакам на основе вредоносных возмущений могут быть уязвимы системы, реализующие технологии искусственного интеллекта, вне зависимости от решаемых задач.

**Классификация по характеристикам вносимого возмущения.** По характеристикам искажения можно выделить следующие типы атак [16]:

- атаки с ограниченным возмущением (restricted perturbation);
- атаки с неограниченным возмущением (unrestricted perturbation).

$$N(x) = y; N(x + p) = y'; y \neq y'; \|p\|_p \leq \eta, \quad (4)$$

где  $\|p\|_p$  — норма вредоносного искажения;  $\eta$  — некоторая константа.

Атаки с ограниченным возмущением предполагают ограничение возмущения некоторым значением метрики расстояния. Такие атаки обхода целевой модели принято называть конвенциональными состязательными атаками (conventional adversarial attack) [16] (4). Ограничение на вносимые возмущения позволяет им оставаться невидимыми или слабо заметными для человека, однако оказывать значительное влияние на отклик нейронной сети [15].

Следует отметить, что вредоносное возмущение может быть ограничено как в атаках обхода [15–50], так и в атаках модификации [66] целевой модели.

Атаки с неограниченным возмущением предполагают произвольную модификацию входных данных вплоть до их подмены [51–56, 67–69]. К указанным искажениям относятся управление цветом [51], поворот изображения [52], управление атрибутами [53], наложение состязательных патчей [52] и не только.

<sup>1</sup> ФСТЭК. Банк данных угроз безопасности информации [Электронный ресурс]. Режим доступа: <https://bdu.fstec.ru/threat>, свободный (дата обращения: 03.03.2023).

<sup>2</sup> MITRE. Adversarial Threat Landscape for Artificial-Intelligence Systems (ATLAS) [Электронный ресурс]. Режим доступа: <https://atlas.mitre.org/>, свободный (дата обращения: 03.03.2023).

К таким атакам в том числе можно отнести и создание дипфейков (DeepFake) [55, 56].

**Классификация по знаниям атакующего о целевой модели.** В зависимости от известной атакующему информации о целевой системе, атаки на основе внесения вредоносных возмущений можно выделить следующие типы атак:

- атаки по модели белого ящика (white-box attack);
- атаки по модели черного ящика (black-box attack).

В случае атак по модели белого ящика [15, 17–32, 34] атакующему доступна вся информация о целевой системе, в том числе архитектура и параметры нейронной сети. Например, атаки, использующие знание гессиана и градиента функции потерь.

Атаки по модели черного ящика [33, 35–50] являются более приближенными к реальным ситуациям, так как атакующий не владеет данными об атакуемой системе. Такие атаки могут быть основаны: на отклике целевой модели [35–44] и на переносе вредоносного возмущения [45–50]. Оба типа атак для генерации искажения используют отклик целевой модели, однако атаки, основанные на переносе, в качестве базиса для генерации используют заранее подготовленное возмущение, в том числе может быть использовано универсальное возмущение [34].

В отдельных случаях выделяют атаки по модели серого ящика (gray-box attack), при которых атакующему известна лишь некоторая часть информации о целевой системе. Такие атаки нередко приравнивают к другим типам по рассматриваемой классификации, чаще — к атакам по модели черного ящика.

**Классификация по направленности.** По направленности можно выделить атаки: целевые (targeted) [15–19, 24, 32, 33, 35–50] и нецелевые [29–31, 34] (nontargeted) по отклику. Целевые по отклику предполагают внесение искажения, провоцирующего конкретный отклик целевой модели, а нецелевые по отклику — провокацию произвольного некорректного отклика.

Также по направленности может быть предложена другая классификация атак: целевые по модели и универсальные (universal) [34]. В этом случае целевые атаки предполагают внесение искажений, провоцирующих некорректный отклик конкретной целевой модели машинного обучения. Универсальные атаки предполагают создание таких искажений, что атака посредством их встраивания позволит вызвать некорректный отклик произвольной модели соответствующего функционала.

Аналогичная классификация может быть предложена и по влиянию вносимого возмущения на входные данные модели: индивидуальные (целевые) [70] по входным данным и универсальные атаки. Первые предполагают нарушение работы целевой системы на конкретном экземпляре входных данных, вторые сохраняют эффект для некоторого множества входных данных.

**Классификация по числу преобразований.** По числу преобразований можно выделить методы генерации искажения за один шаг (one-step method) [18] и итеративные методы (iterative method) [15–17, 21, 23, 29–31, 35–50]. Следует отметить, что созда-

ние вредоносного возмущения за одно преобразование более характерно для атак по модели белого ящика.

### Существующие методы атаки на основе вредоносных возмущений на нейронные сети обработки изображений

Исторически первой состязательной атакой на нейронные сети обработки изображения стала L-BFGS (Limited Memory Broyden–Fletcher–Goldfarb–Shanno) [15, 17], предполагавшая итеративную генерацию вредоносного возмущения посредством анализа гессиана функции потерь целевой модели и обратного движения в сторону роста значения функции потерь.

Другим методом атаки на основе вредоносного возмущения, оказавшим значительное влияние на дальнейшее развитие состязательных атак, стал FGSM (Fast Gradient Sign Method) [18, 19]. В основе указанного метода лежит использование знака градиента функции потерь целевой модели и генерация искажения в результате одноэтапного движения по градиенту к метке определенного класса. Рассматриваемый метод имеет меньшее время поиска вредоносного возмущения, однако генерируемое возмущение обладает высоким значением нормы. При этом FGSM характеризуется сравнительно низкой вероятностью успеха.

Многие последующие алгоритмы атак использовали FGSM в качестве основы и фактически являлись модификацией рассмотренного метода, в том числе FGVM (Fast Gradient Value Method) [20] и I-FGSM (Iterative Fast Gradient Sign Method) [21]. Атака по методу FGVM, что следует из названия, предполагает использование значения градиента вместо его знака. Метод атаки I-FGSM в свою очередь предусматривает генерацию искажения итеративно. В работах [22] и [23] предлагается ограничение генерируемого возмущения по метрикам  $L_2$  и  $L_\infty$  соответственно. Важно отметить, что атака по методу BIM (Basic Iterative Method) [23] была успешно воспроизведена и на физических объектах. Метод атаки PGD (Projected Gradient Descent) [24] также представляет собой вариацию итеративного FGSM, где для решения задачи оптимизации и нахождения оптимального искажения используется проецирование градиентного спуска. Существуют также и другие методы атаки, основанные на FGSM, в том числе его модификации [25–28].

Метод атаки DeepFool [29] предполагает итеративную генерацию атакующего изображения при помощи выпрямления границ принятия решения целевой модели и аппроксимации по ряду Тейлора для нахождения ближайшего класса к данному до тех пор, пока не будет сгенерировано наиболее подходящее минимальное вредоносное возмущение. Рассматриваемый метод характеризуется высокой вероятностью обхода целевой модели при использовании возмущения малой нормы, однако большое количество итераций приводит к большим вычислительным затратам и росту времени генерации.

Метод C&W (Carlini&Wagner) [30] представляет собой модификацию атаки L-BFGS с заменой нелинейных граничных условий для упрощения решения

задачи оптимизации. Преимуществами рассматриваемой атаки являются скорость генерации и минимизация вредоносного возмущения. Для указанного метода характерно малое смещение вероятностей принадлежности к определенному классу, что позволяет противостоять методам защиты, например методу дистилляции (Defense Distillation) [71].

Метод атаки TR (Trust Region based) [31] предполагает использование доверительных областей [72] для минимизации и ограничения вносимого искажения. Генерация вредоносного возмущения по указанному методу заключается в итеративном вычислении доверительного радиуса для максимизации вероятности некорректного класса внутри доверительной области. Рассмотренный метод достигает сравнимых с C&W показателей по метрикам расстояния для генерируемого возмущения, вследствие чего также способен противостоять методу дистилляции.

В отличие от рассмотренных конвенциональных атак, JSMA (Jacobian-based Saliency Map Attack) [32] и однопиксельная атака (One-pixel attack, Singlepixel attack) [33] вместо ограничения возмущения по метрике расстояния используют ограничение по размеру искажения.

Так, JSMA вычисляет Якобиан (Jacobian) [73] функции целевой модели и по полученным значениям формирует карту значимости (saliency map) [74] пикселей изображения. В соответствии с построенной картой значимости определяются элементы, оказывающие наибольшее влияние на отклик модели. Особенностью рассмотренного метода является получение показателей влияния на отклик модели посредством перебора большого числа входных параметров, что приводит к росту времени генерации вредоносного возмущения.

Однопиксельная атака [33] предполагает формирование искажения, не превышающего один пиксел. Для определения положения и значений RGB-компонент пиксела применяется алгоритм дифференциальной эволюции (Differential Evolution Algorithm, DEA) [75], что позволяет организовать атаку по модели черного ящика. Рассмотренный метод атаки более эффективен для небольших изображений [33].

Рассмотренные выше атаки относятся к атакам по модели белого ящика, их сравнение приведено в табл. 1.

Несмотря на то, что однопиксельная атака соответствует модели черного ящика, для оценки и последующего сравнения указанного метода более актуальны критерии атак по модели белого ящика, потому он также включен в табл. 1. В качестве критериев сравнения для таких атак выбраны размер и норма искажения, а также доля успешных обходов целевой модели. Размер искажения определен относительно исходного изображения размерами  $w \times h$ , где  $w$  и  $h$  — ширина и высота изображения соответственно. В качестве нормы искажения рассмотрены метрики, которые были учтены при генерации вредоносного возмущения при адресации атаки.

Рассмотренные методы генерации атакующих изображений имеют как преимущества, так и недостатки. Генерация вредоносного возмущения минимальной нормы требует больших затрат, однако такие искажения позволяют оставаться незаметными в том числе и для некоторых механизмов защиты [71]. Наиболее эффективным методом генерации вредоносного возмущения, согласно выбранным критериям, является метод TR [31], так как в его основе лежит оптимизация вычисления возмущения наименьшей нормы.

Рассмотрим атаки по модели черного ящика [33, 35–50].

Метод атаки Boundary attack [35] основан на отклике целевой модели. В качестве начального искажения для алгоритма генерации вредоносного возмущения выступает изображение целевого класса. С каждой итерацией алгоритм оптимизирует и уменьшает возмущение до тех пор, пока исходное изображение не будет корректно классифицировано. Тогда используется искажение, полученное на предыдущей итерации алгоритма.

Метод атаки HopSkipJump [36] является развитием Boundary attack и начальное поведение алгоритма генерации возмущения совпадает с предыдущим методом. Однако вблизи границы принятия решения HopSkipJump модифицирует искажение в зависимости от направления градиента. Длина шага вдоль направления градиента уменьшается в геометрической прогрессии для обеспечения минимального значения метрики расстояния возмущения.

Метод qFool [37] минимизирует вносимое искажение посредством нахождения направления градиен-

Таблица 1. Сравнение методов атаки по модели белого ящика

Table 1. Comparison of white-box attacks

Метод	Тип атаки	Размер искажения	Ограничение нормы искажения	Доля успешных атак, %
LBFGS [17]	Конвенциональная итеративная целевая по модели атака обхода	$w \times h$	—	87,65
FGSM [19]		$w \times h$	$\ x\ _2 \leq 6,25$	54,33
PGD [24]		$w \times h$	$\ x\ _\infty \leq 0,20$	91,85
DeepFool [29]	Конвенциональная итеративная целевая по модели и нецелевая по отклику атака обхода	$\leq w \times h$	$\ x\ _\infty \leq 0,20$	92,60
C&W [30]		$w \times h$	$\ x\ _\infty \leq 0,20$	94,80
TR [31]		$w \times h$	$\ x\ _\infty \leq 0,10$	94,77
JSMA [32]	Конвенциональная итеративная целевая по модели атака обхода	$< w \times h$	—	97,05
One-pixel [33]		$1 \times 1$	$\ x\ _0 = 1$	72,85

Таблица 2. Сравнение методов атаки по модели черного ящика [38]

Table 2. Comparison of attack methods of black-box attacks

Метод	Тип атаки	$L_2$	$L_\infty$	Число запросов
Boundary attack	Конвенциональная итеративная целевая по модели и целевая по классу атака обхода	5,13	0,052	800
HopSkipJump		1,85	0,012	42
qFool		1,12	—	3
GeoDA		1,01	0,003	14

та, имеющего кратчайший путь до границы принятия решения. Однако указанная граница нередко имеет значительную степень кривизны, тогда оптимизация процедуры вычисления минимального вредоносного возмущения может быть выполнена с помощью определения оптимальной нормали, что реализовано в методе GeoDA [38].

В работах [39–44] рассмотрены другие методы атаки по модели черного ящика, предполагающие использование отклика целевой модели.

Поскольку методы атаки по модели черного ящика выполняют минимизацию вносимого искажения при сохранении некорректного отклика модели, в качестве критериев сравнения выбраны метрики расстояния и количество итераций преобразования. Сравнение атак приведено в табл. 2.

В результате сравнения видно, что методы qFool [37] и GeoDA [38] достигают наименьших по норме возмущений за меньшее число итераций относительно существующих аналогов. Наиболее эффективным методом атаки по модели черного ящика является GeoDA [38], так как указанный метод решает проблему кривизны геометрической границы принятия решений в общем случае.

Рассмотренные атаки по модели черного ящика использовали только отклик целевой модели для генерации искажения. В то же время для уменьшения числа итераций в качестве основы для вредоносного возмущения могут быть использованы заранее подготовленные искажения. Прежде чем рассматривать такие атаки, необходимо упомянуть универсальные искажения.

В работе [34] были представлены вредоносные возмущения, не зависящие от задачи, решаемой целевой системой, и позволяющие осуществить обход таких систем. Отметим, что такие искажения могут быть перенесены как между различными входными данными, так и нейронными сетями.

Атака Customized Adversarial Boundary (CAB) [45] является развитием Boundary attack. Уменьшение числа запросов для генерации искажения достигается за счет инициализации возмущения, характерного для атак переноса вредоносного искажения, а также вычисления статистического распределения шума в предыдущих запросах.

Метод TRansferable EMBEDding based Black-box Attack (TREMBA) [46] предполагает использование модели архитектуры Encoder-Decoder [76] для ограничения пространства поиска вредоносного возмущения, что позволяет значительно ускорять процесс генерации искажения.

В работе [47] предложено использование дополнительной замещающей модели для генерации начального возмущения. Идея метода заключается в применении генератора для синтеза изображений, впоследствии обрабатываемых целевой моделью, и обучении модели на синтезированных данных и полученных откликах. Такой подход позволяет воспроизвести границы принятия решения целевой модели и использовать синтезированные изображения в качестве исходных данных для создания вредоносного возмущения.

Многие исследования развивают идею атак на основе переноса вредоносного возмущения [48–50].

Кроме рассмотренных конвенциональных состязательных атак, также существуют методы, предполагающие неограниченное возмущение [51–56]. Для нарушения работы целевой системы при таких атаках могут быть использованы такие искажения, как поворот изображения [52], значительное изменение цветовых компонент [51], добавление атрибутов [53] и др. К указанным атакам относятся также дипфейки [55, 56], позволяющие подменять человека на изображениях или в видеопотоке на другое лицо, формируя подделку.

Другой вид атак на основе вредоносных возмущений — модификация модели посредством встраивания бэкдора [66–69]. Внедрение бэкдора в целевую систему возможно путем добавления триггеров в виде специфичных паттернов возмущений на элементы обучающей выборки. При этом могут быть использованы как ограниченные [66], так и неограниченные возмущения, в том числе в виде состязательных патчей [67].

Метод модификации целевой системы DPatch [67] предполагает нанесение на изображения небольшого видимого искажения, представляющего собой квадратную область размером от  $20 \times 20$  пикселей. Указанный метод ориентирован на системы обнаружения объектов на изображении. Основным недостатком состязательных патчей является их заметность для человека, что значительно упрощает их обнаружение. Для придания триггерам более естественного вида предложены методы, адаптирующие патчи под естественные отражения или тени [68] и использующие плавную незначительную деформацию изображения [69]. В работе [66] разработан метод, предполагающий встраивание бэкдора посредством внесения вредоносного возмущения, сравнимого с конвенциональными состязательными атаками.

Рассмотренные методы атаки в основном направлены на некорректный отклик модели при классификации, однако также существуют методы, направленные на нарушение работы нейронных сетей при решении других задач обработки изображений [57–59].

### Методы противодействия атакам на основе вредоносных возмущений

Вследствие актуальности и значимости рассмотренных атак для многих систем, были разработаны методы противодействия им. Существующие методы защиты включают:

- методы модификации целевой модели [17, 77, 78];
- методы модификации целевой системы, реализующей технологии искусственного интеллекта [71, 79, 80];
- предобработку или изменение входных данных [79–83].

Кроме того, некоторые методы предполагают как обнаружение, так и устранение искажения или его влияния на нейронную сеть [17, 77, 78, 80–83], тогда как другие — только обнаружение [71, 79].

Отметим, что методы защиты должны удовлетворять следующим критериям [71]:

- минимальное воздействие на архитектуру;
- минимальное влияние на показатели качества;
- минимальное влияние на быстродействие.

Одним из методов, предполагающих модификацию модели, является состязательное обучение (Adversarial Learning) [17, 77, 79]. Сущность метода заключается в обучении целевой модели в том числе на экземплярах, содержащих вредоносное искажение, однако сопоставляемых с корректным откликом. Согласно исследованиям, при состязательном обучении доля некорректных откликов при адресации атаки, основанной на FGSM, снизилась с 89,4 % до 17,9 % [78]. Однако применение указанного метода защиты влияет и на обработку данных, не содержащих вредоносного возмущения — точность классификации снижается. Сложностью при состязательном обучении является обеспечение репрезентативности обучающей выборки и контроль ее статистического распределения. Кроме того, указанный метод предполагает переобучение целевой нейронной сети, что не всегда возможно.

Такой метод, как защитная дистилляция (Defense Distillation) [71], предполагает обучение дополнительной нейронной сети идентичной архитектуры, однако использующей при обучении в качестве предикторов отклики защищаемой модели, представляющие собой вероятности отнесения входного изображения к тому или иному классу. При генерации и наложении вредоносного искажения вероятность отнесения входных данных к некорректному классу нередко достигает значений, не характерных для «чистых» изображений. Такая уверенность в классификации представляет собой аномалию, которая может быть обнаружена защищающей нейронной сетью. Однако защитная дистилляция не позволяет обнаруживать некоторые виды атак, например, C&W [30].

Другим вариантом противодействия рассмотренным атакам является сжатие параметров входных данных (Feature Squeezing) [79], что может быть достигнуто в том числе при понижении размерности метаданных. Метод предполагает обработку как оригинального изображения моделью машинного обучения защищаемой системы, реализующей технологии искусственного

интеллекта, так и модифицированного изображения дополнительной нейронной сетью. Модификация входных данных нарушает целостность вредоносного возмущения, вносимого атакой, в результате чего нейронные сети вернут различные отклики, что будет свидетельствовать о факте атаки. Сжатие параметров и защитная дистилляция предполагают использование дополнительной нейронной сети, что представляет собой значительную избыточность, а также позволяют только обнаруживать атаку.

Метод сертификационной защиты (Certified Defense) [81] предоставляет гарантию устойчивости модели к атакам на основе вредоносных возмущений определенной нормы. Повышение устойчивости модели достигается с помощью вычисления нижней границы  $L_p$  нормы возмущения для проведения успешной атаки и применения преобразований, направленных на нарушение целостности искажения [81, 82]. Однако такие методы защиты, как и состязательное обучение, негативно влияют на показатели качества модели. Кроме того, методы, реализующие сертификационную защиту, позволяют организовать защиту модели от возмущений до определенной нормы. Иными словами, указанный метод защиты не позволяет организовать защиту от множества различных атак.

Противодействие атакам на основе вредоносных возмущений может быть реализовано посредством шумоподавления [80, 83]. Такие методы защиты предполагают обнаружение и классификацию атаки с последующим устранением искажения. Однако методы шумоподавления не позволяют оказывать эффективное противодействие некоторым атакам, в том числе FGSM [18, 19] и BIM [23]. Кроме того, они предполагают сигнатурное определение атак, что не позволяет организовать защиту от неизвестных системе атак.

Достоинства и недостатки рассмотренных методов приведены в табл. 3.

Согласно табл. 3, существующие методы имеют значительные недостатки. Наиболее распространенный недостаток методов противодействия — использование дополнительной нейронной сети, что требует увеличения вычислительных мощностей. Этот факт накладывает значительные ограничения на область применения методов, обладающих указанным недостатком. Еще один немаловажный недостаток — невозможность устранения возмущения, что в свою очередь ограничивает область применения. Так, например, методы защиты, не позволяющие устранять возмущение неэффективны в системах реального времени, таких как беспилотный транспорт. Кроме того, для некоторых методов защиты имеет место снижение показателей качества целевой модели.

### Возможные подходы к обнаружению и устранению вредоносных возмущений

Угрозы, связанные с атаками на основе вредоносных искажений — актуальны и критичны для многих систем, реализующих технологии искусственного интеллекта в различных прикладных областях. В то же время недостатки и ограничения существующих ме-

Таблица 3. Сравнение методов противодействия атакам на основе вредоносных возмущений  
 Table 3. Comparison of defense methods against attacks based on malicious perturbation

Метод	Достоинства	Недостатки
Состязательное обучение [17, 77, 78]	<ul style="list-style-type: none"> <li>— позволяет игнорировать вредоносное возмущение;</li> <li>— не предполагает дополнительных вычислений</li> </ul>	<ul style="list-style-type: none"> <li>— снижение показателей качества целевой модели;</li> <li>— контроль репрезентативности обучающей выборки;</li> <li>— необходимость переобучения целевой нейронной сети</li> </ul>
Защитная дистилляция [71]	<ul style="list-style-type: none"> <li>— не оказывает значительного влияния на качество целевой нейронной сети;</li> <li>— позволяет обнаруживать атаки</li> </ul>	<ul style="list-style-type: none"> <li>— использование дополнительной нейронной сети;</li> <li>— неустойчивость к некоторым методам атаки;</li> <li>— не позволяет устранять возмущение</li> </ul>
Сжатие параметров [79]	<ul style="list-style-type: none"> <li>— не оказывает значительного влияния на качество целевой нейронной сети;</li> <li>— позволяет обнаруживать атаки</li> </ul>	<ul style="list-style-type: none"> <li>— использование дополнительной нейронной сети;</li> <li>— не позволяет устранять возмущение;</li> <li>— неприменим в системах, где качество изображения критически важно</li> </ul>
Сертификационная защита [81, 82]	<ul style="list-style-type: none"> <li>— позволяет устранять вредоносное возмущение;</li> <li>— сохранение качества при использовании для защиты нейронных сетей схожего функционала;</li> <li>— гарантия защиты от возмущения до определенной нормы</li> </ul>	<ul style="list-style-type: none"> <li>— снижение показателей качества целевой модели;</li> <li>— сложность адаптации ко множеству различных атак</li> </ul>
Шумоподавление [80, 83]	<ul style="list-style-type: none"> <li>— позволяет устранить вредоносное возмущение;</li> <li>— повышение качества изображения</li> </ul>	<ul style="list-style-type: none"> <li>— использование дополнительной нейронной сети;</li> <li>— зависимость эффективности от целевой модели;</li> <li>— неустойчивость к некоторым методам атаки;</li> <li>— низкая эффективность против неизвестных атак;</li> <li>— возможность удаления значимой информации.</li> </ul>

тодов защиты не позволяют оказывать эффективного противодействия указанным атакам. Ввиду изложенного противоречия возникает необходимость улучшения существующих и разработки новых подходов к противодействию.

Обнаружение вредоносного возмущения возможно посредством статистического анализа шумовой компоненты изображения, в том числе при помощи дискретного косинусного преобразования и дискретного преобразования Фурье. Для анализа результатов преобразований возможно использование нечетного байесовского классификатора (Bayesian fuzzy clustering) [84], а для обработки шумовой компоненты в изначальном виде — критерия хи-квадрат (Chi-squared test) [85]. Данный подход может иметь ограничение по минимальной норме возмущения.

Устранение возмущения возможно при внесении обратимых визуальных модификаций и искажений изображения. Тогда использование прямых и обратных преобразований в различном порядке потенциально позволит нарушить целостность вредоносного возмущения. Заметим, что предложенный подход может иметь ограничение по максимальной норме устраняе-

мого искажения и может быть неприменим в некоторых прикладных областях.

В то же время эффективное противодействие атакам, оптимизированным по псевдонорме  $L_0$ , предполагает использование иных подходов, описанных далее. Обнаружение вредоносного искажения, характерного для указанных атак, возможно посредством статистических методов, например: Z-оценки, гистограммного анализа, вычисления расстояния Махаланобиса [86]. Ввиду наличия определенных недостатков в указанных методах, возможно использование их комбинации. Кроме того, некоторые методы также вычисляют предполагаемые RGB-компоненты искаженной области.

Комбинация предложенных подходов потенциально позволяет обнаруживать и устранять вредоносные возмущения вне зависимости от их нормы. Следует отметить, что описанные подходы применимы исключительно для ограниченных искажений.

### Заключение

Угрозы, связанные с атаками на основе вредоносных искажений, являются актуальными и критичными для многих систем, реализующих технологии искус-

ственного интеллекта в различных прикладных областях. В то же время недостатки и ограничения существующих методов защиты не позволяют оказывать эффективного противодействия указанным атакам. Из-за изложенного противоречия возникает необходимость улучшения существующих и разработки новых подходов к противодействию. В работе рассмотрены некоторые возможные варианты решения противоречия.

### Литература

1. Goldberg Y. A primer on neural network models for natural language processing // *Journal of Artificial Intelligence Research*. 2016. V. 57. P. 345–420. <https://doi.org/10.1613/jair.4992>
2. Nassif A.B., Shahin I., Attili I., Azzeh M., Shaalan K. Speech recognition using deep neural networks: A systematic review // *IEEE Access*. 2019. V. 7. P. 19143–19165. <https://doi.org/10.1109/access.2019.2896880>
3. Almabdy S., Elrefaei L. Deep convolutional neural network-based approaches for face recognition // *Applied Sciences*. 2019. V. 9. N 20. P. 4397. <https://doi.org/10.3390/app9204397>
4. Khan M.Z., Harous S., Hassan S. U., Khan M. U. G., Iqbal R., Mumtaz S. Deep unified model for face recognition based on convolution neural network and edge computing // *IEEE Access*. 2019. V. 7. P. 72622–72633. <https://doi.org/10.1109/access.2019.2918275>
5. Zhang Y., Shi D., Zhan X., Cao D., Zhu K., Li Z. Slim-ResCNN: A deep residual convolutional neural network for fingerprint liveness detection // *IEEE Access*. 2019. V. 7. P. 91476–91487. <https://doi.org/10.1109/access.2019.2927357>
6. Sarvamangala D.R., Kulkarni R.V. Convolutional neural networks in medical image understanding: a survey // *Evolutionary Intelligence*. 2022. V. 15. N 1. P. 1–22. <https://doi.org/10.1007/s12065-020-00540-3>
7. Mahmood M., Al-Khateeb B., Alwash W. A review on neural networks approach on classifying cancers // *IAES International Journal of Artificial Intelligence*. 2020. V. 9. N 2. P. 317–326. <http://doi.org/10.11591/ijai.v9.i2.pp317-326>
8. Singh V., Singh S., Gupta P. Real-time anomaly recognition through CCTV using neural networks // *Procedia Computer Science*. 2020. V. 173. P. 254–263. <https://doi.org/10.1016/j.procs.2020.06.030>
9. Severino A., Curto S., Barberi S., Arena F., Pau G. Autonomous vehicles: an analysis both on their distinctiveness and the potential impact on urban transport systems // *Applied Sciences*. 2021. V. 11. N 8. P. 3604. <https://doi.org/10.3390/app11083604>
10. Wang L., Fan X., Chen J., Cheng J., Tan J., Ma X. 3D object detection based on sparse convolution neural network and feature fusion for autonomous driving in smart cities // *Sustainable Cities and Society*. 2020. V. 54. P. 102002. <https://doi.org/10.1016/j.scs.2019.102002>
11. Chen L., Lin S., Lu X., Cao D., Wu H., Guo C., Wang F. Y. Deep neural network based vehicle and pedestrian detection for autonomous driving: A survey // *IEEE Transactions on Intelligent Transportation Systems*. 2021. V. 22. N 6. P. 3234–3246. <https://doi.org/10.1109/tits.2020.2993926>
12. Chen P. Y., Liu S. Holistic adversarial robustness of deep learning models // *Proceedings of the AAAI Conference on Artificial Intelligence*. 2023. V. 37. N 13. P. 15411–15420. <https://doi.org/10.1609/aaai.v37i13.26797>
13. Huang X., Kroening D., Ruan W., Sharp J., Sun Y., Thamo E., Min W., Yi X. A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability // *Computer Science Review*. 2020. V. 37. P. 100270. <https://doi.org/10.1016/j.cosrev.2020.100270>
14. Szegedy C., Zaremba W., Sutskever I., Bruna J., Erhan D., Goodfellow I., Fergus R. Intriguing properties of neural networks // *arXiv*. 2013. arXiv:1312.6199. <https://doi.org/10.48550/arXiv.1312.6199>
15. Song Y., Shu R., Kushman N., Ermon S. Constructing unrestricted adversarial examples with generative models // *Advances in Neural Information Processing Systems*. 2018. V. 31.
16. Sayghe A., Zhao J., Konstantinou C. Evasion attacks with adversarial deep learning against power system state estimation // *Proc. of the 2020 IEEE Power & Energy Society General Meeting (PESGM)*. 2020. P. 1–5. <https://doi.org/10.1109/pesgm41954.2020.9281719>

### References

1. Goldberg Y. A primer on neural network models for natural language processing. *Journal of Artificial Intelligence Research*, 2016, vol. 57, pp. 345–420. <https://doi.org/10.1613/jair.4992>
2. Nassif A.B., Shahin I., Attili I., Azzeh M., Shaalan K. Speech recognition using deep neural networks: A systematic review. *IEEE Access*, 2019, vol. 7, pp. 19143–19165. <https://doi.org/10.1109/access.2019.2896880>
3. Almabdy S., Elrefaei L. Deep convolutional neural network-based approaches for face recognition. *Applied Sciences*, 2019, vol. 9, no. 20, pp. 4397. <https://doi.org/10.3390/app9204397>
4. Khan M.Z., Harous S., Hassan S. U., Khan M. U. G., Iqbal R., Mumtaz S. Deep unified model for face recognition based on convolution neural network and edge computing. *IEEE Access*, 2019, vol. 7, pp. 72622–72633. <https://doi.org/10.1109/access.2019.2918275>
5. Zhang Y., Shi D., Zhan X., Cao D., Zhu K., Li Z. Slim-ResCNN: A deep residual convolutional neural network for fingerprint liveness detection. *IEEE Access*, 2019, vol. 7, pp. 91476–91487. <https://doi.org/10.1109/access.2019.2927357>
6. Sarvamangala D.R., Kulkarni R.V. Convolutional neural networks in medical image understanding: a survey. *Evolutionary Intelligence*, 2022, vol. 15, no. 1, pp. 1–22. <https://doi.org/10.1007/s12065-020-00540-3>
7. Mahmood M., Al-Khateeb B., Alwash W. A review on neural networks approach on classifying cancers. *IAES International Journal of Artificial Intelligence*, 2020, vol. 9, no. 2, pp. 317–326. <http://doi.org/10.11591/ijai.v9.i2.pp317-326>
8. Singh V., Singh S., Gupta P. Real-time anomaly recognition through CCTV using neural networks. *Procedia Computer Science*, 2020, vol. 173, pp. 254–263. <https://doi.org/10.1016/j.procs.2020.06.030>
9. Severino A., Curto S., Barberi S., Arena F., Pau G. Autonomous vehicles: an analysis both on their distinctiveness and the potential impact on urban transport systems. *Applied Sciences*, 2021, vol. 11, no. 8, pp. 3604. <https://doi.org/10.3390/app11083604>
10. Wang L., Fan X., Chen J., Cheng J., Tan J., Ma X. 3D object detection based on sparse convolution neural network and feature fusion for autonomous driving in smart cities. *Sustainable Cities and Society*, 2020, vol. 54, pp. 102002. <https://doi.org/10.1016/j.scs.2019.102002>
11. Chen L., Lin S., Lu X., Cao D., Wu H., Guo C., Wang F. Y. Deep neural network based vehicle and pedestrian detection for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 2021, vol. 22, no. 6, pp. 3234–3246. <https://doi.org/10.1109/tits.2020.2993926>
12. Chen P. Y., Liu S. Holistic adversarial robustness of deep learning models. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2023, vol. 37, no. 13, pp. 15411–15420. <https://doi.org/10.1609/aaai.v37i13.26797>
13. Huang X., Kroening D., Ruan W., Sharp J., Sun Y., Thamo E., Min W., Yi X. A survey of safety and trustworthiness of deep neural networks: Verification, testing, adversarial attack and defence, and interpretability. *Computer Science Review*, 2020, vol. 37, pp. 100270. <https://doi.org/10.1016/j.cosrev.2020.100270>
14. Szegedy C., Zaremba W., Sutskever I., Bruna J., Erhan D., Goodfellow I., Fergus R. Intriguing properties of neural networks. *arXiv*, 2013, arXiv:1312.6199. <https://doi.org/10.48550/arXiv.1312.6199>
15. Song Y., Shu R., Kushman N., Ermon S. Constructing unrestricted adversarial examples with generative models. *Advances in Neural Information Processing Systems*, 2018, vol. 31.
16. Sayghe A., Zhao J., Konstantinou C. Evasion attacks with adversarial deep learning against power system state estimation. *Proc. of the 2020 IEEE Power & Energy Society General Meeting (PESGM)*, 2020, pp. 1–5. <https://doi.org/10.1109/pesgm41954.2020.9281719>

17. Goodfellow I.J., Shlens J., Szegedy C. Explaining and harnessing adversarial examples // arXiv. 2014. arXiv:1412.6572. <https://doi.org/10.48550/arXiv.1412.6572>
18. Paul R., Schabath M., Gillies R., Hall L., Goldgof D. Mitigating adversarial attacks on medical image understanding systems // Proc. of the 2020 IEEE 17<sup>th</sup> International Symposium on Biomedical Imaging (ISBI). 2020. P. 1517–1521. <https://doi.org/10.1109/isbi45749.2020.9098740>
19. Rozsa A., Rudd E.M., Boulton T.E. Adversarial diversity and hard positive generation // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2016. P. 25–32. <https://doi.org/10.1109/cvprw.2016.58>
20. Dong Y., Liao F., Pang T., Su H., Zhu J., Hu X., Li J. Boosting adversarial attacks with momentum // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 9185–9193. <https://doi.org/10.1109/cvpr.2018.00957>
21. Miyato T., Maeda S.I., Koyama M., Ishii S. Virtual adversarial training: a regularization method for supervised and semi-supervised learning // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2019. V. 41. N 8. P. 1979–1993. <https://doi.org/10.1109/tpami.2018.2858821>
22. Kurakin A., Goodfellow I.J., Bengio S. Adversarial examples in the physical world // Artificial Intelligence Safety and Security. Chapman and Hall/CRC, 2018. P. 99–112. <https://doi.org/10.1201/9781351251389-8>
23. Madry A., Makelov A., Schmidt L., Tsipras D., Vladu A. Towards deep learning models resistant to adversarial attacks // Stat. 2017. V. 1050. P. 9.
24. Xie C., Zhang Z., Zhou Y., Bai S., Wang J., Ren Z., Yuille A.L. Improving transferability of adversarial examples with input diversity // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019. P. 2730–2739. <https://doi.org/10.1109/cvpr.2019.00284>
25. Dong X., Han J., Chen D., Liu J., Bian H., Ma Z., Li H., Wang X., Zhang W., Yu N. Robust superpixel-guided attentional adversarial attack // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020. P. 12895–12904. <https://doi.org/10.1109/cvpr42600.2020.01291>
26. Sriramanan G., Addepalli S., Baburaj A. Guided adversarial attack for evaluating and enhancing adversarial defenses // Advances in Neural Information Processing Systems. 2020. V. 33. P. 20297–20308.
27. Rony J., Hafemann L.G., Oliveira L.S., Ayed I.B., Sabourin R., Granger E. Decoupling direction and norm for efficient gradient-based L2 adversarial attacks and defenses // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019. P. 4322–4330. <https://doi.org/10.1109/cvpr.2019.00445>
28. Moosavi-Dezfooli S.M., Fawzi A., Frossard P. DeepFool: a simple and accurate method to fool deep neural networks // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016. P. 2574–2582. <https://doi.org/10.1109/cvpr.2016.282>
29. Carlini N., Wagner D. Towards evaluating the robustness of neural networks // Proc. of the IEEE Symposium on Security and Privacy (SP). 2017. P. 39–57. <https://doi.org/10.1109/sp.2017.49>
30. Yao Z., Gholami A., Xu P., Keutzer K., Mahoney M. W. Trust region based adversarial attack on neural networks // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019. P. 11350–11359. <https://doi.org/10.1109/cvpr.2019.01161>
31. Papernot N., McDaniel P., Jha S., Fredrikson M., Celik Z. B., Swami A. The limitations of deep learning in adversarial settings // Proc. of the 2016 IEEE European Symposium on Security and Privacy (EuroS&P). 2016. P. 372–387. <https://doi.org/10.1109/eurosp.2016.36>
32. Su J., Vargas D.V., Sakurai K. One pixel attack for fooling deep neural networks // IEEE Transactions on Evolutionary Computation. 2019. V. 23. N 5. P. 828–841. <https://doi.org/10.1109/tevc.2019.2890858>
33. Moosavi-Dezfooli S.M., Fawzi A., Fawzi O., Frossard P. Universal adversarial perturbations // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. P. 1765–1773. <https://doi.org/10.1109/cvpr.2017.17>
34. Brendel W., Rauber J., Bethge M. Decision-based adversarial attacks: Reliable attacks against black-box machine learning models // Advances in Reliably Evaluating and Improving Adversarial Robustness. 2021. P. 77.
35. Chen J., Jordan M.I., Wainwright M.J. HopSkipJumpAttack: A query-efficient decision-based attack // Proc. of the 2020 IEEE Symposium on Security and Privacy (SP). 2020. P. 1277–1294. <https://doi.org/10.1109/sp40000.2020.00045>
17. Goodfellow I.J., Shlens J., Szegedy C. Explaining and harnessing adversarial examples. *arXiv*, 2014, arXiv:1412.6572. <https://doi.org/10.48550/arXiv.1412.6572>
18. Paul R., Schabath M., Gillies R., Hall L., Goldgof D. Mitigating adversarial attacks on medical image understanding systems. *Proc. of the 2020 IEEE 17<sup>th</sup> International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1517–1521. <https://doi.org/10.1109/isbi45749.2020.9098740>
19. Rozsa A., Rudd E.M., Boulton T.E. Adversarial diversity and hard positive generation. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2016, pp. 25–32. <https://doi.org/10.1109/cvprw.2016.58>
20. Dong Y., Liao F., Pang T., Su H., Zhu J., Hu X., Li J. Boosting adversarial attacks with momentum. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 9185–9193. <https://doi.org/10.1109/cvpr.2018.00957>
21. Miyato T., Maeda S.I., Koyama M., Ishii S. Virtual adversarial training: a regularization method for supervised and semi-supervised learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, vol. 41, no. 8, pp. 1979–1993. <https://doi.org/10.1109/tpami.2018.2858821>
22. Kurakin A., Goodfellow I.J., Bengio S. Adversarial examples in the physical world. *Artificial Intelligence Safety and Security*. Chapman and Hall/CRC, 2018, pp. 99–112. <https://doi.org/10.1201/9781351251389-8>
23. Madry A., Makelov A., Schmidt L., Tsipras D., Vladu A. Towards deep learning models resistant to adversarial attacks. *Stat*, 2017, vol. 1050, pp. 9.
24. Xie C., Zhang Z., Zhou Y., Bai S., Wang J., Ren Z., Yuille A.L. Improving transferability of adversarial examples with input diversity. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 2730–2739. <https://doi.org/10.1109/cvpr.2019.00284>
25. Dong X., Han J., Chen D., Liu J., Bian H., Ma Z., Li H., Wang X., Zhang W., Yu N. Robust superpixel-guided attentional adversarial attack. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 12895–12904. <https://doi.org/10.1109/cvpr42600.2020.01291>
26. Sriramanan G., Addepalli S., Baburaj A. Guided adversarial attack for evaluating and enhancing adversarial defenses. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 20297–20308.
27. Rony J., Hafemann L.G., Oliveira L.S., Ayed I.B., Sabourin R., Granger E. Decoupling direction and norm for efficient gradient-based L2 adversarial attacks and defenses. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4322–4330. <https://doi.org/10.1109/cvpr.2019.00445>
28. Moosavi-Dezfooli S.M., Fawzi A., Frossard P. DeepFool: a simple and accurate method to fool deep neural networks. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2574–2582. <https://doi.org/10.1109/cvpr.2016.282>
29. Carlini N., Wagner D. Towards evaluating the robustness of neural networks. *Proc. of the IEEE Symposium on Security and Privacy (SP)*, 2017, pp. 39–57. <https://doi.org/10.1109/sp.2017.49>
30. Yao Z., Gholami A., Xu P., Keutzer K., Mahoney M. W. Trust region based adversarial attack on neural networks. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 11350–11359. <https://doi.org/10.1109/cvpr.2019.01161>
31. Papernot N., McDaniel P., Jha S., Fredrikson M., Celik Z. B., Swami A. The limitations of deep learning in adversarial settings. *Proc. of the 2016 IEEE European Symposium on Security and Privacy (EuroS&P)*, 2016, pp. 372–387. <https://doi.org/10.1109/eurosp.2016.36>
32. Su J., Vargas D.V., Sakurai K. One pixel attack for fooling deep neural networks. *IEEE Transactions on Evolutionary Computation*, 2019, vol. 23, no. 5, pp. 828–841. <https://doi.org/10.1109/tevc.2019.2890858>
33. Moosavi-Dezfooli S.M., Fawzi A., Fawzi O., Frossard P. Universal adversarial perturbations. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1765–1773. <https://doi.org/10.1109/cvpr.2017.17>
34. Brendel W., Rauber J., Bethge M. Decision-based adversarial attacks: Reliable attacks against black-box machine learning models. *Advances in Reliably Evaluating and Improving Adversarial Robustness*, 2021, pp. 77.
35. Chen J., Jordan M.I., Wainwright M.J. HopSkipJumpAttack: A query-efficient decision-based attack. *Proc. of the 2020 IEEE Symposium*

36. Liu Y., Moosavi-Dezfooli S.M., Frossard P. A geometry-inspired decision-based attack // *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019. P. 4890–4898. <https://doi.org/10.1109/iccv.2019.00499>
37. Rahmati A., Moosavi-Dezfooli S.M., Frossard P., Dai H. GeoDA: a geometric framework for black-box adversarial attacks // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. P. 8446–8455. <https://doi.org/10.1109/cvpr42600.2020.00847>
38. Du J., Zhang H., Zhou J.T., Yang Y., Feng J. Query-efficient meta attack to deep neural networks // *Proc. of the International Conference on Learning Representations*, 2020.
39. Li J., Ji R., Liu H., Liu J., Zhong B., Deng C., Tian Q. Projection & probability-driven black-box attack // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. P. 362–371. <https://doi.org/10.1109/cvpr42600.2020.00044>
40. Li H., Xu X., Zhang X., Yang S., Li B. QEBA: Query-efficient boundary-based blackbox attack // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. P. 1221–1230. <https://doi.org/10.1109/cvpr42600.2020.00130>
41. Cheng M., Singh S., Chen P., Chen P.Y., Liu S., Hsieh C.J. Sign-OPT: A query-efficient hard-label adversarial attack // *Proc. of the International Conference on Learning Representations*, 2020.
42. Brunner T., Diehl F., Le M.T., Knoll A. Guessing smart: Biased sampling for efficient black-box adversarial attacks // *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019. P. 4958–4966. <https://doi.org/10.1109/iccv.2019.00506>
43. Maho T., Furon T., Le Merrer E. SurFree: a fast surrogate-free black-box attack // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. P. 10430–10439. <https://doi.org/10.1109/cvpr46437.2021.01029>
44. Shi Y., Han Y., Tian Q. Polishing decision-based adversarial noise with a customized sampling // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. P. 1030–1038. <https://doi.org/10.1109/cvpr42600.2020.00111>
45. Huang Z., Zhang T. Black-box adversarial attack with transferable model-based embedding // *Proc. of the International Conference on Learning Representations*, 2020.
46. Zhou M., Wu J., Liu Y., Liu S., Zhu C. DaST: Data-free substitute training for adversarial attacks // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020. P. 234–243. <https://doi.org/10.1109/cvpr42600.2020.00031>
47. Zou J., Pan Z., Qiu J., Liu X., Rui T., Li W. Improving the transferability of adversarial examples with resized-diverse-inputs, diversity-ensemble and region fitting // *Lecture Notes in Computer Science*, 2020. V. 12367. P. 563–579. [https://doi.org/10.1007/978-3-030-58542-6\\_34](https://doi.org/10.1007/978-3-030-58542-6_34)
48. Wang X., He K. Enhancing the transferability of adversarial attacks through variance tuning // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. P. 1924–1933. <https://doi.org/10.1109/cvpr46437.2021.00196>
49. Wu W., Su Y., Lyu M.R., King I. Improving the transferability of adversarial samples with adversarial transformations // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. P. 9024–9033. <https://doi.org/10.1109/cvpr46437.2021.00891>
50. Hosseini H., Poovendran R. Semantic adversarial examples // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018. P. 1614–1619. <https://doi.org/10.1109/cvprw.2018.00212>
51. Engstrom L., Tran B., Tsipras D., Schmidt L., Madry A. A rotation and a translation suffice: Fooling cnns with simple transformations [Электронный ресурс]. URL: <https://openreview.net/forum?id=BJfvknCqFQ> (дата обращения: 29.05.2023).
52. Joshi A., Mukherjee A., Sarkar S., Hegde C. Semantic adversarial attacks: Parametric transformations that fool deep classifiers // *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019. P. 4773–4783. <https://doi.org/10.1109/iccv.2019.00487>
53. Liu A., Wang J., Liu X., Cao B., Zhang C., Yu H. Bias-based universal adversarial patch attack for automatic check-out // *Lecture Notes in Computer Science*, 2020. V. 12358. P. 395–410. [https://doi.org/10.1007/978-3-030-58601-0\\_24](https://doi.org/10.1007/978-3-030-58601-0_24)
54. Swathi P., Sk S. DeepFake creation and detection: A survey // *Proc. of the 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2021. P. 584–588. <https://doi.org/10.1109/icirca51532.2021.9544522>
36. Liu Y., Moosavi-Dezfooli S.M., Frossard P. A geometry-inspired decision-based attack. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 4890–4898. <https://doi.org/10.1109/sp40000.2020.00045>
37. Rahmati A., Moosavi-Dezfooli S.M., Frossard P., Dai H. GeoDA: a geometric framework for black-box adversarial attacks. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 8446–8455. <https://doi.org/10.1109/cvpr42600.2020.00847>
38. Du J., Zhang H., Zhou J.T., Yang Y., Feng J. Query-efficient meta attack to deep neural networks. *Proc. of the International Conference on Learning Representations*, 2020.
39. Li J., Ji R., Liu H., Liu J., Zhong B., Deng C., Tian Q. Projection & probability-driven black-box attack. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 362–371. <https://doi.org/10.1109/cvpr42600.2020.00044>
40. Li H., Xu X., Zhang X., Yang S., Li B. QEBA: Query-efficient boundary-based blackbox attack. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 1221–1230. <https://doi.org/10.1109/cvpr42600.2020.00130>
41. Cheng M., Singh S., Chen P., Chen P.Y., Liu S., Hsieh C.J. Sign-OPT: A query-efficient hard-label adversarial attack. *Proc. of the International Conference on Learning Representations*, 2020.
42. Brunner T., Diehl F., Le M.T., Knoll A. Guessing smart: Biased sampling for efficient black-box adversarial attacks. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 4958–4966. <https://doi.org/10.1109/iccv.2019.00506>
43. Maho T., Furon T., Le Merrer E. SurFree: a fast surrogate-free black-box attack. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 10430–10439. <https://doi.org/10.1109/cvpr46437.2021.01029>
44. Shi Y., Han Y., Tian Q. Polishing decision-based adversarial noise with a customized sampling. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 1030–1038. <https://doi.org/10.1109/cvpr42600.2020.00111>
45. Huang Z., Zhang T. Black-box adversarial attack with transferable model-based embedding. *Proc. of the International Conference on Learning Representations*, 2020.
46. Zhou M., Wu J., Liu Y., Liu S., Zhu C. DaST: Data-free substitute training for adversarial attacks. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 234–243. <https://doi.org/10.1109/cvpr42600.2020.00031>
47. Zou J., Pan Z., Qiu J., Liu X., Rui T., Li W. Improving the transferability of adversarial examples with resized-diverse-inputs, diversity-ensemble and region fitting. *Lecture Notes in Computer Science*, 2020, vol. 12367, pp. 563–579. [https://doi.org/10.1007/978-3-030-58542-6\\_34](https://doi.org/10.1007/978-3-030-58542-6_34)
48. Wang X., He K. Enhancing the transferability of adversarial attacks through variance tuning. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 1924–1933. <https://doi.org/10.1109/cvpr46437.2021.00196>
49. Wu W., Su Y., Lyu M.R., King I. Improving the transferability of adversarial samples with adversarial transformations. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 9024–9033. <https://doi.org/10.1109/cvpr46437.2021.00891>
50. Hosseini H., Poovendran R. Semantic adversarial examples. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2018, pp. 1614–1619. <https://doi.org/10.1109/cvprw.2018.00212>
51. Engstrom L., Tran B., Tsipras D., Schmidt L., Madry A. A rotation and a translation suffice: Fooling cnns with simple transformations. Available at: <https://openreview.net/forum?id=BJfvknCqFQ> (accessed: 29.05.2023).
52. Joshi A., Mukherjee A., Sarkar S., Hegde C. Semantic adversarial attacks: Parametric transformations that fool deep classifiers. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 4773–4783. <https://doi.org/10.1109/iccv.2019.00487>
53. Liu A., Wang J., Liu X., Cao B., Zhang C., Yu H. Bias-based universal adversarial patch attack for automatic check-out. *Lecture Notes in Computer Science*, 2020, vol. 12358, pp. 395–410. [https://doi.org/10.1007/978-3-030-58601-0\\_24](https://doi.org/10.1007/978-3-030-58601-0_24)
54. Swathi P., Sk S. DeepFake creation and detection: A survey. *Proc. of the 2021 Third International Conference on Inventive Research in*

55. Chadha A., Kumar V., Kashyap S., Gupta M. Deepfake: An Overview // *Lecture Notes in Networks and Systems*. 2021. V. 203. P. 557–566. [https://doi.org/10.1007/978-981-16-0733-2\\_39](https://doi.org/10.1007/978-981-16-0733-2_39)
56. Nakka K.K., Salzmann M. Indirect local attacks for context-aware semantic segmentation networks // *Lecture Notes in Computer Science*. 2020. V. 12350. P. 611–628. [https://doi.org/10.1007/978-3-030-58558-7\\_36](https://doi.org/10.1007/978-3-030-58558-7_36)
57. He Y., Rahimian S., Schiele B., Fritz M. Segmentations-leak: Membership inference attacks and defenses in semantic image segmentation // *Lecture Notes in Computer Science*. 2020. V. 12368. P. 519–535. [https://doi.org/10.1007/978-3-030-58592-1\\_31](https://doi.org/10.1007/978-3-030-58592-1_31)
58. Choi J.H. Zhang H., Kim J.H., Hsieh C.J., Lee J.S. Evaluating robustness of deep image super-resolution against adversarial attacks // *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019. P. 303–311. <https://doi.org/10.1109/iccv.2019.00039>
59. Jiang L., Ma X., Chen S., Bailey J., Jiang Y.G. Black-box adversarial attacks on video recognition models // *Proc. of the 27th ACM International Conference on Multimedia*. 2019. P. 864–872. <https://doi.org/10.1145/3343031.3351088>
60. Li S., Aich A., Zhu S., Asif S., Song C., Roy-Chowdhury A., Krishnamurthy S. Adversarial attacks on black box video classifiers: Leveraging the power of geometric transformations // *Advances in Neural Information Processing Systems*. 2021. V. 34. P. 2085–2096.
61. Chen X., Li S., Huang H. Adversarial attack and defense on deep neural network-based voice processing systems: An overview // *Applied Sciences*. 2021. V. 11. N 18. P. 8450. <https://doi.org/10.3390/app11188450>
62. Kwon H., Kim Y., Yoon H., Choi D. Selective audio adversarial example in evasion attack on speech recognition system // *IEEE Transactions on Information Forensics and Security*. 2020. V. 15. P. 526–538. <https://doi.org/10.1109/tifs.2019.2925452>
63. Usama M., Qayyum A., Qadir J., Al-Fuqaha A. Black-box adversarial machine learning attack on network traffic classification // *Proc. of the 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*. 2019. P. 84–89. <https://doi.org/10.1109/iwcmc.2019.8766505>
64. Imam N.H., Vassilakis V.G. A survey of attacks against twitter spam detectors in an adversarial environment // *Robotics*. 2019. V. 8. N 3. P. 50. <https://doi.org/10.3390/robotics8030050>
65. Zhong H., Liao C., Squicciarini A.C., Zhu S., Miller D. Backdoor embedding in convolutional neural network models via invisible perturbation // *Proc. of the Tenth ACM Conference on Data and Application Security and Privacy*. 2020. P. 97–108. <https://doi.org/10.1145/3374664.3375751>
66. Liu X., Yang H., Liu Z., Song L., Li H., Chen Y. Dpatch: An adversarial patch attack on object detectors // *arXiv*. 2018. arXiv:1806.02299. <https://doi.org/10.48550/arXiv.1806.02299>
67. Liu Y., Ma X., Bailey J., Lu F. Reflection backdoor: A natural backdoor attack on deep neural networks // *Lecture Notes in Computer Science*. 2020. V. 12355. P. 182–199. [https://doi.org/10.1007/978-3-030-58607-2\\_11](https://doi.org/10.1007/978-3-030-58607-2_11)
68. Nguyen A., Tran A. WaNet - imperceptible warping-based backdoor attack // *Proc. of the International Conference on Learning Representations*. 2021.
69. Костюмов В.В. Обзор и систематизация атак уклонением на модели компьютерного зрения // *International Journal of Open Information Technologies*. 2022. T. 10. № 10. C. 11–20.
70. Papernot N., McDaniel P., Wu X., Jha S., Swami A. Distillation as a defense to adversarial perturbations against deep neural networks // *Proc. of the 2016 IEEE Symposium on Security and Privacy (SP)*. 2016. P. 582–597. <https://doi.org/10.1109/sp.2016.41>
71. Steihaug T. The conjugate gradient method and trust regions in large scale optimization // *SIAM Journal on Numerical Analysis*. 1983. V. 20. N 3. P. 626–637. <https://doi.org/10.1137/0720042>
72. Curtis A.R., Powell M.J.D., Reid J.K. On the estimation of sparse Jacobian matrices // *IMA Journal of Applied Mathematics*. 1974. V. 13. N 1. P. 117–120. <https://doi.org/10.1093/imamat/13.1.117>
73. Niebur E. Saliency map // *Scholarpedia*. 2007. V. 2. N 8. C. 2675. <https://doi.org/10.4249/scholarpedia.2675>
74. Das S., Suganthan P.N. Differential evolution: A survey of the state-of-the-art // *IEEE Transactions on Evolutionary Computation*. 2011. V. 15. N 1. P. 4–31. <https://doi.org/10.1109/tevc.2010.2059031>
75. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. *Computing Applications (ICIRCA)*, 2021, pp. 584–588. <https://doi.org/10.1109/icirca51532.2021.9544522>
55. Chadha A., Kumar V., Kashyap S., Gupta M. Deepfake: An Overview. *Lecture Notes in Networks and Systems*, 2021, vol. 203, pp. 557–566. [https://doi.org/10.1007/978-981-16-0733-2\\_39](https://doi.org/10.1007/978-981-16-0733-2_39)
56. Nakka K.K., Salzmann M. Indirect local attacks for context-aware semantic segmentation networks. *Lecture Notes in Computer Science*, 2020, vol. 12350, pp. 611–628. [https://doi.org/10.1007/978-3-030-58558-7\\_36](https://doi.org/10.1007/978-3-030-58558-7_36)
57. He Y., Rahimian S., Schiele B., Fritz M. Segmentations-leak: Membership inference attacks and defenses in semantic image segmentation. *Lecture Notes in Computer Science*, 2020, vol. 12368, pp. 519–535. [https://doi.org/10.1007/978-3-030-58592-1\\_31](https://doi.org/10.1007/978-3-030-58592-1_31)
58. Choi J.H. Zhang H., Kim J.H., Hsieh C.J., Lee J.S. Evaluating robustness of deep image super-resolution against adversarial attacks. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 303–311. <https://doi.org/10.1109/iccv.2019.00039>
59. Jiang L., Ma X., Chen S., Bailey J., Jiang Y.G. Black-box adversarial attacks on video recognition models. *Proc. of the 27th ACM International Conference on Multimedia*, 2019, pp. 864–872. <https://doi.org/10.1145/3343031.3351088>
60. Li S., Aich A., Zhu S., Asif S., Song C., Roy-Chowdhury A., Krishnamurthy S. Adversarial attacks on black box video classifiers: Leveraging the power of geometric transformations. *Advances in Neural Information Processing Systems*, 2021, vol. 34, pp. 2085–2096.
61. Chen X., Li S., Huang H. Adversarial attack and defense on deep neural network-based voice processing systems: An overview. *Applied Sciences*, 2021, vol. 11, no. 18, pp. 8450. <https://doi.org/10.3390/app11188450>
62. Kwon H., Kim Y., Yoon H., Choi D. Selective audio adversarial example in evasion attack on speech recognition system. *IEEE Transactions on Information Forensics and Security*, 2020, vol. 15, pp. 526–538. <https://doi.org/10.1109/tifs.2019.2925452>
63. Usama M., Qayyum A., Qadir J., Al-Fuqaha A. Black-box adversarial machine learning attack on network traffic classification. *Proc. of the 15th International Wireless Communications & Mobile Computing Conference (IWCMC)*, 2019, pp. 84–89. <https://doi.org/10.1109/iwcmc.2019.8766505>
64. Imam N.H., Vassilakis V.G. A survey of attacks against twitter spam detectors in an adversarial environment. *Robotics*, 2019, vol. 8, no. 3, pp. 50. <https://doi.org/10.3390/robotics8030050>
65. Zhong H., Liao C., Squicciarini A.C., Zhu S., Miller D. Backdoor embedding in convolutional neural network models via invisible perturbation. *Proc. of the Tenth ACM Conference on Data and Application Security and Privacy*, 2020, pp. 97–108. <https://doi.org/10.1145/3374664.3375751>
66. Liu X., Yang H., Liu Z., Song L., Li H., Chen Y. Dpatch: An adversarial patch attack on object detectors. *arXiv*, 2018, arXiv:1806.02299. <https://doi.org/10.48550/arXiv.1806.02299>
67. Liu Y., Ma X., Bailey J., Lu F. Reflection backdoor: A natural backdoor attack on deep neural networks. *Lecture Notes in Computer Science*, 2020, vol. 12355, pp. 182–199. [https://doi.org/10.1007/978-3-030-58607-2\\_11](https://doi.org/10.1007/978-3-030-58607-2_11)
68. Nguyen A., Tran A. WaNet - imperceptible warping-based backdoor attack. *Proc. of the International Conference on Learning Representations*, 2021.
69. Kostyumov V. A Survey and systematization of evasion attacks in computer vision. *International Journal of Open Information Technologies*, 2022, vol. 10, no. 10, pp. 11–20. (in Russian)
70. Papernot N., McDaniel P., Wu X., Jha S., Swami A. Distillation as a defense to adversarial perturbations against deep neural networks. *Proc. of the 2016 IEEE Symposium on Security and Privacy (SP)*, 2016, pp. 582–597. <https://doi.org/10.1109/sp.2016.41>
71. Steihaug T. The conjugate gradient method and trust regions in large scale optimization. *SIAM Journal on Numerical Analysis*, 1983, vol. 20, no. 3, pp. 626–637. <https://doi.org/10.1137/0720042>
72. Curtis A.R., Powell M.J.D., Reid J.K. On the estimation of sparse Jacobian matrices. *IMA Journal of Applied Mathematics*, 1974, vol. 13, no. 1, pp. 117–120. <https://doi.org/10.1093/imamat/13.1.117>
73. Niebur E. Saliency map. *Scholarpedia*, 2007, vol. 2, no. 8, pp. 2675. <https://doi.org/10.4249/scholarpedia.2675>
74. Das S., Suganthan P.N. Differential evolution: A survey of the state-of-the-art. *IEEE Transactions on Evolutionary Computation*, 2011, vol. 15, no. 1, pp. 4–31. <https://doi.org/10.1109/tevc.2010.2059031>

2017. V. 39. N 12. P. 2481–2495. <https://doi.org/10.1109/tpami.2016.2644615>
76. Lowd D., Meek C. Adversarial learning // Proc. of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining. 2005. P. 641–647. <https://doi.org/10.1145/1081870.1081950>
  77. Xu W., Evans D., Qi Y. Feature squeezing: Detecting adversarial examples in deep neural networks // Proc. of the 2018 Network and Distributed System Security Symposium (NDSS), 2018.
  78. Liao F., Liang M., Dong Y., Pang T., Hu X., Zhu J. Defense against adversarial attacks using high-level representation guided denoiser // Proc. of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018. P. 1778–1787. <https://doi.org/10.1109/cvpr.2018.00191>
  79. Zhang D., Ye M., Gong C., Zhu Z., Liu Q. Black-box certification with randomized smoothing: A functional optimization based framework // Advances in Neural Information Processing Systems. 2020. V. 33. P. 2316–2326.
  80. Fischer M., Baader M., Vechev M. Certified defense to image transformations via randomized smoothing // Advances in Neural Information Processing Systems. 2020. V. 33. P. 8404–8417.
  81. Yang R., Chen X.Q., Cao T.J. APE-GAN++: An improved APE-GAN to eliminate adversarial perturbations // IAENG International Journal of Computer Science. 2021. V. 48. N 3. P. 827–844.
  82. Glenn T.C., Zare A., Gader P.D. Bayesian fuzzy clustering // IEEE Transactions on Fuzzy Systems. 2015. V. 23. N 5. P. 1545–1561. <https://doi.org/10.1109/TFUZZ.2014.2370676>
  83. Plackett R.L. Karl Pearson and the chi-squared test // International Statistical Review / Revue Internationale de Statistique. 1983. V. 51. N 1. P. 59–72. <https://doi.org/10.2307/1402731>
  84. McLachlan G.J. Mahalanobis distance // Resonance. 1999. V. 4. N 6. P. 20–26. <https://doi.org/10.1007/BF02834632>
  75. Badrinarayanan V., Kendall A., Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, vol. 39, no. 12, pp. 2481–2495. <https://doi.org/10.1109/tpami.2016.2644615>
  76. Lowd D., Meek C. Adversarial learning. *Proc. of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, 2005, pp. 641–647. <https://doi.org/10.1145/1081870.1081950>
  77. Xu W., Evans D., Qi Y. Feature squeezing: Detecting adversarial examples in deep neural networks. *Proc. of the 2018 Network and Distributed System Security Symposium (NDSS)*, 2018.
  78. Liao F., Liang M., Dong Y., Pang T., Hu X., Zhu J. Defense against adversarial attacks using high-level representation guided denoiser. *Proc. of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1778–1787. <https://doi.org/10.1109/cvpr.2018.00191>
  79. Zhang D., Ye M., Gong C., Zhu Z., Liu Q. Black-box certification with randomized smoothing: A functional optimization based framework. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 2316–2326.
  80. Fischer M., Baader M., Vechev M. Certified defense to image transformations via randomized smoothing. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 8404–8417.
  81. Yang R., Chen X.Q., Cao T.J. APE-GAN++: An improved APE-GAN to eliminate adversarial perturbations. *IAENG International Journal of Computer Science*, 2021, vol. 48, no. 3, pp. 827–844.
  82. Glenn T.C., Zare A., Gader P.D. Bayesian fuzzy clustering. *IEEE Transactions on Fuzzy Systems*, 2015, vol. 23, no. 5, pp. 1545–1561. <https://doi.org/10.1109/TFUZZ.2014.2370676>
  83. Plackett R.L. Karl Pearson and the chi-squared test. *International Statistical Review / Revue Internationale de Statistique*, 1983, vol. 51, no. 1, pp. 59–72. <https://doi.org/10.2307/1402731>
  84. McLachlan G.J. Mahalanobis distance. *Resonance*, 1999, vol. 4, no. 6, pp. 20–26. <https://doi.org/10.1007/BF02834632>

#### Авторы

**Есипов Дмитрий Андреевич** — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0003-4467-5117>, [some1else.d.ma@gmail.com](mailto:some1else.d.ma@gmail.com)

**Бучаев Абдухамид Яхьяевич** — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57219568840](https://orcid.org/0009-0001-1058-9125), <https://orcid.org/0009-0001-1058-9125>, [abdulhamid0055@yandex.ru](mailto:abdulhamid0055@yandex.ru)

**Керимбай Акылжан** — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0009-9945-9906>, [akerimbai@itmo.ru](mailto:akerimbai@itmo.ru)

**Пузикова Яна Владиславовна** — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0007-7604-3022>, [Yanapuzikova19@ya.ru](mailto:Yanapuzikova19@ya.ru)

**Сайдумаров Семен Кириллович** — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0008-0774-9803>, [semen.say@ya.ru](mailto:semen.say@ya.ru)

**Сулименко Никита Сергеевич** — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0009-0007-3218-9249>, [n.s.sulimenko@mail.ru](mailto:n.s.sulimenko@mail.ru)

**Попов Илья Юрьевич** — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57202195632](https://orcid.org/0000-0002-6407-7934), <https://orcid.org/0000-0002-6407-7934>, [ilyapopov27@gmail.com](mailto:ilyapopov27@gmail.com)

**Кармановский Николай Сергеевич** — кандидат технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57192385103](https://orcid.org/0000-0002-0533-9893), <https://orcid.org/0000-0002-0533-9893>, [karmanov50@mail.ru](mailto:karmanov50@mail.ru)

#### Authors

**Dmitry A. Esipov** — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0003-4467-5117>, [some1else.d.ma@gmail.com](mailto:some1else.d.ma@gmail.com)

**Abdulhamid Y. Buchaev** — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57219568840](https://orcid.org/0009-0001-1058-9125), <https://orcid.org/0009-0001-1058-9125>, [abdulhamid0055@yandex.ru](mailto:abdulhamid0055@yandex.ru)

**Akylzhan Kerimbay** — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0009-9945-9906>, [akerimbai@itmo.ru](mailto:akerimbai@itmo.ru)

**Yana V. Puzikova** — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0007-7604-3022>, [Yanapuzikova19@ya.ru](mailto:Yanapuzikova19@ya.ru)

**Semen K. Saidumarov** — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0008-0774-9803>, [semen.say@ya.ru](mailto:semen.say@ya.ru)

**Nikita S. Sulimenko** — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0009-0007-3218-9249>, [n.s.sulimenko@mail.ru](mailto:n.s.sulimenko@mail.ru)

**Ilya Yu. Popov** — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57202195632](https://orcid.org/0000-0002-6407-7934), <https://orcid.org/0000-0002-6407-7934>, [ilyapopov27@gmail.com](mailto:ilyapopov27@gmail.com)

**Nikolay S. Karmanovskiy** — PhD, Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57192385103](https://orcid.org/0000-0002-0533-9893), <https://orcid.org/0000-0002-0533-9893>, [karmanov50@mail.ru](mailto:karmanov50@mail.ru)

Статья поступила в редакцию 27.03.2023

Одобрена после рецензирования 08.06.2023

Принята к печати 30.07.2023

Received 27.03.2023

Approved after reviewing 08.06.2023

Accepted 30.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-734-742

## Brain MRT image super resolution using discrete cosine transform and convolutional neural network

Pooja Singh<sup>1</sup>✉, Dinesh Ganotra<sup>2</sup>

<sup>1,2</sup> Indira Gandhi Delhi Technical University for Women, New Delhi, 110006, India

<sup>1</sup> [pujasingh0409@gmail.com](mailto:pujasingh0409@gmail.com)✉, <https://orcid.org/0000-0002-8603-5954>

<sup>2</sup> [dinesh\\_ganotra@hotmail.com](mailto:dinesh_ganotra@hotmail.com), <https://orcid.org/0000-0002-3720-8716>

### Abstract

High Resolution (HR) images have numerous applications, such as video conferencing, remote sensing, medical imaging, etc. Furthermore, a few challenges with the super resolution algorithms of magnetic resonance brain images are now obtainable, namely, low sensitivity, significant frequency noise as well as poor resolution. To fix these problems, a Convolutional Neural Network (CNN) based Discrete Cosine Transform (DCT) singular frame quality improvement method is described. There are two stages in this proposed method, involving training and testing. During the training stage, the HR, and Low Resolution (LR) pictures are employed as input, and they are preprocessed to create blocks of images. The histogram and DCT are used for extracting the features from the LR and HR blocks, and these extracted features are assigned with class id. The CNN, which extracts the features and allocates class id, receives its feature extractor as its final input. An LR input image is once more divided into  $[2 \times 2]$  blocks during the testing stage, so each block histogram and DCT feature are estimated. Each feature vector is fed into the neural network as well as the results are contrasted with a set of feature vectors that have been recorded, in addition to the class id that has been allocated to a certain vector. In order to generate a Super resolution image with an LR image, a relevant HR block is then swapped out for this LR block. These results indicated that the initial dataset can achieve 22.4 and 19.5 Peak Signal to Noise Ratio (PSNR) and Root Mean Square Error (RMSE) values while measuring the effectiveness of this proposed method using RMSE and PSNR. Then, the second dataset illustrates that the PSNR and RMSE values are 20.1 and 25.5. For the third dataset, the values are 45.7 and 12.3, respectively. However, the presented method works better than the neural method of Super Resolution Channel Spatial Modulation Network and resolution enhancement technique.

### Keywords

high resolution, low resolution, discrete cosine transform, resolution enhancement, RMSE, PSNR, convolutional neural network

**For citation:** Singh P., Ganotra D. Brain MRT image super resolution using discrete cosine transform and convolutional neural network. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 734–742. doi: 10.17586/2226-1494-2023-23-4-734-742

УДК 004.032.26

## Сверхвысокое разрешение изображения магнитно-резонансной томографии головного мозга с использованием дискретного косинусного преобразования и сверточной нейронной сети

Пуджа Сингх<sup>1</sup>✉, Динеш Ганотра<sup>2</sup>

<sup>1,2</sup> Делийский технический университет Индиры Ганди для женщин, Нью-Дели, 110006, Индия

<sup>1</sup> [pujasingh0409@gmail.com](mailto:pujasingh0409@gmail.com)✉, <https://orcid.org/0000-0002-8603-5954>

<sup>2</sup> [dinesh\\_ganotra@hotmail.com](mailto:dinesh_ganotra@hotmail.com), <https://orcid.org/0000-0002-3720-8716>

### Аннотация

Изображения с высоким разрешением (High Resolution, HR) имеют широкое применение, например при проведении видеоконференций, дистанционного зондирования, медицинской визуализации и других. С использованием алгоритмов сверхвысокого разрешения появилась возможность решить несколько проблем

© Singh P., Ganotra D., 2023

магнитно-резонансной томографии изображений мозга, связанных с низкой чувствительностью, значительным частотным шумом, а также низким разрешением. Чтобы устранить данные проблемы, предложен метод улучшения качества сингулярного кадра на основе сверточной нейронной сети (Convolutional Neural Network, CNN) с дискретным косинусным преобразованием (Discrete Cosine Transform, DCT). Метод состоит из двух этапов, включающих обучение и тестирование. На этапе обучения изображения HR и низкого разрешения (Low Resolution, LR) используются в качестве входных данных и проходят предварительную обработку для создания блоков изображений. Для извлечения признаков из блоков LR и HR применены гистограмма и DCT. Извлеченным признакам присваивается идентификатор класса. CNN извлекает функции DCT, назначает идентификатор класса и получает свой экстрактор функций для окончательного ввода. Входное изображение LR на этапе тестирования повторно делится на блоки  $[2 \times 2]$ , с помощью гистограммы оценивается каждый блок и функции DCT. Каждый вектор признаков передается в нейронную сеть, полученные результаты сравниваются с набором векторов признаков, которые были записаны, в дополнение к идентификатору класса и назначены определенному вектору. Для генерации изображения сверхвысокого разрешения с изображением LR соответствующий блок HR заменяется на блок LR. Полученные результаты показали, что эффективность предложенного метода исходного набора данных достигла значений отношения пикового сигнала к шуму (PSNR) и среднеквадратичной ошибки (RMSE) 22,4 и 19,5 соответственно. Второй набор данных показал значения PSNR и RMSE, равные 20,1 и 25,5, а третий набор — 45,7 и 12,3. Таким образом, представленный метод работает лучше, чем нейронная сеть пространственной модуляции канала сверхвысокого разрешения и метод повышения разрешения.

#### Ключевые слова

высокое разрешение, низкое разрешение, дискретное косинусное преобразование, повышение разрешения, среднеквадратическая ошибка, PSNR, сверточная нейронная сеть

**Ссылка для цитирования:** Сингх П., Ганотра Д. Сверхвысокое разрешение изображения магнитно-резонансной томографии головного мозга с использованием дискретного косинусного преобразования и сверточной нейронной сети // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 734–742 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-4-734-742

## Introduction

High Resolution (HR) images have numerous applications, such as video conferencing, remote sensing, medical imaging, etc. Low Resolution (LR) pictures are converted into HR pictures through an algorithm using resolution improvement techniques [1, 2]. It is relatively easier to retrieve an HR image from numerous LR images rather than from only one LR photo. Computer vision is highly dependent upon the challenge of retrieving an HR image from a single LR image. This problem is known as an under-determined inverse problem having no unique solution. Most often, the resolution to this issue helps to limit the possible solutions using solid previous knowledge.

In most cases, an illustration method is used to compare the most recent and prior province methods. Some methods can focus on an image internal similarity [3] or outside LR and HR pairings to identify mapping operations [4]. These methods based on single-image Super Resolution (SR) are categorized into two classes: sparse coding [5, 6] based on deep learning [7, 8]. Sparse coding is an example of one of the typical external example-based resolution enhancement techniques [9, 10]. The coding coefficients are identical in the patch space of each pair of patches in HR and LR images. The overlapped patch were densely cropped and pre-processed from source image (e.g., subtracting mean and normalization).

A LR dictionary then encodes these patches. An HR dictionary is used to store the sparse coefficients for reconstructing HR patches. To generate the final output, the overlapping reconstructed patches are aggregated. One of the earlier approaches that use Convolutional Neural Network (CNN) for resolution enhancement is Super Resolution Convolutional Neural Network (SRCNN). It is connected to the conventional resolution enhancement method focused on sparse coding, according to the nature

of the network structure. This network comprises three convolutional layers [11]. The scale of the convolution kernel is respectively  $[1 \times 1]$ ,  $[3 \times 3]$ , ...  $[9 \times 9]$ . While SRCNN would be an end-to-end method that proceeds completely forwards, their speed is better compared to conventional techniques.

Deep convolution networks have been shown to produce better results than the conventional resolution enhancement approach [12]. As improved network length, deep neural networks have become more effective. In order to improve the network size, both numbers of layers as well as the units at every layer are grown. Specifically, when working with such a big amount of training data set, this is a rapid and effective method for developing predictive method [13]. If there are only few identified samples in the training set, having a greater variety of parameters makes it harder for the network training to convergence and renders the enlarged networks highly susceptible to overloading.

Direct end-to-end mapping from LR photos to HR images is the target of deep learning based methods. In this paper, the designed method consists of two phases: training and testing. At first, the HR and LR images are taken as input and are preprocessed for dividing the image pixels into blocks such as  $[2 \times 2]$  and  $[4 \times 4]$ . The histogram and Discrete Cosine Transform (DCT) are used for extracting the features from the LR and HR blocks. Each and every extracted feature vector from LR and HR blocks using the DCT is aligned with the class id. In the end, the DCT and histogram characteristics are merged, but this acts as an input for the CNN. Through the use of a class id which is aligned, the features from the input element are extracted to create the SR image by swapping out the LR values into HR values. The LR image should be first compressed to a  $[2 \times 2]$  block in the testing stage, before the histogram and DCT are evaluated, and the LR block is therefore swapped out for the proper HR block to generate the superior image

quality. An earlier resolution improvement method for head imaging used to have a drawback that the proposed scheme has fixed.

The following list contains the main goals of the proposed method.

- LR Magnetic Resonance Tomography (MRT) brain scans can be improved to produce HR MRT scans utilizing CNN.
- Blocks made from the image pixels are employed to effectively features extracted.
- The histogram and DCT methods are used to extract the features in order to minimize the time training in CNN.

The main accomplishments of the proposed method include training the CNN with HR brain MRT images to transform LR MRT brain images into SR MRT brain images. In order to extract features using HR and LR images, the image pixels are split into blocks. Both histogram and DCT method are then used to retrieve the coefficient of feature vectors from LR and HR pictures.

### Literature Review

This section addresses some of the essential studies related to the Brain MRT SR image for transforming the LR image to HR image previously proposed utilizing diverse deep learning approaches.

Dong et al. [14] used an end-to-end mapping between the LR–HR images which is straightforwardly learned in a deep learning method that has been proposed. Considering LR image input and HR image output, this mapping is represented as a deep CNN. They also demonstrated that conventional resolution enhancement methods based on sparse-coding could also be observed as a deep convolutional network. This proposed method optimizes all layers, unlike conventional approaches that treat each variable separately, and proposed deep CNN has a smaller network but shows state-of-the-art quality in restoration. However, CNN end performance is poor due to its capacity limitations.

Aharon et al. [15] proposed a method using the representation of sparse signals for single-image SR. An over-complete dictionary describes image patches using a sparse linear combination of components. For the LR image, sparse representation is calculated for each patch, and the HR image is produced with the help of this coefficient. These outcomes demonstrate that signals that were down-sampled can still accurately retrieve the excess representation. It is possible to determine exactly comparable the dense representation of the LR and HR image patch pairings can be with regard to respective vocabularies through simultaneously training 2 dictionaries for the LR and HR image patches. When compared to earlier approaches, the learned vocabulary pairing samples a greater number of picture patch pairings, which minimizes the computational expense. HR images are generated by using this algorithm that is comparable or better in quality to images formed by other similar resolution enhancement approaches. Furthermore, this approach is inherently robust to noise.

Rueda et al. [16] introduced an SR method based on sparse coding that integrates low and high-frequency

information to generate an HR image from an LR brain MRT image which was adapted for conveniently integrating prior knowledge. With a 3D HR reconstruction, the proposed method considerably improves computationally both speed and precision by incorporating a whole-image multi-scale edge assessment with dimensionality reduction technique. Reconstructed and interpolated reconstructed types of 29 MRT brain images were compared to the novel images learned in a 3T scanner to validate the method, yielding a 70 % reduction in Root Mean Square Error (RMSE) and a 10.3 dB increase in Peak Signal to Noise Ratio (PSNR). Its method outperforms a present state algorithm, demonstrating it will have a big impact on voxel-based morphometry analyses. These approaches are quick and simple to understand and use, but they struggle with edges and complex textures.

Wang and Jiang [17] proposed a research method of the Adaptive Regulation of Thought-Rational resolution enhancement reconstruction method. In order to train a more advanced end-to-end deep threshold recurrent neural network, these authors also added a recurrent layer to the basic CNN during the matching step. During the recruitment process, the neighbourhood embedding approach is employed to constitute for the lost image data. According to the experimental outcomes, the methods has a great potential for photo restoration and has achieved good outcomes in calculating metrics for both PSNR and Structural Similarity Index. However, the quality of reconstruction begins to decline.

Liu et al. [18] have suggested an improvement of image resolution used in medicine. A method was proposed for increasing resolution that first retrieves an HR picture from its counterpart using lower ranking as well as nonlocal self-similar consistency. By iteratively applying a subsampling accuracy restriction on low rank restoration, then they improve the recovered HR value. Results obtained on Magnetic Resonance and Computer Tomography images illustrate that the proposed method works better than standard estimation techniques and therefore it is comparable with current state-of-the-art technology in terms of both measures and picture quality. It results in a large computational cost, though.

Liu et al. [19] had developed a Positron Emission Tomography (PET) Imaging Image Enhancement Using Artificial Intelligence (AI). Excessive noises as well as a restricted spatial resolution are two important influencing factors that limit the qualitative and statistical efficiency of PET scans. AI methods for image denoising and de-blurring are growing in popularity for post-reconstruction enhancement of PET images. The proposed method offers a comprehensive analysis of prior AI-based PET image enhancement efforts, with a focusing on network architectures, file formats, loss functions and evaluation criteria.

According to the mentioned study, brain imaging super resolution faces a number of important difficulties. Due to its capacity limitations, EES end performance is poor (Dong et al. [14]). This approach is inherently robust to noise Aharon et al. [15], quick and simple to understand and use, but they struggle with edges and complex textures used by Rueda et al. [16], and the quality of reconstruction

begins to decline Wang et al. [17]; it leads to a high computational cost of Liu H. et al. [18]. CNN was proposed for brain image SR as a solution to this issue.

### Proposed Method for Brain Image Super Resolution

Resolution is generated by image sensors or image collection hardware. The size or quantity of sensor elements affects an images spatial resolution. Converting LR photos to HR is termed as brain image improvement. For improving picture quality from extracted features in this method, CNN is suggested. Video conferencing, remote sensing, medical imaging, and several other uses are available for HR images. LR photos are transformed into HR images by an algorithm utilizing resolution enhancement techniques. A CNN is suggested in this method because previous Brain-computer interface technology mainly depends on artificial stimuli for resolution improvement.

The process flow of the proposed method is illustrated in Fig. 1. The HR  $[256 \times 256]$  and LR  $[128 \times 128]$  images are initially selected as input for the training phase, then they pre-processed to split the image pixels into blocks. After that, the histogram and DCT methods are utilized to extract the features from HR and LR blocks. The CNN receives these extracted features as input and utilizes them to extract the features and establish a class id. The histogram and DCT feature of an LR input image were computed for each block of the source images throughout testing. While feeding the neural network with this feature data, the output is assessed against a previously established recorded feature vector which is given a class id. In order to create a SR image from an LR image, each LR block is again swapped out with the equivalent HR block. Three distinct dataset of brain MRT images are utilized to test the proposed method.

#### Data Augmentation

When including slightly modified versions of either existing data or brand-new synthetic data that is derived from previous data, data augmentation in data analysis is utilized to expand the quantity of data. Less data is employed in the datasets for such method design for

training and testing the classifier. Consequently, the data augmentation technique increases the dataset size.

#### Pre-processing

The pixels are divided into blocks using pre-processing in this method. HR and LR images are utilized for the pre-processing for it. From the first dataset<sup>1</sup>. LR images of size  $[128 \times 128]$  with matching HR images of size  $[256 \times 256]$  from the SR of Brain MRT Images Dataset Download was employed for training. The LR image is separated into blocks of size  $[2 \times 2]$  but each HR image is split into blocks of size  $[4 \times 4]$  pixels, obtaining 4096 blocks of  $[2 \times 2]$  as well as  $[4 \times 4]$  pixel resolution by each LR and HR image, correspondingly. For 21 LR images total  $[21 \times 4096]$  blocks were used for training.

Another dataset: 1 of 49 images was used for training method land from another database  $[2 \times 46]$  images were used for training, resulting in  $[49 \times 4096]$  and  $[46 \times 4096]$  blocks respectively.

#### 1. Feature extraction

The process of feature extraction is fundamental to the improvement of images. These features are extracted of the pre-processed images using the DCT method. Each LR and HR block was utilized to extract the histogram with 2D DCT. Even though the input to the 2D DCT was a matrix of size  $[2 \times 2]$ , the output of the 2D DCT function was chosen as a matrix of size  $[3 \times 3]$  to improve the image quality and reduce the complexity. For the histogram, 16 bins were chosen. This shape of each HR block now consists of a feature vector with a size of  $[1 \times 25]$ . The sixteen bins were from 0–15, 16–31, ..., 240–255 grayscale values.

It is possible to add cosine functions which fluctuate at specific frequencies to describe a continuous sequence of points using a DCT. Extraction of DCT features occurs in two phases. First, the entire image is subjected to the DCT to generate DCT coefficients, while in the following phase, feature vectors are built using a subset of the coefficients that have been chosen in the initial process. Approximately the same size to the input image is the DCT coefficient matrix [20].

<sup>1</sup> Available at: <https://projecttunnel.com/Super-Resolution-of-Brain-MRI-Images-Dataset-Download> (accessed: 14.07.2023).

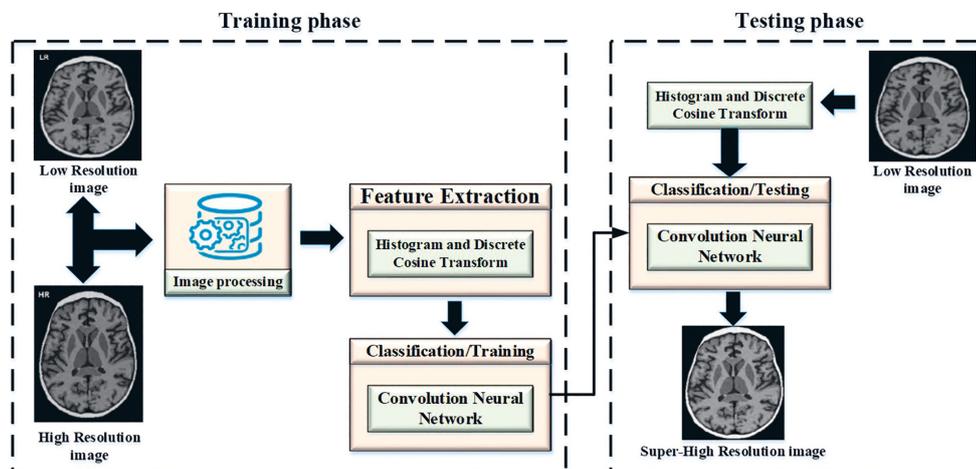


Fig. 1. Architecture of the proposed method

For image of  $M \times N$  size, in which each photo relates to a 2D matrix, DCT coefficients are calculated.

$$\mathbf{F}(u, v) = \frac{1}{\sqrt{MN}} \alpha(u) \alpha(v) \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \times \cos\left(\frac{(2x+1)u\pi}{2M}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right),$$

where  $u = 0, 1, \dots, M-1$  and  $v = 0, 1, \dots, N-1$ ,  $\alpha(w)$  is defined by,

$$\alpha(w) = \begin{cases} \frac{1}{\sqrt{2}}, & w = 0 \\ 1, & \text{otherwise} \end{cases}$$

The image brightness function is known as  $f(x, y)$ , and the 2D matrix  $\mathbf{F}(u, v)$  includes the DCT coefficients. Applying the DCT to the entire image gives the frequencies coefficients matrices of the same dimensions. A method is utilized for feature extraction that has a significant impact on accurate identification is the second phase. The purpose of selecting the coefficient is used to minimise the reconstruction error.

## 2. Convolutional Neural Network

In this method, CNN is proposed for extracting features from the previously extracted image. After using CNN for extracting the features, the LR values are replaced by the HR values. An already-trained version of the networks, which has been trained on more than a million pictures, can be loaded from the ImageNet dataset. Evaluate this proposed method to handle recognizing various kinds of sign images.

As inputs to the Convolution layers, the feature vector of dimension  $[1 \times 25]$  was utilized. Using  $[2 \times 2]$  kernel in the CNN layer,  $[4 \times 4]$  feature vector is obtained. This feature vector sent to the max-pooling layer results in  $[3 \times 3]$  size. After CNN, each feature vector has size  $[3 \times 3]$ . Each corresponding input and output feature vector pair were assigned a number referred to as class id. For the same feature vector, the same class id is assigned.

The convolution layer, which is the initial layer of the CNN, executes a convolution operation on the input, as illustrated in Fig. 2. A total result is then filtered using an activation function, such as Rectified Linear Unit, and sent to the pooling layer which is the subsequent layer. The following equation will be utilized to estimate the convolution layer operation [21]:

$$O_{i(x,y)} = f' \left( \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} w(i,j) I(x+i, y+j) + bias \right),$$

where  $O$  is the basis function of fully connected layer;  $f'$  is the representation of function;  $I$  is the representation of input image;  $x$  and  $y$  are the mapping row and column of image;  $m$  is the maximum size of filter in coevolution;  $bias$  is the bias value.

This equation includes map dimension as well as the network parameters which are minimized through the use of the pooling layer. The features, such as edges and points, are extracted using max pooling. In order to link the neurons from the current layer to the following layer, a completely connected layer is therefore used. An input image is classified into different classes using the training dataset. The neural network output  $y$  is identified as in [21]

$$y = \sigma(w^L \dots \sigma(w^2 \sigma(w^1 x + b^1) + b^2) \dots + b^L),$$

where  $\sigma$  is an activation function,  $w$  is a network parameter,  $b$  is a bias, and  $x$  is input,  $L$  is layer number.

## 3. Training

Depending on the class id of the feature vector for LR block, blocks were recognized during training. The trained vector, which has a size of  $[86,016 \times 9]$  for the primary dataset, has size  $[200,704 \times 9]$  for second dataset and  $[188,416 \times 9]$  for third dataset, is created by obtaining the class ids for the LR and HR photo blocks.

A histogram feature with 16 bins and a DCT of 9 values are used in the proposed method. Thus, we have an input vector of 25 elements. This 25-element vector is placed through a CNN method, and the resulting 9-element vector includes what the initial hidden layer will use as input. This output of the initial hidden layer has 128 items and also the 2 items hidden layer includes another 9 elements. Finally, there is only one output element in the output layer which denotes the class id. At the input to the hidden layer and the hidden layer output layer, sigmoid functions are employed as activation functions. In order to identify the correct class id for the LR block, the neural network obtains its inputs feature vector. It is replaced with an HR block that matched this LR block.

## 4. Testing

This stage again divides the LR input image into  $[2 \times 2]$  blocks for testing, so each block histogram and DCT features are obtained.

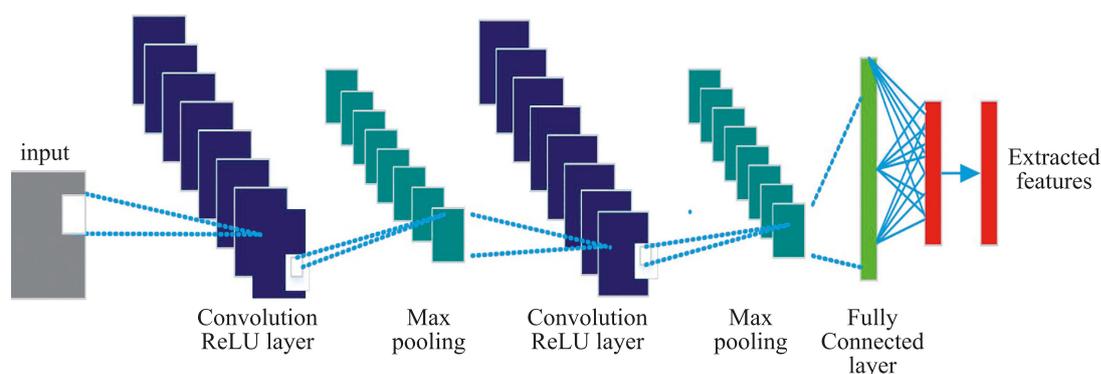


Fig. 2. Generic architecture of Convolutional Neural Network

The class id is given to the feature vector when the outputs of the neural network are evaluated with a group of feature vectors that have been recorded. In order to create an HR image with an LR image, an appropriate HR block is then exchanged for this LR block. Ten pictures from the initial dataset were utilized to test the proposed method. 40 images were used to test the method for the second and third datasets.

### Result and Discussion

The purpose of this method is to transform a low quality picture into a high-quality image. This section assesses the CNN. MATLAB 2021a is utilized to execute the testing, with an Intel Core i5, NVidia GeForce GTX 1650 and 16GB of RAM. In order to evaluate if it transforms the LR image into an HR image, CNN examines the data entered. Whenever the data was first collected, it was utilized in two phases: training and testing. The CNN classifier in the present method receives its input from the DCT features that have been retrieved.

Both LR and HR pictures are utilized as input initially, then pre-processing follows. Pre-processing is used to decompose picture pixels into blocks, while extracted features serves to feature extraction from the pre-processed pictures discrete cosine transform. Block splitting is employed by the proposed method for both the HR and LR images which have pixel sizes of  $[256 \times 256]$  and  $[128 \times 128]$ , respectfully. Different class ids are given after extracting the histogram and DCT features from such LR and HR blocks. These feature vectors are used to train the method to identify the relevant class id. CNN is utilized to train the method. Three distinct datasets of brain MRT images are employed to test the proposed framework. The proposed method results are compared to Super Resolution Channel Spatial (SRCS) and neural network-based resolution enhancement. The performance of this proposed method with neural network-based learning and the SRCS method was calculated using three datasets of brain MRT images.

#### Dataset description

Dataset 1: There are 62 photographs in this set, including 31 images in LR and 31 in HR. Each low-resolution image is  $[128 \times 128]$  pixels in size<sup>1</sup>. Similarly, each HR image is  $256 \times 256$  pixels in size.

Dataset 2: The data are collected from the Kaggle dataset. There are 155 non-tumorous data images in this dataset, and 98 tumorous images<sup>2</sup>.

Dataset 3: the data are collected from the figshare dataset. In this dataset, 86 images are used<sup>3</sup>.

The above chosen datasets consist of 31, 89 and 86 images of LR for training and testing using the classifier. These fewer amounts of data are not feasible for training

the CNN. Instead of searching for a new dataset, the collected dataset is augmented 10 times, resulting in more data for training the CNN.

In Table 1, in the first dataset totally of 31 images are used for training and testing. For training, 21 images are used, and 10 images are taken from the total images for testing purposes.

In Fig. 3, from the first dataset of 31 brain MRT images, 21 were used as training, and 10 were used for the testing. In comparison to the SRCS and neural network based method, the PSNR values and on average, raised by 6 dB and 9.6 dB, respectively, based on the outcomes of these 10 testing images. RMSE is decreased by 12.8 dB and 22.2 dB than the neural network based and SRCS methods.

The average RMSE and PSNR of 10 test images are shown in Table 2. For the parameter PSNR, the SRCS value is 12.8, neural network based value is 18.8, and our method reaches 22.4. For RMSE, the value of SRCS is 41.7, the neural network based value reaches 28.9, and the proposed method reaches 19.5, respectively.

In Table 3, in the first dataset totally 89 images are used for training and testing: 49 images are used for training,

Table 1. First dataset description

Data Augmentation	Total number of images	Images used for training	Images used for testing
Before	31	21	10
After	310	248	62

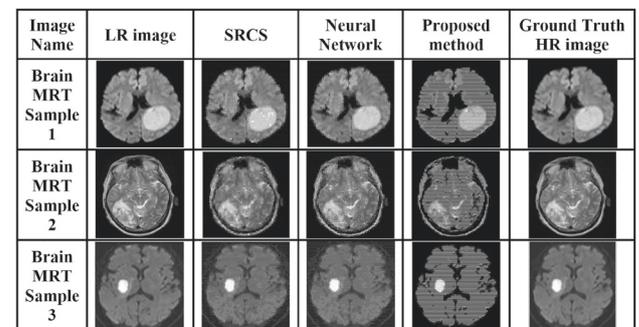


Fig. 3. Results of various techniques for Dataset 1

Table 2. Average RMSE and PSNR of 10 test images from dataset 1

Parameter	SRCS	Neural Network	Proposed Method
PSNR	12.8	18.8	22.4
RMSE	41.7	28.9	19.5

Table 3. Second dataset description

Data Augmentation	Total number of images	Images used for training	Images used for testing
Before	89	49	40
After	890	712	178

<sup>1</sup> Dataset 2: <https://projecttunnel.com/Super-Resolution-of-Brain-MRI-Images-Dataset-Download> (accessed: 14.07.2023).

<sup>2</sup> Chakrabarty N. (2019, April 14). Brain MRI images for Brain tumor detection. Kaggle. Available at: <https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection> (accessed: 14.07.2023).

<sup>3</sup> Dataset 3: [https://figshare.com/articles/dataset/brain\\_tumor\\_dataset/1512427](https://figshare.com/articles/dataset/brain_tumor_dataset/1512427) (accessed: 14.07.2023).

and 40 images are taken from the total images for testing purposes.

In Fig. 4, for the second dataset of 89 images, 49 were used for training the network, and the proposed method was tested on 40 images. After averaging the RMSE and PSNR values across the 40 images, it was found that the RMSE is lower by 2.2 dB and 9.3 dB than the neural network based method as well as SRCS method, respectively, and also the PSNR values are greater by 20.6 dB and 23.6 dB than the neural network based method and SRCS, respectively.

The average RMSE and PSNR of 40 test images are shown in Table 4. For the parameter PSNR, the SRCS value is 10.8, neural network based value is 16, and our method reaches 20.1. For RMSE, the value of SRCS is 49.1, the neural network based value reaches 34.5, and the proposed method reaches 25.5, respectively.

In Table 5, in the first dataset totally 86 images are used for training and testing: 46 images are used for training, and 40 images are taken from the total images for testing purposes [22].

In Fig. 5, for the third dataset having 86 images, 46 were used for training purposes and testing was done on 40 images. The outcomes of these 40 test images were averaged, while it was found that the RMSE was reduced by 4.4 dB and 9.7 dB compared to the neural network-based approach and also the SRCS method, respectively, and thus the PSNR values significantly improved by 10.5 dB and 26.4 dB. DCT and histogram features were used to enhance the proposed method accuracy through training.

The average RMSE and PSNR of 40 test images are shown in Table 6. For the parameter PSNR, the SRCS value is 19.3, neural network based value is 35.2, and our method reaches 45.7. For RMSE, the value of SRCS is 22.0, the neural network based reaches 16.7, and the proposed method reaches 12.3, respectively.

Table 7 shows the comparison of PSNR and RMSE for proposed and existing approach. Proposed approach attain 45.7 % PSNR value which is high when compared

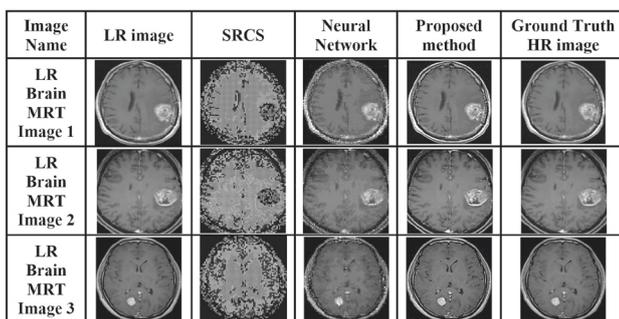


Fig. 4. Results of various techniques for Dataset 2

Table 4. Average RMSE and PSNR of 40 test images from dataset 2

Parameter	SRCS	Neural Network	Proposed Method
PSNR	10.8	16	20.1
RMSE	49.1	34.5	25.5

Table 5. Third dataset description

Data Augmentation	Total number of images	Images used for training	Images used for testing
Before	86	46	40
After	860	688	172

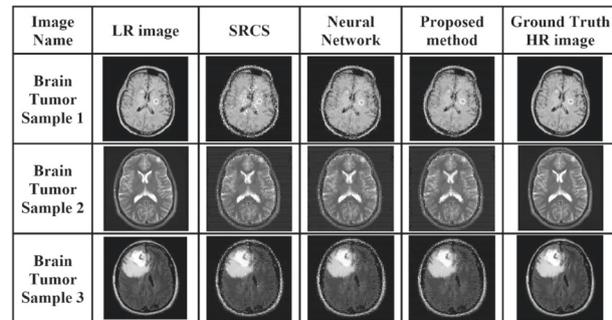


Fig. 5. Results of various techniques for Dataset 3

Table 6. Average RMSE and PSNR of 40 test images from dataset 3

Parameter	SRCS	Neural Network	Proposed Method
PSNR	19.3	35.2	45.7
RMSE	22.0	16.7	12.3

Table 7. Average RMSE and PSNR for proposed and existing approaches

Parameter	Proposed method	DBPN [23]	CSAM [24]	DRPB [25]
PSNR	45.7	31.86	31.42	30.17
RMSE	12.3	17.8	22.4	28.11

to other current approaches, such as Dual residual Path Block (DBPN), Channel Spatial Attention Module (CSAM) and Deep Back Projection Network (DRPB) whose values are 31.86 %, 31.42 % and 30.17 %, respectively. RMSE values for proposed approach is 12.3 % which is low when compared to the current approaches. It clearly shows that the proposed approach yield better performance when compared to other current approaches.

## Conclusion

In this research, DCT and CNN based resolution improvement method called "Brain Image" was developed. CNN is utilized to extract the features and change the LR value with the HR value. The HR image and LR images are divided into smaller blocks. The histogram and DCT features are extracted for each HR and LR block. The method is trained using these extracted features, and the unique class id is allotted to each feature vector. The proposed method attained 22.4 and 19.5 PSNR and RMSE values. And for the second dataset, the PSNR and RMSE

values are 20.1 and 25.5. For the third dataset, the values are 45.7 and 12.3. The analysis of three various datasets of brain MRT images reveals comparisons between the proposed approach and SRCS and resolution improvement using a neural network. The RMSE and PSNR values are

improved as compared to these two methods. As a result, the proposed approach may be a good improvement over the methods already being used. The processing of large-scale medical images and pixel fusion picture registration should be enhanced in future work.

## References

1. Chen Q., Huang J., Feris R., Brown L.M., Dong J., Yan S. Deep domain adaptation for describing people based on fine-grained clothing attributes. *Proc. of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 5315–5324. <https://doi.org/10.1109/cvpr.2015.7299169>
2. Denton E.L., Chintala S., Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks. *Advances in Neural Information Processing Systems*, 2015, vol. 28.
3. Cui Z., Chang H., Shan S., Zhong B., Chen X. Deep network cascade for image super-resolution. *Lecture Notes in Computer Science*, 2014, vol. 8693, pp. 49–64. [https://doi.org/10.1007/978-3-319-10602-1\\_4](https://doi.org/10.1007/978-3-319-10602-1_4)
4. Farhadifard F., Abar E., Nazzal M., Ozkaraman H. Single image super resolution based on sparse representation via directionally structured dictionaries. *Proc. of the 22<sup>nd</sup> Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2014, pp. 1718–1721. <https://doi.org/10.1109/siu.2014.6830580>
5. Ahmed J., Memon R.A., Waqas M., Mangrio M.I., Ali S. Selective sparse coding based coupled dictionary learning algorithm for single image super-resolution. *Proc. of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, 2018, pp. 1–5. <https://doi.org/10.1109/icomet.2018.8346357>
6. Choi J.H., Kim J.H., Cheon M., Lee J.S. Deep learning-based image super-resolution considering quantitative and perceptual quality. *Neurocomputing*, 2020, vol. 398, pp. 347–59. <https://doi.org/10.1016/j.neucom.2019.06.103>
7. Dong C., Loy C.C., He K., Tang X. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, vol. 38, no. 2, pp. 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>
8. Dong C., Loy C.C., Tang X. Accelerating the super-resolution convolutional neural network. *Lecture Notes in Computer Science*, 2016, vol. 9906, pp. 391–407. [https://doi.org/10.1007/978-3-319-46475-6\\_25](https://doi.org/10.1007/978-3-319-46475-6_25)
9. Ayas S., Ekinici M. Single image super resolution using dictionary learning and sparse coding with multi-scale and multi-directional Gabor feature representation. *Information Sciences*, 2020, vol. 512, pp. 1264–1278. <https://doi.org/10.1016/j.ins.2019.10.040>
10. Gu S., Zuo W., Xie Q., Meng D., Feng X., Zhang L. Convolutional sparse coding for image super-resolution. *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1823–1831. <https://doi.org/10.1109/iccv.2015.212>
11. Dosovitskiy A., Springenberg J.T., Brox T. Learning to generate chairs with convolutional neural networks. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1538–1546. <https://doi.org/10.1109/cvpr.2015.7298761>
12. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error. *Proc. of the 4<sup>th</sup> International Conference on Learning Representations (ICLR)*, 2016.
13. Alec R., Luke M., Soumith C. Unsupervised representation learning with deep convolutional generative adversarial networks. *Proc. of the International Conference on Learning Representations (ICLR)*, 2015, pp. 1–16.
14. Dong C., Loy C.C., He K., Tang X. Learning a deep convolutional network for image super-resolution. *Lecture Notes in Computer Science*, 2014, vol. 8692, pp. 184–199. [https://doi.org/10.1007/978-3-319-10593-2\\_13](https://doi.org/10.1007/978-3-319-10593-2_13)
15. Aharon M., Elad M., Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, vol. 54, no. 11, pp. 4311–4322. <https://doi.org/10.1109/tsp.2006.881199>
16. Rueda A., Malpica N., Romero E. Single-image super-resolution of brain MR images using overcomplete dictionaries. *Medical Image Analysis*, 2013, vol. 17, no. 1, pp. 113–132. <https://doi.org/10.1016/j.media.2012.09.003>

## Литература

1. Chen Q., Huang J., Feris R., Brown L.M., Dong J., Yan S. Deep domain adaptation for describing people based on fine-grained clothing attributes // *Proc. of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015. P. 5315–5324. <https://doi.org/10.1109/cvpr.2015.7299169>
2. Denton E.L., Chintala S., Fergus R. Deep generative image models using a Laplacian pyramid of adversarial networks // *Advances in Neural Information Processing Systems*. 2015. V. 28.
3. Cui Z., Chang H., Shan S., Zhong B., Chen X. Deep network cascade for image super-resolution // *Lecture Notes in Computer Science*. 2014. V. 8693. P. 49–64. [https://doi.org/10.1007/978-3-319-10602-1\\_4](https://doi.org/10.1007/978-3-319-10602-1_4)
4. Farhadifard F., Abar E., Nazzal M., Ozkaraman H. Single image super resolution based on sparse representation via directionally structured dictionaries // *Proc. of the 22<sup>nd</sup> Signal Processing and Communications Applications Conference (SIU)*. IEEE, 2014. P. 1718–1721. <https://doi.org/10.1109/siu.2014.6830580>
5. Ahmed J., Memon R.A., Waqas M., Mangrio M.I., Ali S. Selective sparse coding based coupled dictionary learning algorithm for single image super-resolution // *Proc. of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. 2018. P. 1–5. <https://doi.org/10.1109/icomet.2018.8346357>
6. Choi J.H., Kim J.H., Cheon M., Lee J.S. Deep learning-based image super-resolution considering quantitative and perceptual quality // *Neurocomputing*. 2020. V. 398. P. 347–59. <https://doi.org/10.1016/j.neucom.2019.06.103>
7. Dong C., Loy C.C., He K., Tang X. Image super-resolution using deep convolutional networks // *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2016. V. 38. N 2. P. 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>
8. Dong C., Loy C.C., Tang X. Accelerating the super-resolution convolutional neural network // *Lecture Notes in Computer Science*. 2016. V. 9906. P. 391–407. [https://doi.org/10.1007/978-3-319-46475-6\\_25](https://doi.org/10.1007/978-3-319-46475-6_25)
9. Ayas S., Ekinici M. Single image super resolution using dictionary learning and sparse coding with multi-scale and multi-directional Gabor feature representation // *Information Sciences*. 2020. V. 512. P. 1264–1278. <https://doi.org/10.1016/j.ins.2019.10.040>
10. Gu S., Zuo W., Xie Q., Meng D., Feng X., Zhang L. Convolutional sparse coding for image super-resolution // *Proc. of the IEEE International Conference on Computer Vision (ICCV)*. 2015. P. 1823–1831. <https://doi.org/10.1109/iccv.2015.212>
11. Dosovitskiy A., Springenberg J.T., Brox T. Learning to generate chairs with convolutional neural networks // *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015. P. 1538–1546. <https://doi.org/10.1109/cvpr.2015.7298761>
12. Mathieu M., Couprie C., LeCun Y. Deep multi-scale video prediction beyond mean square error // *Proc. of the 4<sup>th</sup> International Conference on Learning Representations (ICLR)*. 2016.
13. Alec R., Luke M., Soumith C. Unsupervised representation learning with deep convolutional generative adversarial networks // *Proc. of the International Conference on Learning Representations (ICLR)*. 2015. P. 1–16.
14. Dong C., Loy C.C., He K., Tang X. Learning a deep convolutional network for image super-resolution // *Lecture Notes in Computer Science*. 2014. V. 8692. P. 184–199. [https://doi.org/10.1007/978-3-319-10593-2\\_13](https://doi.org/10.1007/978-3-319-10593-2_13)
15. Aharon M., Elad M., Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation // *IEEE Transactions on Signal Processing*. 2006. V. 54. N 11. P. 4311–4322. <https://doi.org/10.1109/tsp.2006.881199>
16. Rueda A., Malpica N., Romero E. Single-image super-resolution of brain MR images using overcomplete dictionaries // *Medical Image Analysis*. 2013. V. 17. N 1. P. 113–132. <https://doi.org/10.1016/j.media.2012.09.003>

17. Wang H., Jiang K. Research on image super-resolution reconstruction based on transformer. *Proc. of the 2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID)*, 2021, pp. 226–230. <https://doi.org/10.1109/aiid51893.2021.9456580>
18. Liu H., Guo Q., Wang G., Gupta B.B., Zhang C. Medical image resolution enhancement for healthcare using nonlocal self-similarity and low-rank prior. *Multimedia Tools and Applications*, 2019, vol. 78, no. 7, pp. 9033–9050. <https://doi.org/10.1007/s11042-017-5277-6>
19. Liu J., Malekzadeh M., Mirian N., Song T.A., Liu C., Dutta J. Artificial intelligence-based image enhancement in PET imaging: Noise reduction and resolution enhancement. *PET Clinics*, 2021, vol. 16, no. 4, pp. 553–576. <https://doi.org/10.1016/j.cpet.2021.06.005>
20. Dabbaghchian S., Ghaemmaghami M.P., Aghagolzadeh A. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology. *Pattern Recognition*, 2010, vol. 43, no. 4, pp. 1431–1440. <https://doi.org/10.1016/j.patcog.2009.11.001>
21. Liew W.S., Tang T.B., Lin C.H., Lu C.K. Automatic colonic polyp detection using integration of modified deep residual convolutional neural network and ensemble learning approaches. *Computer Methods and Programs in Biomedicine*, 2021, vol. 206, pp. 106114. <https://doi.org/10.1016/j.cmpb.2021.106114>
22. Timofte R., De V., Van Gool L. Anchored neighborhood regression for fast example-based super-resolution. *Proc. of the IEEE International Conference on Computer Vision*, 2013, pp. 1920–1927. <https://doi.org/10.1109/iccv.2013.241>
23. Haris M., Shakhnarovich G., Ukita N. Deep back-project networks for single image super-resolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, vol. 43, no. 12, pp. 4323–4337. <https://doi.org/10.1109/tpami.2020.3002836>
24. Niu B., Wen W., Ren W., Zhang X., Yang L., Wang S., Zhang K., Cao X., Shen H. Single image super-resolution via a holistic attention network. *Lecture Notes in Computer Science*, 2020, vol. 12357, pp. 191–207. [https://doi.org/10.1007/978-3-030-58610-2\\_12](https://doi.org/10.1007/978-3-030-58610-2_12)
25. Lan R., Sun L., Liu Z., Lu H., Pang C., Luo X. MADNet: a fast and lightweight network for single-image super resolution. *IEEE Transactions on Cybernetics*, 2021, vol. 51, no. 3, pp. 1443–1453. <https://doi.org/10.1109/tcyb.2020.2970104>

#### Authors

**Pooja Singh** — Magister, Researcher, Indira Gandhi Delhi Technical University for Women, New Delhi, 110006, India, [sc 57225030639](https://orcid.org/0000-0002-8603-5954), <https://orcid.org/0000-0002-8603-5954>, [pujasingh0409@gmail.com](mailto:pujasingh0409@gmail.com)

**Dinesh Ganotra** — PhD, Associate Professor, Indira Gandhi Delhi Technical University for Women, New Delhi, 110006, India, [sc 6506229541](https://orcid.org/0000-0002-3720-8716), <https://orcid.org/0000-0002-3720-8716>, [dinesh\\_ganotra@hotmail.com](mailto:dinesh_ganotra@hotmail.com)

Received 23.01.2023

Approved after reviewing 12.06.2023

Accepted 24.07.2023

#### Авторы

**Сингх Пуджа** — магистр, исследователь, Делийский технический университет Индиры Ганди для женщин, Нью-Дели, 110006, Индия, [sc 57225030639](https://orcid.org/0000-0002-8603-5954), <https://orcid.org/0000-0002-8603-5954>, [pujasingh0409@gmail.com](mailto:pujasingh0409@gmail.com)

**Ганотра Динеш** — PhD, доцент, Делийский технический университет Индиры Ганди для женщин, Нью-Дели, 110006, Индия, [sc 6506229541](https://orcid.org/0000-0002-3720-8716), <https://orcid.org/0000-0002-3720-8716>, [dinesh\\_ganotra@hotmail.com](mailto:dinesh_ganotra@hotmail.com)

Статья поступила в редакцию 23.01.2023

Одобрена после рецензирования 12.06.2023

Принята к печати 24.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-743-749

УДК 004.89

## Метод аугментации текстовых данных с сохранением стиля речи и лексики персоны

Анастасия Андреевна Матвеева<sup>1</sup>✉, Олеся Владимировна Махныткина<sup>2</sup>

<sup>1,2</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>1</sup> [anastasiamatveevaitmo@gmail.com](mailto:anastasiamatveevaitmo@gmail.com)✉, <https://orcid.org/0000-0002-2935-991X>

<sup>2</sup> [makhnytkina@itmo.ru](mailto:makhnytkina@itmo.ru), <https://orcid.org/0000-0002-8992-9654>

### Аннотация

**Введение.** В настоящее время часто для различных задач обработки естественного языка требуются большие наборы данных. Однако для многих задач сбор большого набора данных является трудоемким, дорогостоящим процессом и требует привлечения экспертов. Увеличение объема информации возможно достичь с использованием методов аугментации данных. Вместе с тем использование классических подходов может привести к включению в корпус данных фраз, которые отличаются по стилю речи и лексикону целевой персоны. Это сопровождается изменением целевого класса и появлением реплик с неестественным использованием лексики и отсутствием смысла. Предложен новый метод аугментации текстовых данных, учитывающий стиль и лексикон персоны. **Метод.** В работе разработан новый метод аугментации текстовых данных, сохраняющий индивидуальные речевые характеристики и словарный запас. Основная идея метода заключается в формировании индивидуальных шаблонов для каждого человека на основе анализа синтаксических деревьев высказываний и последующего создания новых реплик по сформированным шаблонам. **Основные результаты.** Метод апробирован на задаче оценки эмоционального состояния пользователя в диалоге. Исследования проведены для датасетов на английском и русском языках. Предложенный метод позволил повысить качество решения этих задач как для английского, так и для русского языков. Получено увеличений значений метрик accuracy и weighted F1 для разных моделей. **Обсуждение.** Результаты работы могут быть применены для повышения метрик accuracy и weighted F1 моделей, предназначенных для решения различных задач для английского и русского языков.

### Ключевые слова

аугментация текстовых данных, распознавание эмоций, оценка валентности высказываний

### Благодарности

Исследование выполнено за счет гранта Российского научного фонда (№ 22-11-00128, <https://www.rscf.ru/project/22-11-00128/>).

**Ссылка для цитирования:** Матвеева А.А., Махныткина О.В. Метод аугментации текстовых данных с сохранением стиля речи и лексики персоны // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 743–749. doi: 10.17586/2226-1494-2023-23-4-743-749

## Text augmentation preserving persona speech style and vocabulary

Anastasia A. Matveeva<sup>1</sup>✉, Olesia V. Makhnytkina<sup>2</sup>

<sup>1,2</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>1</sup> [anastasiamatveevaitmo@gmail.com](mailto:anastasiamatveevaitmo@gmail.com)✉, <https://orcid.org/0000-0002-2935-991X>

<sup>2</sup> [makhnytkina@itmo.ru](mailto:makhnytkina@itmo.ru), <https://orcid.org/0000-0002-8992-9654>

### Abstract

Currently, various natural language processing tasks often require large data sets. However, for many tasks, collecting large datasets is quite tedious and expensive, and requires the involvement of experts. An increase in the amount of data can be achieved using methods of data augmentation, however, the use of classical approaches can lead to the inclusion of phrases in the data corpus that differ in the speech style and vocabulary of the target person, which can lead to both

a change in the target class as well as the appearance of replicas with unnatural vocabulary use and lack of meaning. In this context, a new method for test data enrichment is proposed that takes into account the person's style and vocabulary. In this article, a new method for expanding text data that preserves individual language features and vocabulary is proposed. The core of the method is to create individual templates for each person based on the analysis of syntactic trees of propositions and then to create new replicas according to the generated templates. The method was tested on the task of assessing the user's emotional state in a dialogue. The search was carried out for data sets in English and Russian. The proposed method made it possible to improve the quality of solving these problems for both the English and Russian languages. Up to a 2 % increase in accuracy and weighted F1 metrics has been noted for various models. The results of the work can be applied to improve the accuracy and weighted F1 metrics of models designed to solve various problems for the English and Russian languages.

#### Keywords

text data augmentation, emotion recognition, statement valence evaluation

#### Acknowledgements

This research was supported by a grant from the Russian Science Foundation (22-11-00128 <https://www.rscf.ru/project/22-11-00128/>).

**For citation:** Matveeva A.A., Makhnytkina O.V. Text augmentation preserving persona speech style and vocabulary. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 743–749 (in Russian). doi: 10.17586/2226-1494-2023-23-4-743-749

### Введение

В настоящее время активно развивается направление обработки естественного языка. В основном повышение качества решения задач происходит за счет появления более сложных и глубоких архитектур нейронных сетей, требующих больших датасетов для обучения. Использование предобученных моделей и их дообучение на целевых датасетах зависит от объема целевых датасетов. Для получения качественных моделей необходимо наличие больших датасетов. Увеличение набора данных возможно с использованием различных методов аугментации текста.

Аугментация текстовых данных может осуществляться на двух уровнях: текста и векторных представлений слов и реплик.

Аугментация на уровне векторных представлений слов и реплик подразумевает модификацию не текстовых сообщений, а векторов слов, полученных с использованием различных языковых моделей. В работах [1, 2] рассмотрены методы аугментации на уровне векторных представлений слов: интерполяции, экстраполяции, добавления случайного шума. Различные варианты добавления шума в векторные представления слов, полученные с использованием модели Word2Vec, исследованы в работе [3]: гауссовский шум, шум Бернулли, враждебный шум (Adversarial Noise) и т. д. Исследования добавления шума в векторные представления слов, приведенные в [1–3], в целом показали неплохие результаты, однако использование таких методов аугментации подразумевает использование эмбедингов не существующих слов, что потенциально может привести к рассогласованию с метками классов сообщений, нарушению семантической целостности, изменению лексики и стиля.

Аугментацию текстовых данных на уровне слов и реплик можно условно подразделить на методы, способные к сохранению стиля и словарного состава текста, и методы неспособные к их сохранению. Методы аугментации данных на уровне текста, неспособные к сохранению стиля и словарного состава, могут включать в себя простые методы редактирования текста. Например, набор простых техник аугментации пред-

ставлен в алгоритме EDA [4], который состоит из четырех операций: замена синонимами; случайные вставка, перестановка и удаление. При аугментации текстовых данных важно сохранение смысла текста, в связи с этим часто используются: замена слов на синонимы; различные словари (например, WordNet [1, 4]) или предобученные языковые модели (BERT [5, 6], GPT2 [6], Word2Vec [2], Glove [1]) и другие.

Основываясь на гипотезе о том, что предложения являются естественными, даже когда слова в предложениях заменяются другими словами с парадигматическими отношениями, в работе [7] предложен подход контекстного дополнения. Еще одним подходом может быть использование обратного перевода вместо создания парафраз.

В [8] рассмотрена аугментация данных посредством перевода сообщений с английского на французский и с французского на английский языки для датасета Stanford Question Answering Dataset. В работе [9] использовано несколько этапов машинного перевода с помощью Google Translate<sup>1</sup>. При этом максимальная длина цикла перевода задавалась до трех языков, например, английский → немецкий → датский → английский. Исследование глубоких нейронных сетей на предмет эффективности обратного перевода привело к положительным результатам в работе [10]. Нетривиальная техника аугментации представлена в [11], которая выполнена путем перемешивания текста с использованием нейронной сети. Однако ни один из вышеперечисленных методов не сохраняет исходный стиль сообщения и лексику.

При изменении текстовых данных преобразования могут вносить искажения в текст, делая его грамматически или семантически неверным, или стилистически отличным от исходного текста. По этой причине требуются приемы, которые могут аугментировать текст, при этом обеспечив сохранение стиля, лексики и синтаксической целостности.

<sup>1</sup> Google Translate [Электронный ресурс]. Режим доступа: <https://cloud.google.com/translate> (дата обращения: 14.09.2022).

Существуют также методы аугментации, способные сохранять стиль и словарный состав текста, например, генератор парафраз, основанный на преобразовании синтаксических деревьев [12]. В этом методе текст модифицируется путем преобразования синтаксического дерева на основе общеупотребимых синтаксических грамматик. Аугментация текста с помощью синтаксических деревьев с генерацией новых данных на основе синтаксических шаблонов рассмотрена в [13, 14]. Данные методы действительно позволяют сохранить словарный состав, обеспечить сохранение семантической целостности, но не позволяют сохранить синтаксические особенности речи, в связи с чем стиль речи может быть значительно искажен.

### Описание метода аугментации

В настоящей работе предложен новый метод аугментации, сохраняющий уникальные стиль речи и лексику (рисунок). Под стилем речи подразумевается совокупность отличительных языковых признаков речи персоны, например, синтаксические признаки (сохранение синтаксических структур, наиболее часто используемых персоной) и морфологические признаки (сохранение частотности использования определенных частей речи). Данный метод основан на схеме аугментации данных, предложенной в работе [12]. Важно, что этап преобразования и генерации перефразированных данных в предлагаемом методе — единый процесс. В качестве синтаксических шаблонов использованы синтаксические структуры, которые характерны для речи рассматриваемой персоны, что позволяет сохранить синтаксические особенности речи этой персоны и, таким образом, обеспечить высокую вероятность сохранения стиля речи. Отметим, что метод адаптирован для работы с данными на русском языке.

Пусть дано  $n$  персон, и каждая персона имеет  $m$  (для каждой персоны свое число) реплик  $A_0 \dots A_m$ . Для

каждой реплики извлекается синтаксическое дерево  $S_i$  с помощью парсера Stanford Core NLP для английского и русского языков [15]. Получим синтаксическое дерево предложения «This is a test»:

$$\begin{aligned}
 &(\text{ROOT} \\
 &(\text{S} \\
 &(\text{NP} (\text{DT} \text{This})) \\
 &(\text{VP} (\text{VBZ} \text{is}) \\
 &(\text{NP} (\text{DT} \text{a}) (\text{NN} \text{test})) \\
 &)))
 \end{aligned} \tag{1}$$

На основе синтаксического дерева (1) с помощью изъятия всех слов из исходного предложения и сохранения только синтаксических структурных единиц, получим синтаксический шаблон  $T_j$ . Синтаксический шаблон для предложения «This is a test» имеет вид:

$$(\text{ROOT}(\text{S}(\text{NP}(\text{DT}))(\text{VP}(\text{VBZ})(\text{NP}(\text{DT})(\text{NN})))))) \tag{2}$$

где ROOT — корень предложения; DT — определитель; VP — глагольная фраза; VBZ — глагол настоящего времени 3-го лица единственного числа; NP — именная группа; NN — существительное, нарицательное, единственное число.

Для каждой персоны  $P_i \dots P_n$  и множества принадлежащих ей синтаксических шаблонов (2), составим частотность использования, т. е. создадим пару наборов шаблонов  $T_i$  и  $f$  (количество использований данного шаблона и похожих на него более чем на 96 % шаблонов, при этом похожие шаблоны удаляются из множества). Из оставшегося множества выберем  $r$  наиболее встречающихся шаблонов.

К извлеченным наборам шаблонов, полученных для каждой персоны, применим предобработку для получения формата, принимаемого OpenAttacker Syntactically Controlled Paraphrase Network (SCPNet) — в конец строки добавим тег окончания строк (End-Of-String, EOS).

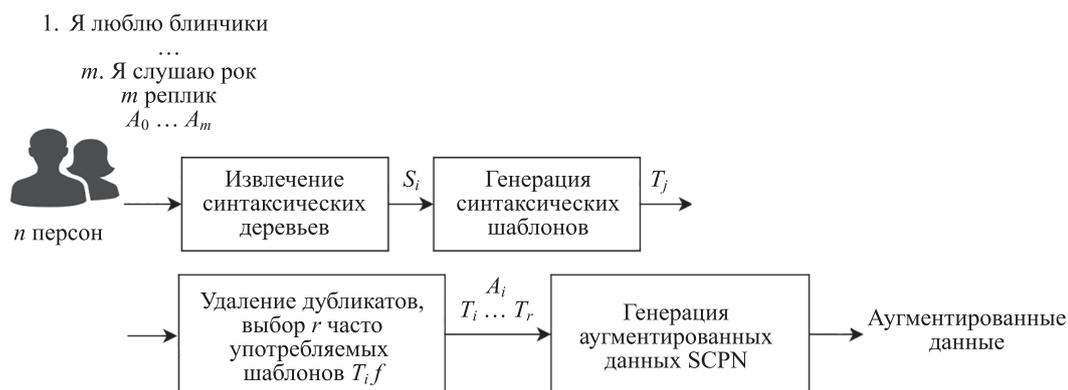


Рисунок. Блок-схема метода аугментации.

$A_0$  — первая реплика персоны;  $S_i$  — синтаксическое дерево;  $A_i$  — одна из реплик персоны;  $T_j$  — синтаксический шаблон для реплики  $A_i$ , где  $j = i$ ;  $f$  — частота использования синтаксического шаблона;  $r$  — настраиваемый параметр количества вариантов аугментации одной реплики;  $T_i$  — один из синтаксических шаблонов после очистки дубликатов

Figure. Augmentation method flowchart.

$A_0$  — person's first replica;  $S_i$  — syntax tree;  $A_i$  — one of the person's replicas;  $T_j$  — syntactic template for the replica  $A_i$ , where  $j = i$ ;  $f$  — frequency of using a syntactic template;  $r$  — configurable parameter for the number of augmentation options for one replica;  $T_i$  — one of the syntax patterns after clearing duplicates

Вместе с исходной репликой человека теги отправляются в SCPN — модель кодера-декодера для синтаксически контролируемой генерации парафраз из фреймворка OpenAttacker для генерации аугментированных данных. Таким образом, каждая реплика персоны может быть преобразована  $r$  различными способами. В настоящей работе выбрано  $r = 5$  и проведены эксперименты с учетом данного значения.

Поскольку для аугментации данных использованы только характерные для человека синтаксические конструкции, то сохранены синтаксические особенности речи человека. Лексика и стиль частично сохранены, в связи с тем, что аугментация основана на преобразовании синтаксического дерева, а при данном преобразовании в первую очередь используется лексика исходного предложения в случае, если какая-либо из частей речи отсутствует в исходной реплике, SCPN добавляет необходимые части речи (союзы, предлоги, частицы) для сохранения синтаксической согласованности дополненных реплик. Если остальные части речи отсутствуют, они добавляются с помощью Long short-term memory, генерирующей слова. Например, фраза «Oh, unbelievable, we had the best time» с синтаксическим шаблоном

(ROOT(S(INTJ(UH))(NP(PRPP))VP(VBD)×  
×(NP(NP(DT)(JJS)(NN)(SBAR(IN)(PRP)×  
(VP(VBD)))))(ADJP(JJ))))

преобразуется во фразу «Oh, it was the best time that we had, unbelievable». Для работы с русскоязычными данными выполнена дополнительная подготовка. Для SCPN она заключалась в дообучении SCPN-модели на русскоязычных данных. Для этого был сформирован h5 файл, в котором содержались примеры синтаксических деревьев, для каждой из синтаксических структур были сформированы входные данные из одной фразы и выходные данные из похожей фразы с ее синтаксической

структурой. Данные использовались из датасета Тайга корпуса Наташа. Помимо этого, для датасета Russian Tweet Corpus (RuTweetCorp) проведена дополнительная очистка данных — из сообщений удалены обращения и хэштеги.

### Эксперименты и результаты

Эксперименты проведены на следующих наборах данных:

- Multimodal EmotionLines Dataset (MELD) — набор данных, содержащий более 1400 диалогов и 13 000 высказываний из сериала «Друзья». В диалогах участвовали несколько дикторов. Каждое высказывание в диалоге было отмечено любой из семи эмоций: Гнев, Отвращение, Печаль, Радость, Нейтральность, Удивление и Страх;
- MELDRU — набор данных MELD, переведенный на русский язык с помощью модели машинного перевода MarianMT;
- RuTweetCorp — русскоязычный корпус коротких текстов на основе русскоязычных постов из микроблога Twitter, состоящий из 17 639 674 записей. Корпус содержит более 139 000 различных дикторов. Данные размечены на два класса по эмотивной валентности: позитивные и негативные.

Пример аугментированных данных с помощью синтаксического перефразирования представлен в табл. 1.

В связи с тем, что в рамках работы рассмотрены датасеты, содержащие неформальную разговорную речь (реплики диалогов и твиты), в данных присутствовало небольшое количество высказываний, обогащенных сложными синтаксическими конструкциями.

Для аугментированных данных оценивалась сохранность языковых характеристик человека с помощью Mean Opinion Score (MOS) оценки. Данная оценка часто используется для предоставления численного значения о качестве аугментированной аудио информации. Для оценки было опрошено 10 носителей рус-

Таблица 1. Пример аугментированных данных с помощью синтаксического перефразирования

Table 1. Examples of data augmented with syntactic paraphrasing

Исходный текст	Датасет	Результат
Oh hey, don't thank me, thank yourself. You're the one who faced her fears and ultimately overcame them.	MELD	['You're the one who faced her fears and eventually stopped them. Thank yourself.', 'Don't thank me, the one who faced her fears and surpassed them is you.', 'Oh, thank yourself, you 're the one who faced her fears , and you missed them.', 'Of course, you 're the one who faced her fears and eventually stopped them', «Oh, you don't thank me , you 're the one who faced her fears and eventually stopped them .»]
Так ты говоришь мне, что между тобой и Чендлером ничего нет.	MELDRU	['Ты говоришь мне, что ничего не происходит между тобой и Чендлером.', 'Так что скажи мне, что это не происходит между тобой и Чендлером.', 'Так ты говоришь мне, что это не происходит между тобой и Чендлером.', 'Так ты говоришь мне, что ничего не происходит между тобой и Чендлером.', 'Так что скажи мне, что с тобой и Чендлером ничего нет.']
Самое тяжкое, что долго не живут животные. Привыкаешь.	RuTweetCorp	['Самое трудное — животные живут недолго, а ты привыкаешь', 'Самое трудное в том, что животные не живут долго. Привыкаешься', 'Трудная вещь, когда долго не живут питомцы. Привыкаешься.', 'Тяжкое, животные не живут долго']

ского языка по трем анкетам на основе наборов данных MELD, MELDRU и RuTweetCorp. В связи с тем, что все персоны, принявшие участие в опросе, являются носителями русского языка, выполнена оценка только русскоязычных наборов данных, так как результаты оценки англоязычного набора данных не были бы объективными из-за отсутствия соответствующего культурного отпечатка и менталитета. В анкетах опросов содержались примеры фраз, характерные для персон. Данные примеры представляли из себя список из пяти фраз, соотнесенных с их персоной. Для примеров из используемых датасетов выбирались наиболее отражающие стиль речи персоны фразы (фразы, в которых присутствовала наиболее часто употребляемая персоной лексика или синтаксическая структура предложения). Далее респонденты выбирали из реплик, аугментированных с помощью различных методик (eda, mt5, предложенный метод), наиболее подходящую реплику для указанной персоны. Для обеспечения чистоты эксперимента каждым методом аугментации было сгенерировано равное количество аугментированных предложений. Выбранная реплика оценивалась значениями от 1 до 5 по шкале осмысленности и логичности, где 5 — «фраза полностью логична», «фраза полностью осмыслена». Оценки представлены в табл. 2.

В результате отметим, что респонденты чаще выбирали реплики, сгенерированные с помощью разработанного авторами данной работы метода, как более подходящие персоне. Заметим, что предложенный метод генерирует реплики с достаточно высоким уровнем логичности и осмысленности. Так как предлагаемые респондентам реплики в анкетах никак не были помечены алгоритмами их сгенерировавшими, то они выбирали подходящие персоне реплики без знания об алгоритмах. Иначе говоря, метод сохраняет отличительные характеристики речи лучше, чем аналоги, с которыми проведено сравнение.

Для оценки влияния аугментированных данных на точность моделей, данные были использованы для распознавания эмоций и анализа эмотивной валентности текста. В связи с тем, что анализ эмотивной валентности в данной работе используется лишь для оценки влияния аугментированных данных на метрики моделей, была рассмотрена эмотивная валентность самого высказывания в диалоговом контексте, при этом особенности характера персоны не учитывались.

Распознавание эмоций в полилогах (Emotion recognition in conversation, ERC) с использованием современных архитектур нейронных сетей достигло заметных успехов. Распознавание эмоций может осуществляться для разных модальностей, текстовая модальность при одномодальных исследованиях показывает лучшие результаты [16, 17]. В последних работах в этой области использованы различные виды трансформеров для кодирования текстовых контекстов [18–20] или, например, графовых сетей [21]. Отсутствие доступных и подходящих наборов данных является основным препятствием для решения этих вопросов для русского языка. Однако существует множество таких наборов данных для английского языка. Исходя из этого, в качестве одного из подходов к решению этой проблемы часто предлагается использовать автоматический перевод англоязычных наборов данных на русский язык. В данном случае использование такого подхода дает результаты достаточного качества. Для исследования был выбран EmoBERTa для решения проблемы ERC. Наряду с хорошей точностью (топ-2 для набора данных MELD), он достаточно прост и удобен в использовании:

- модель RoBERTa предварительно обучается как часть задачи MLM (маскированного языкового моделирования) на целевом языке;
- модель дообучается для задачи классификации текста.

Данные для обучения классификации представляют собой конкатенированные диалоговые контексты. Каждый диалоговый контекст содержит целевую реплику и четыре другие реплики, задействованные в том же диалоге, что и целевая. Включение дополнительных реплик позволяет при оценке эмотивной валентности учитывать контекст целевой реплики за счет предыдущих фраз дикторов и реплик собеседников. Средняя длина одной реплики составляет 8 слов (токенов) для датасета MELD, 6 — для MELDRU и 10 — для RuTweetCorp. Для аугментированных данных диалоговый контекст собран следующим образом: для каждой реплики из контекста, в том числе и для целевой, случайным образом с заданной вероятностью берется один из пяти аугментированных вариантов и присоединяется к общему контексту.

Для анализа эмотивной валентности текста сообщений на русском языке была выбрана предобученная Multilingual BERT модель. Сравнение влияния ауг-

Таблица 2. MOS оценка на данных датасетов MELDRU и RuTweetCorp  
Table 2. MOS assessment based on MELDRU and RuTweetCorp datasets

Параметры	Датасет					
	MELDRU			RuTweetCorp		
	EDA (Easy Data Augmentation)	MT5	Предложенный метод	EDA	MT5	Предложенный метод
Средняя оценка осмысленности	4,2	4,4	4,4	4,3	3,8	4,2
Средняя оценка логичности	4,3	4,4	<b>4,5</b>	4,6	3,2	3,8
Процент анкет с преобладанием данного метода	—	12	<b>88</b>	—	30	<b>70</b>
Процент доли выбора метода для всех вопросов	17	18	<b>65</b>	18	32	<b>50</b>

Таблица 3. Результаты точности моделей с аугментированными данными

Table 3. Performance of the tested augmentation approaches

Модель	Датасет	Доля аугментированных данных	Предложенный метод		EDA	
			Accuracy	Weighted F1	Accuracy	Weighted F1
RoBERTa-base	MELDRU	0,0	56,02	50,01	56,02	50,01
RoBERTa-base	MELDRU	0,5	55,86	<b>53,17</b>	53,25	51,14
RoBERTa-base	MELDRU	1,0	52,43	50,51	51,76	48,34
Distill RoBERTa-base	MELD	0,0	60,30	58,19	60,30	58,19
Distill RoBERTa-base	MELD	0,5	<b>61,87</b>	<b>60,40</b>	53,39	48,01
Distill RoBERTa-base	MELD	1,0	53,26	48,77	51,92	47,20
MBERT-base	RuTweetCorp	0,0	68,30	67,27	68,30	67,27
MBERT-base	RuTweetCorp	0,5	<b>70,60</b>	<b>69,40</b>	62,76	57,74
MBERT-base	RuTweetCorp	1,0	67,45	66,29	64,24	61,36

ментированных вариантов предложенным методом проведено с классическим методом аугментации EDA, в котором аугментация текста производится за счет замены синонимами, с помощью случайных вставки, перестановки и удаления.

Эксперимент по распознаванию эмоций и анализу эмотивной валентности выполнен следующим образом. Было опробовано три варианта использования аугментированных данных. Вариант  $\text{aug-prob}=0,0$  соответствует обучению без аугментаций. В варианте  $\text{aug-prob}=1,0$  каждое высказывание заменяется аугментацией. Вариант  $\text{augprob}=0,5$  — точка баланса, при котором половина высказываний заменяется аугментацией, другая половина остается неизменной. Результаты экспериментов приведены в табл. 3.

В результате оценки отметим, что предложенный метод не только позволил достичь повышения точности для разных моделей до 2 %, но и показал результат по влиянию на точность лучше, чем метод аугментации EDA.

## Заключение

Предложен метод аугментации текстовых данных, который сохраняет стиль речи и словарный запас. В работе использованы наборы данных MELD, MELDRU и RuTweetCorp для экспериментов по оценке эмоционального состояния пользователя. Замечено, что аугментация данных с помощью представленного метода приводит к повышению точности при доле аугментированных данных к исходным менее 1,0, потому что в этом случае есть шанс, что исходное сообщение останется неизменным. Исходное сообщение, каким бы хорошим ни было дополнение, содержит данные, которые максимально подходят к домену, из которого оно было взято. С другой стороны, полезно увеличить разнообразие данных с помощью аугментации, повышающей приобретенную способность сохранять стиль речи за счет лучшего обобщения. Наблюдалось увеличение точности для разных моделей до 2 %.

## Литература

1. Giridhara P.K., Mishra C., Venkataramana R.K., Bukhari S.S., Dengel A.R. A study of various text augmentation techniques for relation classification in free text // Proc. of the 8<sup>th</sup> International Conference on Pattern Recognition Applications and Methods. 2019. P. 360–367 <https://doi.org/10.5220/0007311003600367>
2. Papadaki M. Data Augmentation Techniques for Legal Text Analytics: A thesis submitted to Athens University of Economics and Business in fulfillment of the requirements for the degree of Master in Data Science. 2017. 33 p.
3. Zhang Z., Zweigenbaum P. GNEG: Graph-based negative sampling for word2vec // Proc. of the 56<sup>th</sup> Annual Meeting of the Association for Computational Linguistics. V. 2. 2018. P. 566–571. <https://doi.org/10.18653/v1/P18-2090>
4. Wei J., Zou K. EDA: Easy data augmentation techniques for boosting performance on text classification tasks // Proc. of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9<sup>th</sup> International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). 2018. P. 6382–6388. <https://doi.org/10.18653/v1/D19-1670>
5. Wu X., Xia Y., Zhu J., Wu L., Xie S., Fan Y., Qin T. mixSeq: A simple data augmentation method for neural machine translation // Proc. of the 18<sup>th</sup> International Conference on Spoken Language Translation (IWSLT 2021). 2021. P. 192–197. <https://doi.org/10.18653/v1/2021.iwslt-1.23>

## References

1. Giridhara P.K., Mishra C., Venkataramana R.K., Bukhari S.S., Dengel A.R. A study of various text augmentation techniques for relation classification in free text. *Proc. of the 8<sup>th</sup> International Conference on Pattern Recognition Applications and Methods*, 2019, pp. 360–367 <https://doi.org/10.5220/0007311003600367>
2. Papadaki M. *Data Augmentation Techniques for Legal Text Analytics*. A thesis submitted to Athens University of Economics and Business in fulfillment of the requirements for the degree of Master in Data Science, 2017, 33 p.
3. Zhang Z., Zweigenbaum P. GNEG: Graph-based negative sampling for word2vec. *Proc. of the 56<sup>th</sup> Annual Meeting of the Association for Computational Linguistics. V. 2*, 2018, pp. 566–571. <https://doi.org/10.18653/v1/P18-2090>
4. Wei J., Zou K. EDA: Easy data augmentation techniques for boosting performance on text classification tasks. *Proc. of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9<sup>th</sup> International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2018, pp. 6382–6388. <https://doi.org/10.18653/v1/D19-1670>
5. Wu X., Xia Y., Zhu J., Wu L., Xie S., Fan Y., Qin T. mixSeq: A simple data augmentation method for neural machine translation. *Proc. of the 18<sup>th</sup> International Conference on Spoken Language Translation (IWSLT 2021)*, 2021, pp. 192–197. <https://doi.org/10.18653/v1/2021.iwslt-1.23>

6. Kumar V., Choudhary A., Cho E. Data augmentation using pre-trained transformer models // Proc. of the 2<sup>nd</sup> Workshop on Life-long Learning for Spoken Language Systems. 2020. P. 18–26.
7. Kobayashi S. Contextual augmentation: Data augmentation by words with paradigmatic relations // Proc. of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers). 2018. P. 452–457. <https://doi.org/10.18653/v1/N18-2072>
8. Yu A., Dohan D., Luong M., Zhao R., Chen K., Norouzi M., Le Q. QANet: Combining local convolution with global self-attention for reading comprehension // Proc. of the ICLR Conference. 2018.
9. Mehdi R., Meyer M., Goutal S. Text Data Augmentation: Towards better detection of spear-phishing emails // arXiv. 2020. arXiv:2007.02033. <https://doi.org/10.48550/arXiv.2007.02033>
10. Edunov S., Ott M., Auli M., Grangier D. Understanding back-translation at scale // Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing. 2018. P. 489–500. <https://doi.org/10.18653/v1/D18-1045>
11. Guo H., Mao Y., Zhang R. Augmenting data with mixup for sentence classification: An empirical study // arXiv. 2019. arXiv:1905.08941. <https://doi.org/10.48550/arXiv.1905.08941>
12. Coulombe C. Text data augmentation made simple by leveraging NLP cloud APIs // arXiv. 2018. arXiv:1812.04718. <https://doi.org/10.48550/arXiv.1812.04718>
13. Shen T., Lei T., Barzilay R., Jaakkola T. Style transfer from non-parallel text by cross-alignment // Advances in Neural Information Processing Systems. 2017. V. 30.
14. Yang S., Huang X., Lau J.H., Erfani S. Robust task-oriented dialogue generation with contrastive pre-training and adversarial filtering // Findings of the Association for Computational Linguistics (EMNLP 2022). 2022. P. 1220–1234.
15. Kovrigin L., Shilin I., Shipilo A., Putintseva A. Russian tagging and dependency parsing models for stanford CoreNLP natural language toolkit // Communications in Computer and Information Science. 2017. V. 786. P. 101–111. [https://doi.org/10.1007/978-3-319-69548-8\\_8](https://doi.org/10.1007/978-3-319-69548-8_8)
16. Matveev Y., Matveev A., Frolova O., Lyakso E., Ruban N. Automatic speech emotion recognition of younger school age children // Mathematics. 2022. V. 10. N 14. P. 2373. <https://doi.org/10.3390/math10142373>
17. Lyakso E., Frolova O., Matveev A., Matveev Y., Grigorev A., Makhnytkina O., Ruban N. Recognition of the emotional state of children with down syndrome by video, audio and text modalities: human and automatic // Lecture Notes in Computer Science. 2022. V. 13721. P. 438–450. [https://doi.org/10.1007/978-3-031-20980-2\\_38](https://doi.org/10.1007/978-3-031-20980-2_38)
18. Kim T., Vossen P. EmoBERTa: Speaker-Aware Emotion Recognition in Conversation with RoBERTa // arXiv. 2021. arXiv:2108.12009. <https://doi.org/10.48550/arXiv.2108.12009>
19. Song X., Zang L., Zhang R., Hu S., Huang L. Emotionflow: Capture the dialogue level emotion transitions // Proc. of the ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2022. P. 8542–8546. <https://doi.org/10.1109/ICASSP43922.2022.9746464>
20. Shen W., Chen J., Quan X., Xie Z. DialogXL: All-in-One XLNet for multi-party conversation emotion recognition // Proceedings of the AAAI Conference on Artificial Intelligence. 2021. V. 35. N 15. P. 13789–13797 <https://doi.org/10.1609/aaai.v35i15.17625>
21. Shen W., Wu S., Yang Y., Quan X. Directed acyclic graph network for conversational emotion recognition // Proc. of the 59<sup>th</sup> Annual Meeting of the Association for Computational Linguistics and the 11<sup>th</sup> International Joint Conference on Natural Language Processing (Volume 1: Long Papers). 2021. P. 1551–1560.
6. Kumar V., Choudhary A., Cho E. Data augmentation using pre-trained transformer models. *Proc. of the 2<sup>nd</sup> Workshop on Life-long Learning for Spoken Language Systems*, 2020, pp. 18–26.
7. Kobayashi S. Contextual augmentation: Data augmentation by words with paradigmatic relations. *Proc. of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers)*, 2018, pp. 452–457. <https://doi.org/10.18653/v1/N18-2072>
8. Yu A., Dohan D., Luong M., Zhao R., Chen K., Norouzi M., Le Q. QANet: Combining local convolution with global self-attention for reading comprehension. *Proc. of the ICLR Conference*, 2018.
9. Mehdi R., Meyer M., Goutal S. Text Data Augmentation: Towards better detection of spear-phishing emails. *arXiv*, 2020, arXiv:2007.02033. <https://doi.org/10.48550/arXiv.2007.02033>
10. Edunov S., Ott M., Auli M., Grangier D. Understanding back-translation at scale. *Proc. of the 2018 Conference on Empirical Methods in Natural Language Processing*, 2018, pp. 489–500. <https://doi.org/10.18653/v1/D18-1045>
11. Guo H., Mao Y., Zhang R. Augmenting data with mixup for sentence classification: An empirical study. *arXiv*, 2019, arXiv:1905.08941. <https://doi.org/10.48550/arXiv.1905.08941>
12. Coulombe C. Text data augmentation made simple by leveraging NLP cloud APIs. *arXiv*, 2018, arXiv:1812.04718. <https://doi.org/10.48550/arXiv.1812.04718>
13. Shen T., Lei T., Barzilay R., Jaakkola T. Style transfer from non-parallel text by cross-alignment. *Advances in Neural Information Processing Systems*, 2017, vol. 30.
14. Yang S., Huang X., Lau J.H., Erfani S. Robust task-oriented dialogue generation with contrastive pre-training and adversarial filtering. *Findings of the Association for Computational Linguistics (EMNLP 2022)*, 2022, pp. 1220–1234.
15. Kovrigin L., Shilin I., Shipilo A., Putintseva A. Russian tagging and dependency parsing models for stanford CoreNLP natural language toolkit. *Communications in Computer and Information Science*, 2017, vol. 786, pp. 101–111. [https://doi.org/10.1007/978-3-319-69548-8\\_8](https://doi.org/10.1007/978-3-319-69548-8_8)
16. Matveev Y., Matveev A., Frolova O., Lyakso E., Ruban N. Automatic speech emotion recognition of younger school age children. *Mathematics*, 2022, vol. 10, no. 14, pp. 2373. <https://doi.org/10.3390/math10142373>
17. Lyakso E., Frolova O., Matveev A., Matveev Y., Grigorev A., Makhnytkina O., Ruban N. Recognition of the emotional state of children with down syndrome by video, audio and text modalities: human and automatic. *Lecture Notes in Computer Science*, 2022, vol. 13721, pp. 438–450. [https://doi.org/10.1007/978-3-031-20980-2\\_38](https://doi.org/10.1007/978-3-031-20980-2_38)
18. Kim T., Vossen P. EmoBERTa: Speaker-Aware Emotion Recognition in Conversation with RoBERTa. *arXiv*, 2021, arXiv:2108.12009. <https://doi.org/10.48550/arXiv.2108.12009>
19. Song X., Zang L., Zhang R., Hu S., Huang L. Emotionflow: Capture the dialogue level emotion transitions. *Proc. of the ICASSP 2022–2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022, pp. 8542–8546. <https://doi.org/10.1109/ICASSP43922.2022.9746464>
20. Shen W., Chen J., Quan X., Xie Z. DialogXL: All-in-One XLNet for multi-party conversation emotion recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, vol. 35, no. 15, pp. 13789–13797 <https://doi.org/10.1609/aaai.v35i15.17625>
21. Shen W., Wu S., Yang Y., Quan X. Directed acyclic graph network for conversational emotion recognition. *Proc. of the 59<sup>th</sup> Annual Meeting of the Association for Computational Linguistics and the 11<sup>th</sup> International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, 2021, pp. 1551–1560.

#### Авторы

**Матвеева Анастасия Андреевна** — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57204215042](https://orcid.org/0000-0002-2935-991X), <https://orcid.org/0000-0002-2935-991X>, [anastasiamatveevaitmo@gmail.com](mailto:anastasiamatveevaitmo@gmail.com)

**Махныткина Олеся Владимировна** — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57208002090](https://orcid.org/0000-0002-8992-9654), <https://orcid.org/0000-0002-8992-9654>, [makhnytkina@itmo.ru](mailto:makhnytkina@itmo.ru)

#### Authors

**Anastasia A. Matveeva** — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57204215042](https://orcid.org/0000-0002-2935-991X), <https://orcid.org/0000-0002-2935-991X>, [anastasiamatveevaitmo@gmail.com](mailto:anastasiamatveevaitmo@gmail.com)

**Olesia V. Makhnytkina** — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57208002090](https://orcid.org/0000-0002-8992-9654), <https://orcid.org/0000-0002-8992-9654>, [makhnytkina@itmo.ru](mailto:makhnytkina@itmo.ru)

doi: 10.17586/2226-1494-2023-23-4-750-756

## Verification of event-driven software systems using the specification language of cooperating automata objects

Irina V. Afanasieva<sup>1</sup>, Fedor A. Novikov<sup>2</sup>, Ludmila N. Fedorchenko<sup>3</sup>✉

<sup>1</sup> Special Astrophysical Observatory of the Russian Academy of Sciences (SAO RAS), Nizhny Arkhyz, 369167, Russian Federation

<sup>2</sup> Peter the Great St. Petersburg Polytechnic University (SPbPU), Saint Petersburg, 195251, Russian Federation

<sup>2</sup> Alferov Academic University, Saint Petersburg, 194021, Russian Federation

<sup>3</sup> St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

<sup>1</sup> riv615@gmail.com, <https://orcid.org/0000-0003-4225-4124>

<sup>2</sup> fedornovikov51@gmail.com, <https://orcid.org/0000-0003-4450-0173>

<sup>3</sup> Inf@iiias.spb.su✉, <https://orcid.org/0000-0002-4008-9316>

### Abstract

The CIAO (Cooperative Interaction Automata Objects) specification language is intended to describe the behavior of distributed and parallel event-driven systems. This class of systems includes various software and hardware systems for control, monitoring, data collection, and processing. The ability to verify compliance with requirements is desirable competitive advantage for such systems. The CIAO language extends the concept of state machines of the UML (Unified Modeling Language) with the possibility of cooperative interaction of several automata through strictly defined interfaces. The cooperative interaction of automata objects is defined by a link scheme that defines how the provided and required interfaces of different automata objects are connected. Thus, the behavior of the system as a whole could be described as a set of execution protocols, each of which is a sequence of interface calls, possibly with guard conditions. We represent a set of protocols using a semantic graph in which all possible paths from the initial nodes to the final nodes define sequences of interface method calls. Because the interfaces are strictly defined in advance by the connection scheme, it is possible to construct a semantic graph automatically according to a given system of interacting automaton objects. To verify the system behavior, one only has to check if each path in the semantic graph does satisfy the requirements. System requirements are formally described using conditional regular expressions that define patterns of acceptable and forbidden behavior. This article proposes methods and algorithms that allow you to check the compliance of programs in the CIAO language with the requirements automatically and, thereby, check the semantics of the developed program. The proposed method narrows the specification formalism to the class of regular languages and the programming language to a language with a simple and predefined structure. In many practical cases, this is sufficient for effective verification.

### Keywords

verification, event-driven programs, critical system development, regular expressions, elevator control system

### Acknowledgments

The work of Irina Afanasieva was carried out within the framework of the SAO RAS State Assignment approved by the Ministry of Science and Higher Education of the Russian Federation no. 121092300060-6.

**For citation:** Afanasieva I.V., Novikov F.A., Fedorchenko L.N. Verification of event-driven software systems using the specification language of cooperating automata objects. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 750–756. doi: 10.17586/2226-1494-2023-23-4-750-756

УДК 004.415.52, 004.434

## Верификация событийно-управляемых программных систем с использованием языка спецификации взаимодействующих автоматных объектов

Ирина Викторовна Афанасьева<sup>1</sup>, Федор Александрович Новиков<sup>2</sup>,  
Людмила Николаевна Федорченко<sup>3</sup>✉

<sup>1</sup> Специальная астрофизическая обсерватория Российской академии наук, Нижний Архыз, 369167, Российская Федерация

<sup>2</sup> Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, 195251, Российская Федерация

<sup>2</sup> Санкт-Петербургский национальный исследовательский Академический университет имени Ж.И. Алфёрова Российской академии наук, Санкт-Петербург, 194021, Российская Федерация

<sup>3</sup> Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация

<sup>1</sup> riv615@gmail.com, <https://orcid.org/0000-0003-4225-4124>

<sup>2</sup> fedornovikov51@gmail.com, <https://orcid.org/0000-0003-4450-0173>

<sup>3</sup> lnf@ias.spb.su✉, <https://orcid.org/0000-0002-4008-9316>

### Аннотация

**Введение.** Язык спецификации Cooperative Interaction Automata Objects (CIAO) предназначен для описания поведения распределенных и параллельных систем, управляемых событиями. К этому классу систем относятся различные программно-аппаратные комплексы управления, контроля, сбора и обработки данных. Возможность автоматической проверки соответствия требованиям является желательным конкурентным преимуществом событийно-управляемых программных систем. Язык CIAO расширяет концепцию конечных автоматов (Unified Modeling Language, UML) возможностью кооперативного взаимодействия нескольких автоматов через строго определенные интерфейсы. Кооперативное взаимодействие автоматных объектов определяется схемой связи, которая связывает предоставленные и требуемые интерфейсы различных автоматных объектов. Таким образом, поведение системы в целом можно описать как набор протоколов выполнения, каждый из которых представляет собой последовательность вызовов интерфейса, возможно со сторожевыми условиями. **Метод.** Представлен набор протоколов с помощью семантического графа, в котором все возможные пути от начальных к конечным узлам определены последовательностью вызовов методов интерфейса. Благодаря тому, что интерфейсы заранее строго определены схемой связи, возможно автоматическое построение семантического графа по заданной системе взаимодействующих автоматных объектов. Для проверки поведения системы достаточно убедиться, что каждый путь в семантическом графе удовлетворяет требованиям. Системные требования формально описаны с помощью условных регулярных выражений, определяющих шаблоны допустимого и запрещенного поведения. **Основные результаты.** Предложены методы и алгоритмы, позволяющие автоматически проверить соответствие программ на языке CIAO заданным требованиям. **Обсуждение.** Разработанный метод сужает формализм для задания спецификаций до класса регулярных языков, а язык программирования — до языка с простой и предопределенной структурой. Во многих практических случаях этого достаточно для эффективной верификации.

### Ключевые слова

верификация, событийно-ориентированные программы, разработка критических систем, регулярные выражения, система управления лифтом

### Благодарности

Работа Ирины Афанасьевой выполнена в рамках государственного задания Специальной астрофизической обсерватории РАН № 121092300060-6, утвержденного Министерством науки и высшего образования Российской Федерации.

**Ссылка для цитирования:** Афанасьева И.В., Новиков Ф.А., Федорченко Л.Н. Верификация событийно-управляемых программных систем с использованием языка спецификации взаимодействующих автоматных объектов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 750–756 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-4-750-756

### Introduction

Event-driven programs, otherwise called discrete reactive systems [1, 2], are often found in the tasks of control, monitoring, data collection, and processing. An event-driven system responds to emerging events (stimuli) by performing certain actions (reactions). Such systems are often classified as critical systems [3, 4] for which the formulation of requirements and verification of compliance with the requirements are nontrivial tasks. Ordinary verbal formulations and selective testing are not enough for

critical systems, and the use of formal verification methods [2, 5] is necessary.

At the same time, for event-driven systems, it is not enough to specify a logical precondition that must be met before starting work and a logical postcondition that must be met because of the implementation of a certain sequence of events/actions, since the same set of actions can be performed both in permissible sequence and in an undesirable forbidden sequence.

Thus, the formal requirements for the systems of the class under consideration must be set in the form

of a description of both admissible and inadmissible sequences of events/actions. Various methods are known for the formal description of a set of sequences which vary depending on how diverse the elements of the sequences are and how complex the sequences are organized.

In this case, elementary events/actions can be interpreted as symbols of a finite alphabet, sequences of actions can be interpreted as words in this alphabet, and in this case, formal requirements put some language in a given alphabet [6].

The purpose of this article is to present methods and algorithms that allow for a certain class of reactive systems, namely for systems described in the Cooperative Interaction Automata Objects (CIAO) language [7, 8], to build a formal description of a set of sequences of actions (a set of possible execution protocols) in a form very close to conditional regular expressions [9] automatically.

### The application of the CIAO specification language for building event-driven systems

The practical application of the CIAO language in the field of creating control and data processing systems has shown good results, in particular, a high degree of reliability of software created using this language [10]. The CIAO language is based on the use of state transition graphs to describe the behavior of reactive systems, and Unified Modeling Language state machine diagrams [11] are used as transition graphs extended with additional constructions and conventions to increase the expressive power of the language. The most significant innovation of the CIAO language is, on the one hand, the multiplicity of interacting automata objects and, on the other hand, the strict typing of interaction interfaces.

In the article [8], we considered the use of the CIAO language for building control systems using the example of the elevator control problem described by D. Knuth [12].

In this problem, the actions that the elevator can perform, the conditions that the control algorithm can check, and the requirements that the control algorithm must satisfy are specified. Because of applying the technique described in the article [8], the specification of the elevator control algorithm in the CIAO language was obtained, as shown in Fig. 1.

Here  $x$  is the number of the starting floor,  $y$  is the number of the target floor,  $t_1$  is the maximum waiting time for the passenger to start service,  $t_2$  is the maximum waiting time for the elevator to enter the passenger

This specification contains states, events, actions, and guard conditions. Abbreviated identifiers according to the approach proposed in [13] accompanied by short descriptions are given in Table 1.

For a passenger, all events come from within, from free will, and only the *ez5* event comes from the elevator. The elevator, on the other hand, has no free will, and all the events of the elevator are the actions of the passenger.

In article [8], the requirements for the control system are formulated as follows.

1. All requests to move to the floors inside the elevator must be serviced.
2. All elevator call requests from floors must be serviced.

In the article [8], it is also shown how the specification in Fig. 1 guaranties the fulfillment of requirements 1 and 2. Here we go further.

When developing critical systems, it is necessary to be able to verify compliance with requirements at all stages of development, including the ability to change requirements if necessary or to incorporate new requirements into an existing system. Consider, for example, the following additional requirement.

3. If there is a passenger in the elevator car, the light must be turned on.

In the following sections, we show how this new additional requirement can be verified.

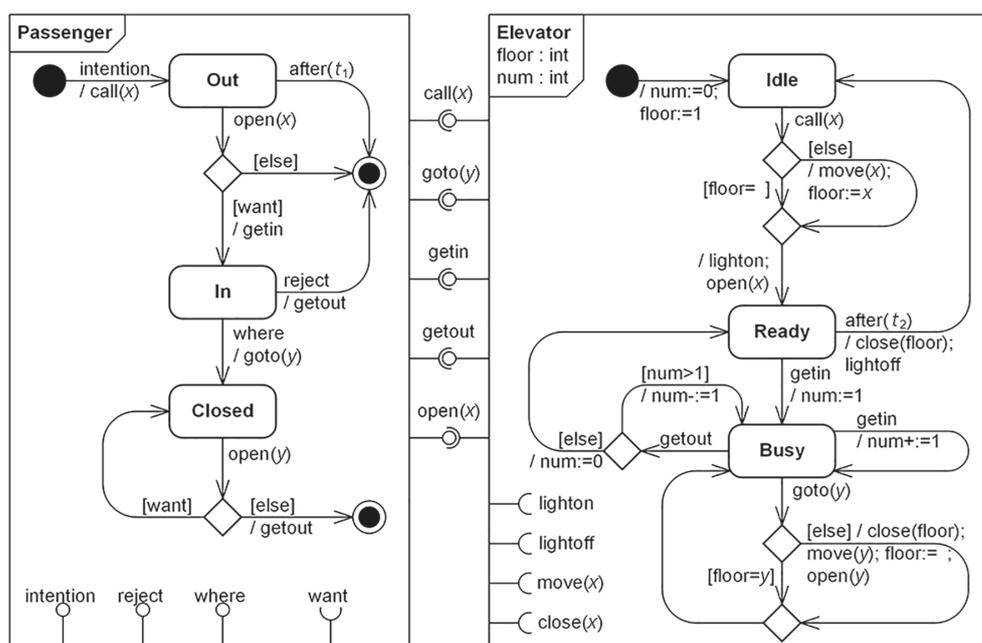


Fig. 1. Specification of elevator control system in CIAO language

Table 1. States, Events, Actions and Guard Conditions IDs and descriptions

ID	Name	Kind	Description
<i>e1</i>	intention	Event	The passenger had the idea to take the elevator from floor <i>x</i> to floor <i>y</i>
<i>e2</i>	reject	Event	The passenger changed his mind about taking the elevator
<i>e3</i>	where	Event	The passenger has decided which floor he should get to
<i>ez1</i>	call ( <i>x</i> )	Event/Action	The passenger presses the elevator call button on floor <i>x</i>
<i>ez2</i>	goto ( <i>y</i> )	Event/Action	The passenger in the elevator cabin presses the departure button to floor <i>y</i>
<i>ez3</i>	getin	Event/Action	The passenger enters the elevator
<i>ez4</i>	getout	Event/Action	The passenger exits the elevator
<i>ez5</i>	open ( <i>x</i> )	Event/Action	Open elevator and shaft doors on floor <i>x</i>
<i>z1</i>	move	Action	Move the elevator to floor <i>x</i>
<i>z2</i>	lighton	Action	Switch the lights on
<i>z3</i>	lightoff	Action	Switch the lights off
<i>z4</i>	close	Action	Close the doors of the elevator car and the doors of the shaft on floor <i>x</i>
<i>c1</i>	want	Guard	The passenger wants to use the elevator
<i>c2</i>	floor = <i>x</i>	Guard	The elevator is on floor <i>x</i>
<i>s1</i>	Out	State	The passenger is outside the elevator car
<i>s2</i>	In	State	The passenger is in the elevator car, the doors are open, and the elevator is stopped
<i>s3</i>	Closed	State	The passenger is in the elevator car, the doors are closed, and the elevator is moving
<i>s4</i>	Idle	State	The elevator is idle, the doors are closed, and the lights are off
<i>s5</i>	Ready	State	The elevator is ready for service, the doors are open, and the lights are on
<i>s6</i>	Busy	State	The elevator is serving a passenger, the doors are closed, the lights are on, and the elevator is moving

### The description of the semantics for reactive systems

The basis of the proposed verification methods is the use of the technique of formal description of languages.

In these notations and conventions, elementary actions are symbols of some alphabet (in this example, the symbols of the alphabet are listed in the first column ID of Table 1). Thus, the program execution protocol is a word in a given alphabet, and the entire set of protocols, that is, the semantics of the program, is a language over this alphabet.

In terms of formal languages, the requirements for the system are expressed as some statements about the structure of the words of the language that define the semantics. It is most convenient to set the structure of words using a regular expression, and then checking compliance with the requirements is reduced to the problem of parsing [14].

Let us explain what has been said with an example. In the notation of Table 1, we introduce the event/action class *zz*, which includes all actions except *z2*, *z3*, *z4*, and *ez5*. Then, requirement 3 is written as the following regular expression (where \* denotes the Kleene iteration, as usual).

$$(zz)^*, z2, ez5, ((zz)^*, z4, (zz)^*, ez5, (zz)^*)^*, z4, z3. \quad (1)$$

In fact, requirement 3 means that the actions of turning on/off the light and entering/exiting the passenger must always be performed strictly in the specified sequence relative to each other and nothing else.

The fulfillment of the requirements described in this way is easily established by the algorithm for checking

that each instance of the program execution protocol conforms to the regular language specified by the regular expression (1). Therefore, we assume that the requirements for the system are specified by a set of regular expressions, possibly with guard conditions, and the system itself is specified by a set of transition graphs in the CIAO language. The arcs of this transition graph are marked by events/actions which are considered as symbols of regular expressions.

### Algorithm for automatic verification of semantics

Verification of compliance with the requirements for a given system in the CIAO language is carried out in three stages.

At the first stage, according to the given transition graphs of interacting automata objects and link scheme, a single graph is built, which is similar to a language source graph [15]. This source graph actually defines the semantics of the program, so we call it the semantic graph. The nodes in this graph correspond to events/actions, and the arcs can be labeled with guard conditions.

The graph turns out to be single since the action in one graph of transitions through the corresponding interface is unified with the event in another graph of transitions. The paths in the constructed graph correspond to the system execution protocols, that is, the words of the language of the semantics being verified. Fig. 2 shows a semantic graph for an elevator control system.

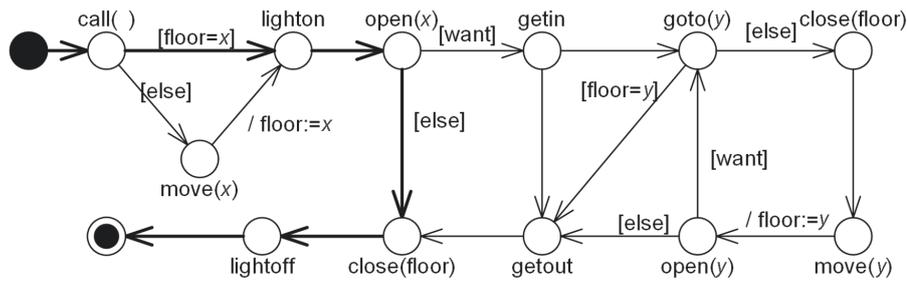


Fig. 2. Semantic graph of the elevator control system

In the second stage, conditional regular expressions from the graph that describe the language of the semantics of the system being verified are constructed. If the language of semantics turns out to be regular (which is often encountered in practice, for example, in communication protocols [16]), then the proposed methods turn out to be universal and allow one to extract a regular expression describing the semantics of a program directly from a given program in the CIAO language automatically.

If the semantic language of the program being checked is not regular, then it is impossible to obtain a single regular expression, but it is possible to obtain a certain set of expressions provided by guard conditions. We can say that the behavior of the program as a whole can be described as a set of several descriptions of the behavior of the program in special “modes”, each mode being characterized by its own guard condition. The division of the program into different modes of operation is not automatic, but is determined by the programmer in the CIAO language when setting guard conditions. In this example, the following conditional regular expression is obtained, matching the bold path in the semantic graph in Fig. 2.

$$ez1, [c2], z2, ez5, [!c1], z4, z3. \quad (2)$$

At the third stage, it is necessary to match the requirements given by the templates with the regular expressions obtained at the second stage. The matching consists of checking that the received regular expressions describing the semantics (paths from the initial node to the final one in the semantic graph) really have the structure prescribed by the requirements templates. For example, it is easy to check that the resulting regular expression (2) is indeed a special case of pattern (1), which means that the introduced requirement 3 is satisfied.

### Discussion and conclusion

Thus, if the terms of reference for a responsible reactive system are described in terms of acceptable and unacceptable sequences of elementary actions using conditional regular expressions, i.e., the desired and undesirable behavior is indicated, then the proposed methods allow, without testing, a mathematically strict check of the compliance of the developed system in the CIAO language to the requirements specified automatically.

Last but not least, this is the positioning of the proposed semantic verification method by constructing

a semantic graph among other verification methods. When comparing verification methods, it is necessary to consider two significant factors that affect the theoretical significance and practical applicability of the methods. First, what is the expressive power of the formalism used for the specification? The more powerful the formalism used; more complex requirements may be specified in the specification. Second, what is the expressive power of the programming language being used? The more powerful the language used, the wider the class of automatically checked programs.

In his outstanding book [17], Dijkstra developed the method of predicate transformers which is in fact a full-scale verification. This method is, in a certain sense, extremely general, since Dijkstra’s Guarded Command Language is Turing complete, and the first-order predicate calculus language is sufficient in all reasonable cases, according to David Hilbert. However, subsequent studies have shown that the method of predicate transformers cannot be completely automated because the verification problem turns out to be algorithmically unsolvable if the power of the specification language is not limited and the power of the programming language is not limited.

The next step was the development of a family of methods for model checking [2], among which, as applied to automaton programs, work [5] stands out. Model checking allows automation by using temporal logic languages for the specification and application of automaton models as programs to be verified. These restrictions make it possible to narrow the verification problem to an algorithmically solvable one. However, in the model checking method, both the class of specifications and the class of automaton models are still very wide, and therefore, the automatic verification methods turn out to be computationally laborious.

In the proposed verification method, the specification formalism is narrowed down to a well-studied class of regular languages, and the programming language is narrowed down to the CIAO language, programs in which have a very simple and predefined structure. Due to this narrowing, it was possible to construct an efficient algorithm for automatic verification. In fact, this algorithm automatically checks whether the program performs some actions in a given order or does not perform some other actions in a forbidden order. No more, but no less. If we carefully structure the required behavior into elementary actions and translate informal requirements into conditional

regular expressions, then in many practical cases the proposed limited means will be sufficient for effective verification. The main intellectual effort in automatic

programming according to this method is required in the formalization of informal specifications, in full accordance with a recent letter from Prof. A.A. Shalyto [18].

## References

1. Harel D., Pnueli D. On the development of reactive systems. *Logics and Models of Concurrent Systems*. Berlin, Heidelberg, Springer, 1985, pp. 477–498. [https://doi.org/10.1007/978-3-642-82453-1\\_17](https://doi.org/10.1007/978-3-642-82453-1_17)
2. Karpov Iu.G. *Model Checking. Verification of Parallel and Distributed Software Systems*. St. Petersburg, BHV-Petersburg Publ., 2010, 560 p. (in Russian)
3. Sommerville I. *Software Engineering*. 10<sup>th</sup> ed. Boston, Pearson, 2020.
4. Hinchey M., Coyle L. Evolving critical systems: a research agenda for computer-based systems. *Proc. of the 17<sup>th</sup> IEEE International Conference and Workshops on Engineering of Computer Based Systems*, 2010, pp. 430–435. <https://doi.org/10.1109/ECBS.2010.56>
5. Velder S.E., Lukin M.A., Shalyto A.A., Iaminov B.R. *Automata-Based Program Verification*. St. Petersburg, Nauka Publ., 2011, 244 p. (in Russian)
6. Fedorchenko L., Baranov S. Equivalent transformations and regularization in context-free grammars. *Cybernetics and Information Technologies*, 2015, vol. 14, no. 4, pp. 29–44. <https://doi.org/10.1515/cait-2014-0003>
7. Novikov F.A., Afanasieva I.V. Cooperative interaction of automata objects. *Information and Control Systems*, 2016, no. 6, pp. 50–64. (in Russian). <https://doi.org/10.15217/issn1684-8853.2016.6.50>
8. Afanasieva I., Novikov F., Fedorchenko L. Methodology for development of event-driven software systems using ciao specification language. *SPIIRAS Proceedings*, 2020, no. 19, no. 3, pp. 481–514. (in Russian). <https://doi.org/10.15622/sp.2020.19.3.1>
9. Aho A.V., Lam M.S., Sethi R., Ullman J.D. *Compilers: Principles, Techniques, and Tools*. 2<sup>nd</sup> ed. Boston, Pearson/Addison-Wesley, 2007, 1009 p.
10. Afanasieva I.V., Novikov F.A. Software architecture for optical detector systems. *Information and Control Systems*, 2016, no. 3, pp. 51–63. (in Russian). <https://doi.org/10.15217/issn1684-8853.2016.3.51>
11. Novikov F.A., Ivanov D.Iu. *UML Modeling. Theory, Practice, Video Course*. St. Petersburg, Professional'naja literature Publ., 2010, 649 p.
12. Knuth D.E. *The Art of Computer Programming, V. 1. Fundamental Algorithms*. 3<sup>rd</sup> ed. Addison-Wesley Professional, 1997, 672 p.
13. Polikarpova N.I., Shalyto A.A. *Automata-Based Programming*. St. Petersburg, Piter Publ., 2011, 176 p. (in Russian)
14. Fan W., Li J., Ma S., Tang N., Wu Y., Wu Y. Graph pattern matching: From intractable to polynomial time. *Proceedings of the VLDB Endowment*, 2010, vol. 3, no. 1–2, pp. 264–275. <https://doi.org/10.14778/1920841.1920878>
15. Avdoshin S.M., Nabebin A.A. *Discrete Mathematics. Formal Logic Systems and Languages*. Moscow, DMK Press Publ., 2018, 390 p. (in Russian)
16. Levonevskiy D., Novikov F., Fedorchenko L., Afanasieva I. Verification of internet protocol properties using cooperating automaton objects. *Proc. of the 12<sup>th</sup> International Conference on Security of Information and Networks (SIN'19)*, 2019, pp. 1–4. <https://doi.org/10.1145/3357613.3357639>
17. Dijkstra E.W. *A Discipline of Programming*. 3<sup>rd</sup> ed. Englewood Cliffs, N.J., Prentice Hall, 1976, 217 p.
18. Shalyto A.A. Validation of state machine specifications. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 2, pp. 436–438. (in Russian). <https://doi.org/10.17586/2226-1494-2023-23-2-436-438>

## Литература

1. Harel D., Pnueli D. On the development of reactive systems // *Logics and Models of Concurrent Systems*. Berlin, Heidelberg: Springer, 1985. P. 477–498. [https://doi.org/10.1007/978-3-642-82453-1\\_17](https://doi.org/10.1007/978-3-642-82453-1_17)
2. Карпов Ю.Г. *Model Checking. Верификация параллельных и распределенных программных систем*. СПб.: БХВ-Петербург, 2010. 560 с.
3. Sommerville I. *Software Engineering / 10<sup>th</sup> ed.* Boston: Pearson, 2020.
4. Hinchey M., Coyle L. Evolving critical systems: a research agenda for computer-based systems // *Proc. of the 17<sup>th</sup> IEEE International Conference and Workshops on Engineering of Computer Based Systems*. 2010. P. 430–435. <https://doi.org/10.1109/ECBS.2010.56>
5. Вельдер С.Э., Лукин М.А., Шальто А.А., Яминов Б.Р. *Верификация автоматных программ*. СПб.: Наука, 2011. 244 с.
6. Fedorchenko L., Baranov S. Equivalent transformations and regularization in context-free grammars // *Cybernetics and Information Technologies*. 2015. V. 14. N 4. P. 29–44. <https://doi.org/10.1515/cait-2014-0003>
7. Новиков Ф.А., Афанасьева И.В. Кооперативное взаимодействие автоматных объектов // *Информационно-управляющие системы*. 2016. № 6. С. 50–64. <https://doi.org/10.15217/issn1684-8853.2016.6.50>
8. Афанасьева И.В., Новиков Ф.А., Федорченко Л.Н. Методика построения событийно-управляемых программных систем с использованием языка спецификации CIAO // *Труды СПИИРАН*. 2020. Т. 19. № 3. С. 481–514. <https://doi.org/10.15622/sp.2020.19.3.1>
9. Aho A.V., Lam M.S., Sethi R., Ullman J.D. *Compilers: Principles, Techniques, and Tools / 2<sup>nd</sup> ed.* Boston: Pearson/Addison-Wesley, 2007. 1009 p.
10. Афанасьева И.В., Новиков Ф.А. Архитектура программного обеспечения систем оптической регистрации // *Информационно-управляющие системы*. 2016. № 3. С. 51–63. <https://doi.org/10.15217/issn1684-8853.2016.3.51>
11. Новиков Ф.А., Иванов Д.Ю. *Моделирование на UML. Теория, практика, видеокурс*. СПб.: Профессиональная литература, 2010. 640 с.
12. Knuth D.E. *The Art of Computer Programming, V. 1. Fundamental Algorithms / 3<sup>rd</sup> ed.* Addison-Wesley Professional, 1997. 672 p.
13. Поликарпова Н.И., Шальто А.А. *Автоматное программирование*. СПб.: Питер, 2011. 176 с.
14. Fan W., Li J., Ma S., Tang N., Wu Y., Wu Y. Graph pattern matching: From intractable to polynomial time // *Proceedings of the VLDB Endowment*. 2010. V. 3. N 1–2. P. 264–275. <https://doi.org/10.14778/1920841.1920878>
15. Авдошин С.М., Набебин А.А. *Дискретная математика. Формально-логические системы и языки*. М.: ДМК Пресс, 2018. 390 с.
16. Levonevskiy D., Novikov F., Fedorchenko L., Afanasieva I. Verification of internet protocol properties using cooperating automaton objects // *Proc. of the 12<sup>th</sup> International Conference on Security of Information and Networks (SIN'19)*. 2019. P. 1–4. <https://doi.org/10.1145/3357613.3357639>
17. Dijkstra E.W. *A Discipline of Programming / 3<sup>rd</sup> ed.* Englewood Cliffs, N.J.: Prentice Hall, 1976. 217 p.
18. Шальто А.А. Валидация автоматных спецификаций // *Научно-технический вестник информационных технологий, механики и оптики*. 2023. Т. 23. № 2. С. 436–438. <https://doi.org/10.17586/2226-1494-2023-23-2-436-438>

## Authors

**Irina V. Afanasieva** — PhD, Head of Laboratory, Special Astrophysical Observatory of the Russian Academy of Sciences (SAO RAS), Nizhny Arkhyz, 369167, Russian Federation, [sc 57210431774](https://orcid.org/0000-0003-4225-4124), [https://orcid.org/0000-0003-4225-4124](mailto:riv615@gmail.com), [riv615@gmail.com](mailto:riv615@gmail.com)

**Fedor A. Novikov** — D.Sc., Senior Researcher, Professor, Peter the Great St. Petersburg Polytechnic University (SPbPU), Saint Petersburg, 195251, Russian Federation; Professor, Alferov Academic University, Saint Petersburg, 194021, Russian Federation, [sc 16441904500](https://orcid.org/0000-0003-4450-0173), <https://orcid.org/0000-0003-4450-0173>, [fedornovikov51@gmail.com](mailto:fedornovikov51@gmail.com)

**Ludmila N. Fedorchenko** — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, [sc 36561350100](https://orcid.org/0000-0002-4008-9316), <https://orcid.org/0000-0002-4008-9316>, [lnf@iias.spb.su](mailto:lnf@iias.spb.su)

*Received 02.05.2023*

*Approved after reviewing 16.06.2023*

*Accepted 26.07.2023*

## Авторы

**Афанасьева Ирина Викторовна** — кандидат технических наук, заведующий лабораторией, Специальная астрофизическая обсерватория Российской академии наук, Нижний Архыз, 369167, Российская Федерация, [sc 57210431774](https://orcid.org/0000-0003-4225-4124), <https://orcid.org/0000-0003-4225-4124>, [riv615@gmail.com](mailto:riv615@gmail.com)

**Новиков Федор Александрович** — доктор технических наук, старший научный сотрудник, профессор, Санкт-Петербургский политехнический университет Петра Великого, Санкт-Петербург, 195251, Российская Федерация; профессор, Санкт-Петербургский национальный исследовательский Академический университет имени Ж.И. Алфёрова, Российской академии наук, Санкт-Петербург, 194021, Российская Федерация, [sc 16441904500](https://orcid.org/0000-0003-4450-0173), <https://orcid.org/0000-0003-4450-0173>, [fedornovikov51@gmail.com](mailto:fedornovikov51@gmail.com)

**Федорченко Людмила Николаевна** — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 36561350100](https://orcid.org/0000-0002-4008-9316), <https://orcid.org/0000-0002-4008-9316>, [lnf@iias.spb.su](mailto:lnf@iias.spb.su)

*Статья поступила в редакцию 02.05.2023*

*Одобрена после рецензирования 16.06.2023*

*Принята к печати 26.07.2023*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-757-766

УДК 004.89, 004.4

## Интеллектуальная система адаптивного тестирования

Лилия Фаритовна Тагирова<sup>1</sup>, Татьяна Михайловна Зубкова<sup>2</sup>✉

<sup>1,2</sup> Оренбургский государственный университет, Оренбург, 460018, Российская Федерация

<sup>1</sup> LG-77@mail.ru, <https://orcid.org/0000-0002-3388-9462>

<sup>2</sup> bars87@mail.ru✉, <https://orcid.org/0000-0001-6831-1006>

### Аннотация

**Введение.** Современное обучение невозможно без систем автоматизированной проверки знаний. Наиболее прогрессивными в настоящее время являются адаптивные модели тестирования, в которых сложность заданий меняется в зависимости от правильности ответов испытуемого. Представлена разработка интеллектуальной системы адаптивного тестирования с использованием аппарата нечеткой математики. **Метод.** Разработана интеллектуальная система адаптивного тестирования, в модуле, реализующим экспертную систему, используется продукционная база правил. Входными параметрами при работе экспертной системы являются: процент верных ответов, степень правильности ответа, длительность ответа, число попыток. Выходными — изменение текущего уровня подготовки обучаемого, на основании которого подобраны тестовые вопросы соответствующей сложности. В качестве метода логического вывода применен метод Мамдани, который состоит из шести операционных действий: фаззификации — преобразование точных значений входных переменных в значения лингвистических переменных посредством функций принадлежности; проектирования нечеткой базы правил экспертной системы; агрегирования подусловий — определение истинности условий для каждого лингвистического правила системы нечеткого вывода; активации подзаклучений — нахождение степени истинности каждого из подзаклучений в лингвистическом правиле; аккумуляции заключений — нахождение функции принадлежности для каждой из выходных лингвистических переменных; дефаззификации — нахождение численного значения для каждой из выходных лингвистических переменных. **Основные результаты.** Представлена разработанная интеллектуальная система адаптивного тестирования, которая позволяет на основании анализа результатов прохождения тестов определить текущий уровень подготовки обучающегося и адаптировать материал к уровню его подготовки. Система динамически, в режиме реального времени, предоставляет вопросы соответствующей сложности. **Обсуждение.** При использовании рассмотренной интеллектуальной системы адаптивного тестирования обучающимся будут заданы вопросы соответствующего уровня сложности, что позволит выстроить индивидуальную траекторию обучения. Внедрение предложенной системы обеспечит реализацию персонализированного подхода к организации учебного процесса, повысит точность оценки знаний обучающихся, и, как результат, повысит качество их обучения.

### Ключевые слова

искусственный интеллект, экспертная система, нечеткая логика, нечеткая математика, тестирование обучаемых, адаптивное тестирование, интеллектуальная система

**Ссылка для цитирования:** Тагирова Л.Ф., Зубкова Т.М. Интеллектуальная система адаптивного тестирования // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 757–766. doi: 10.17586/2226-1494-2023-23-4-757-766

**Intelligent adaptive testing system**Liliya F. Tagirova<sup>1</sup>, Tatyana M. Zubkova<sup>2</sup>✉<sup>1,2</sup> Orenburg State University, Orenburg, 460000, Russian Federation<sup>1</sup> lg-77@mail.ru, <https://orcid.org/0000-0002-3388-9462><sup>2</sup> bars87@mail.ru✉, <http://orcid.org/0000-0001-6831-1006>**Abstract**

Modern learning is impossible without automated knowledge testing systems. At present, the most progressive are adaptive testing models in which the complexity of tasks varies depending on the correctness of the patient's answers. This article describes the development of an intelligent adaptive testing system using a fuzzy mathematics device. An intelligent adaptive testing system has been developed; the module that implements the expert system uses the production base of the rules. The input parameters of testing are the percentage of correct responses, the degree of correctness of the response, the duration of the response, and the number of attempts. The output is a change in the current level of training of the student on the basis of which test questions of related complexity are selected. As a method of logical inference, the Mamdani method is used which consists of six operational actions: phazification — conversion of exact values of input variables into values of linguistic variables through belonging functions, this served as the basis for designing a fuzzy base of rules of the expert system; aggregation of sub-conditions — determination of the truth of conditions for each linguistic rule of the fuzzy inference system; activating sub-conclusions — finding the degree of truth of each of the sub-conclusions in the linguistic rule; accumulation of conclusions — finding the belonging function for each of the output linguistic variables; defuzzification — finding a numerical value for each of the output linguistic variables. A developed intelligent adaptive testing system (ISAT) is presented that allows, based on the analysis of test results, to determine the current level of training of students, to adapt the material to the level of their training. This system allows you to dynamically present questions of appropriate complexity in real time. When using the developed intelligent adaptive testing system, students will be provided with questions of the appropriate level of complexity, this will allow building an individual learning trajectory. The introduction of a predefined system will ensure the implementation of a personalized approach for organizing the learning process; will increase the accuracy of assessing students' knowledge. The use of the technology of fuzzy expert systems allows for automated, intelligent control of students' knowledge.

**Keywords**

artificial intelligence, expert system, fuzzy logic, fuzzy mathematics, trainee testing, adaptive testing, intelligent system

**For citation:** Tagirova L.F., Zubkova T.M. Intelligent adaptive testing system. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 757–766 (in Russian). doi: 10.17586/2226-1494-2023-23-4-757-766

**Введение**

Современные тенденции развития мирового электронного образовательного пространства определяют необходимость цифровой трансформации образования за счет разработки и внедрения различных электронных обучающих систем и ресурсов.

Выпущен ряд документов, регламентирующих деятельность образовательных организаций в условиях цифровой трансформации образования<sup>1,2,3,4</sup>.

<sup>1</sup> Распоряжение Минпросвещения России от 18.05.2020 N P-44 «Об утверждении методических рекомендаций для внедрения в основные общеобразовательные программы современных цифровых технологий» [Электронный ресурс]. Режим доступа: <https://legalacts.ru/doc/rasporjzhenie-minprosveshchenija-rossii-ot-18052020-n-r-44-ob-utverzhenii/> (дата обращения: 01.04.2023).

<sup>2</sup> Указ Президента Российской Федерации от 21.07.2020 № 474 «О национальных целях развития Российской Федерации на период до 2030 года» [Электронный ресурс]. Режим доступа: <http://www.kremlin.ru/acts/bank/45726> (дата обращения: 01.04.2023).

<sup>3</sup> Паспорт Федерального проекта «Цифровая образовательная среда» [Электронный ресурс]. Режим доступа: [https://minobr.gov-murman.ru/files/Nach\\_proekty/Tsos/fp\\_cifrovaya\\_obrazovatel'naya\\_sreda\\_04102019.pdf](https://minobr.gov-murman.ru/files/Nach_proekty/Tsos/fp_cifrovaya_obrazovatel'naya_sreda_04102019.pdf) (дата обращения: 01.04.2023).

<sup>4</sup> Приказ Министерства просвещения РФ от 2 декабря 2019 г. № 649 «Об утверждении Целевой модели цифровой образовательной среды» [Электронный

Согласно<sup>2</sup>, суть цифровой трансформации образования заключается в достижении необходимых образовательных результатов посредством персонализации учебного процесса на основе использования инновационных достижений в областях аппаратного и программного обеспечения современных цифровых технологий. Процесс цифровой трансформации образования инициирует совершенствование всей системы образования, в том числе средств оценки учебных достижений и управления образовательным процессом [1].

Перспективным направлением в области персонализации образования выступает создание и внедрение интеллектуальных систем адаптивного тестирования [2]. Адаптивными обучающими системами являются электронные образовательные ресурсы, предоставляющие обучающемуся индивидуальную траекторию обучения за счет «подстраивания» персонального образовательного пространства под его собственные особенности [2]. Системы адаптивного тестирования применяются многими отечественными и зарубежными исследователями, предлагающими различные подходы к ее реализации.

Проведенный анализ научных работ [2–13] показал, что наибольшее количество современных теоретических и практических исследований в области разработки систем адаптивного тестирования связано с

[Электронный ресурс]. Режим доступа: <https://cdnstatic.rg.ru/uploads/attachments/180/25/89/56962.pdf> (дата обращения: 01.04.2023).

применением искусственных нейронных сетей (ИНС) в различных предметных областях. Например, в работе [2] предложено использование ИНС для автоматизации процесса выставления оценок при разработке тренажерно-обучающей системы при подготовке специалистов в авиационной области В [3] применена технология ИНС при решении задачи выбора сложности вопроса адаптивной системы тестирования.

В работе [4] использован алгоритм Байеса для проектирования адаптивной модели тестирования нескольких компетенций обучающихся. В [5] с помощью алгоритма Байеса предложено адаптивно проводить персонализированную формирующую оценку учащегося, динамически выбирая элементы и тесты в соответствии с их способностями.

В [6] предложен модульный принцип построения адаптивной системы тестирования для обучения иностранному языку курсантов военных вузов. В работе [7] использована марковская математическая модель динамического адаптивного тестирования активного агента для разработки системы оценки знаний студентов.

На основе биологической теории развития человека Ж. Пиаже предложил алгоритм разработки динамических адаптивных тестов-тренажеров для диагностики процессуальных характеристик обучения учащихся временам английского языка [8]. В работе [9] использован критикующий подход к построению экспертных систем для реализации адаптивного тестирования, а в [10] представлена модель адаптивного тестирования с использованием нечеткой математики.

В работе [11] разработана гибкая онлайн-платформа для компьютеризированного адаптивного тестирования, в которой принята во внимание особенность неоднородности обучающихся. Приложение позволяет динамически подстраивать не только тестовые задания, но и пользовательский интерфейс.

В [12] учтен эффект забывания со временем обучающимися материала дисциплины. Предложено реализовать адаптивную формирующую систему оценки, основанную на компьютерном тестировании и цикле памяти обучения, чтобы обеспечить повторную оценку знаний учащихся.

В [13] разработана компьютеризированная система адаптивного тестирования для измерения уровня владения английским языком обучающихся на основе шестизападной процедуры.

Проблематика разработки адаптивных тестов является актуальной не только в рамках решения задач тестирования обучающихся, но и в других сферах, требующих проверки уровня компетенций, а также личностных интеллектуальных и психофизиологических качеств испытуемого. Например, в медицине: при разработке компьютеризированной адаптивной тест-системы функциональной оценки инсульта [14], психологии при оценке физических навыков мышления компьютеризированного адаптивного тестирования [15].

В настоящей работе в качестве примера представлено решение проблемы разработки системы адаптивного тестирования для обучающихся подготовки в ИТ-сфере при изучении специальной дисциплины.

## Постановка задачи

Для реализации поставленной цели работы выполнено создание интеллектуальной системы адаптивного тестирования (ИСАТ), позволяющей оценивать знания обучающихся. Для этого разработана структура обучающей системы, определены информационные потоки и инструменты для адаптации ранжированного тестового материала.

Структура обучающей системы (рис. 1) состоит из шести блоков: трех базовых (классических структурных единиц): «Блок учебный», «Блок методический» и «Блок электронного журнала» и трех интеллектуальных, направленных на разностороннюю адаптацию обучающей системы к особенностям обучающихся: «Блок адаптации пользовательского интерфейса», «Блок адаптации учебного контента» и «Блок адаптированного тестирования».

Данная электронная обучающая система аналогична действиям, которые проводит преподаватель в условиях традиционной формы на занятиях. Вопросы, задаваемые системой, разбиваются на несколько уровней по мере возрастания сложности вопросов. На первом этапе всем испытуемым выдается одинаковый входной тест, цель которого – осуществление предварительной дифференциации обучающихся. По результатам дифференциации на втором этапе организуется адаптивный режим, и строятся соответствующие тесты [4].

Результаты ответов на вопросы каждого уровня обуславливают выбор следующего уровня (может повышаться, понижаться, оставаться прежним). Комплексная оценка формируется с учетом процента правильных ответов, длительности ответа, степени правильности ответа, числа попыток.

На рис. 2 представлена диаграмма потоков данных процесса автоматизации адаптивного тестирования в Case-средстве BP-WIN в нотации DFD. С системой могут взаимодействовать преподаватель (выступает в роли «Эксперт») и обучающийся (в роли «Студент»), которые входят в систему с разными правами доступа.

В качестве инструмента для адаптации ранжированного тестового материала использована экспертная система, ядром которой является нечеткая база знаний. В качестве математического метода использован алгоритм нечеткого вывода Мамдани, в котором применена минимаксная композиция нечетких множеств [16].

Таким образом, система адаптивного тестирования позволила максимально приблизить процесс электронного контроля к непосредственному контролю знаний преподавателем, при этом оценка становится более полной и достоверной, чем при использовании обычного тестирования [4].

## Математическая модель экспертной системы

Для реализации интеллектуального контроля знаний использована 12-балльная шкала, преимущество которой, по сравнению с традиционной 5-балльной — более высокая различающая способность [10]. Данная шкала

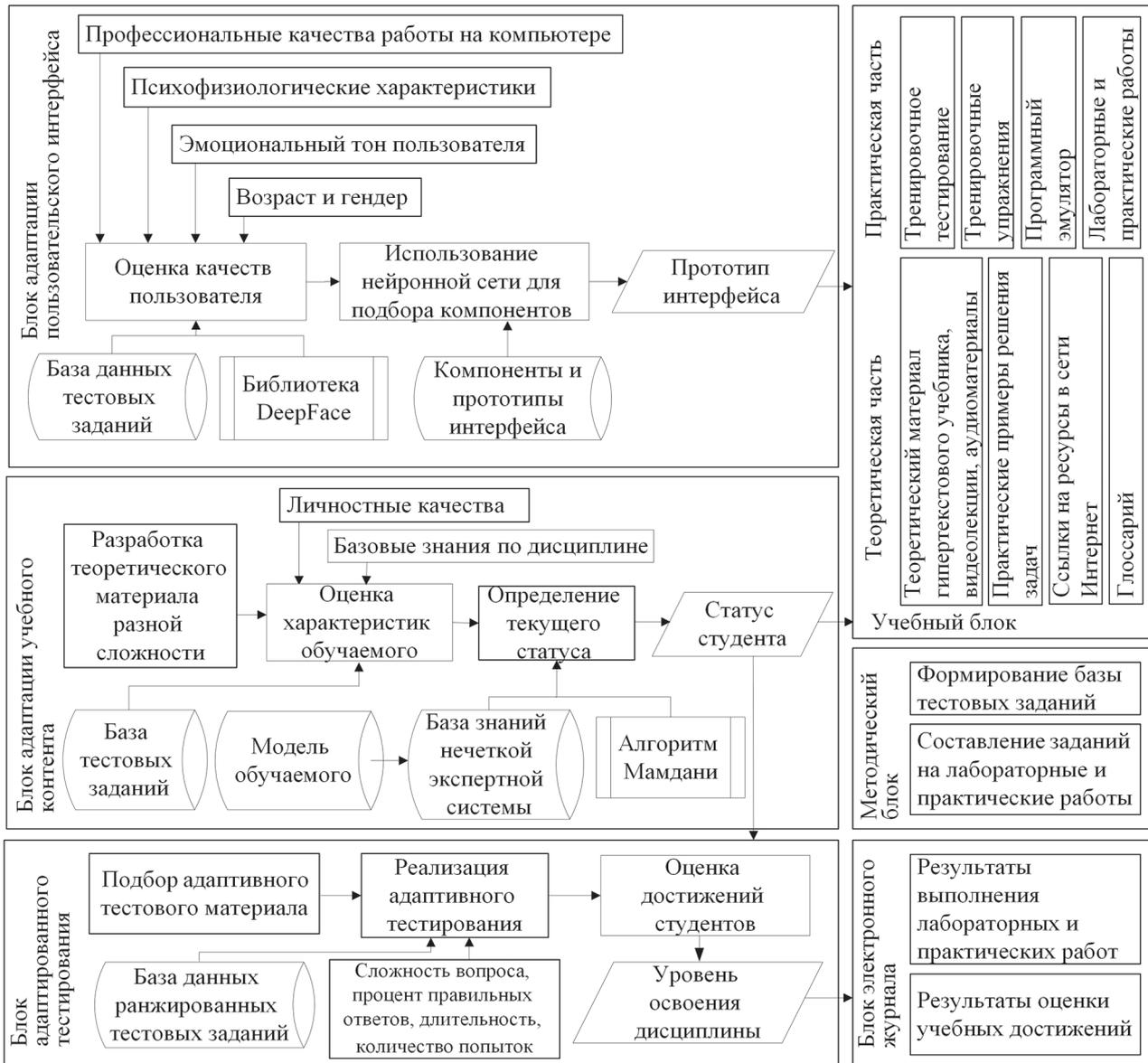


Рис. 1. Структура и взаимосвязь компонентов электронной обучающей системы

Fig. 1. Structure and interconnection of electronic training system components

позволила описать результат с помощью нечетких понятий «менее чем ...» и «более чем», что соответствует оценкам типа «5 с минусом», «3 с плюсом» и т. д., которые часто используются преподавателями для оценки учебных достижений обучающихся.

На рис. 3 показано соответствие 12-балльной и существующей 5-балльной шкал оценивания.

На рис. 3 видно, что в 12-балльной шкале определены возможные оценки для испытуемого на основе нечетких множеств: «нет знаний», «неудовлетворительно», «малоудовлетворительно», «почти удовлетворительно», «удовлетворительно», «более чем удовлетворительно», «почти хорошо», «хорошо», «более чем хорошо», «почти отлично», «отлично», «превосходно». Данные уровни изменяются в зависимости от успеха отвечающего. Для каждого уровня сложности «Эксперту» (преподавателю по дисциплине) необходимо разработать соответствующий набор вопросов. Тестирование

проводится по методу «вопрос — ответ». После ответа на каждое задание производится оценка правильности ответа. В качестве инструмента для определения текущего уровня знаний обучающегося использована нечеткая экспертная система, а в качестве алгоритма — алгоритм нечеткого вывода Мамдани, который состоит из шести операционных действий.

**1. Фаззификация.** Преобразование точных значений входных переменных в значения лингвистических переменных посредством функций принадлежности. На начальном этапе характеристики обучающихся описаны в виде входных лингвистических переменных, заданы их термы и параметры [16].

Модель нечеткого вывода содержит четыре входных переменных: процент верных ответов на тест (OT), длительность ответа (DO), степень правильности ответа (SPO), число попыток (KP) и одну выходную переменную изменение уровня подготовки (CHP).



Рис. 2. Диаграмма потоков данных процесса адаптации тестового материала  
 Fig. 2. Test material adaptation process data flow diagram



Рис. 3. Соответствие 5- и 12-балльных шкал оценивания  
 Fig. 3. Compliance with 5- and 12-point rating scales

Лингвистическая переменная «Процент правильных ответов» описывает в процентах число верных ответов на тесты и содержит шесть термов: «очень низкий», «низкий», «средний», «выше среднего», «высокий», «очень высокий».

График функции принадлежности лингвистической переменной «Процент правильных ответов» представлен на рис. 4.

Для значения входной лингвистической переменной «Длительность ответа» принято единое время для ответа на все тесты — 5 мин (установлено как оптимальное время). Отрицательные значения отклонения означают, что

обучаемый затратил на ответ меньше времени, чем установлено в качестве оптимального для данного задания.

Для оценки степени правильности ответа применен способ с использованием нечеткой математики. Предложена следующая нечеткая шкала оценки: «неправильно», «неточно», «не совсем правильно», «почти правильно», «правильно».

Лингвистическая переменная «число попыток» может принимать значения «небольшое», «среднее число попыток», «большое число попыток».

Выходной лингвистической переменной является переменная «Изменение уровня подготовки», со-

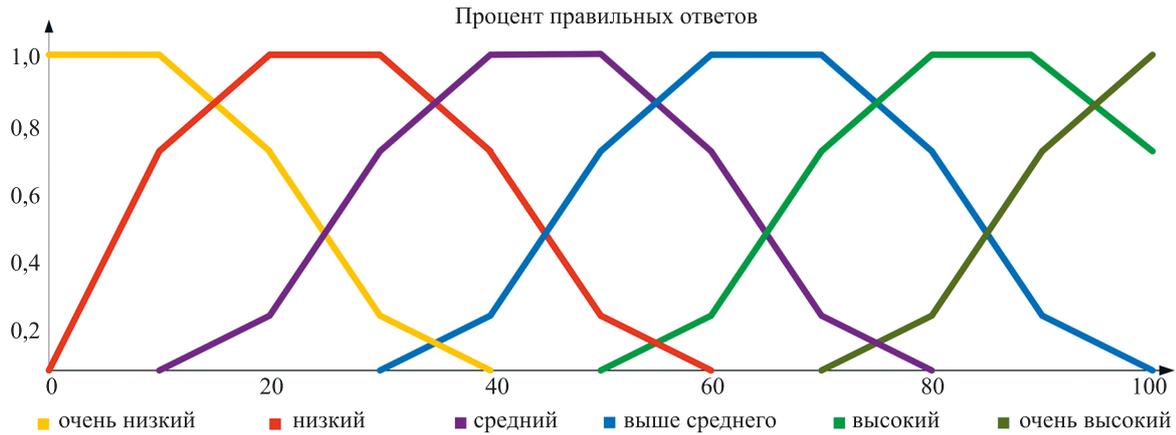


Рис. 4. График функции принадлежности лингвистической переменной «Процент правильных ответов»

Fig. 4. Graph of the membership function of the linguistic variable “Percentage of correct answers”

держащая пять термов: «уменьшить на два уровня», «уменьшить на один уровень», «не изменять уровень подготовки», «увеличить на один уровень», «увеличить на два уровня».

**2. Проектирование нечеткой базы правил экспертной системы.** Описание параметров лингвистических переменных послужило основой для проектирования нечеткой базы правил экспертной системы. Прототип базы правил представлен в таблице. Условные обозначения использованы аналогичные фаззификации.

**3. Агрегирование подусловий.** Определение истинности условий для каждого лингвистического правила системы нечеткого вывода. Так как во всех правилах использована операция «И», то агрегирование проведено по методу логической конъюнкции, которое заключается в нахождении минимального значения истинности всех его подусловий:

$$b_i = \min \{ \mu(OT), \mu(DO), \mu(KP), \mu(SPO) \},$$

где  $i$  — общее количество лингвистических правил.

**4. Активация подзаключений.** Нахождение степени истинности каждого из подзаключений в лингвистическом правиле. Для каждого подзаключения в соответствии с функцией принадлежности выходной лингвистической переменной  $\mu(CHP)$  и результатом предыдущего операционного действия получена степень истинности. Определено минимальное значение функции принадлежности подзакключений:

$$\mu'(CHP) = \min \{ c_i, \mu(CHP) \},$$

где  $\mu(CHP)$  — функция принадлежности выходной лингвистической переменной  $CHP$ ;  $c_i$  — элемент множества  $C = \{c_i\}$  который определен как произведение элементов множества истинности условий  $B = \{b_i\}$  и значений весовых коэффициентов  $F_i$  для каждого из лингвистических правил системы нечеткого вывода.

**5. Аккумуляция правил заключения.** Нахождение функции принадлежности для каждой из выходных лингвистических переменных  $CHP = \{CHP_i\}$ , осуществ-

ляемое на основе max-объединения функций принадлежности:

$$\mu''(CHP) = \max \{ \mu'_i(CHP) \},$$

где  $\mu'_i(CHP)$  — функция принадлежности, полученная в результате активации лингвистических правил.

**6. Дефаззификация.** Нахождение численного значения для каждой из выходных лингвистических переменных множества ST. Результатом данного операционного действия получение количественных значений выходной переменной «Изменение уровня подготовки». Дефаззификация выполнена по методу левого модального значения:

$$y = \min \{ CHP_M \},$$

где  $CHP_M$  — модальное значение нечеткого множества для выходной переменной  $CHP$  после аккумуляции.

В результате работы нечеткого логического вывода для каждого обучающегося определен текущий уровень подготовки, в соответствии с которым подобрана наиболее оптимальная редакция тестового материала, максимально адаптированная под его индивидуальные особенности [16].

### Практическая реализация ИСАТ

На основании описания нечетких множеств и правил вывода реализована ИСАТ [17, 18]. В предложенном программном средстве предусмотрено два уровня доступа: «Студент» и «Эксперт». «Студент» имеет возможность: изучить предложенный теоретический материал; выполнить лабораторные и практические работы; пройти тестирование.

При первом запуске программного средства под уровнем доступа «Студент» необходимо пройти входное тестирование для определения начального уровня подготовки. Окно прохождения тестирования представлено на рис. 5, а, после тестирования будет выведен результат с указанием изменения текущего уровня подготовки обучающегося (рис. 5, б).

Таблица. Фрагмент базы правил

Table. Rule base fragment

ЕСЛИ	ОТ=очень высокий	И	ДО=очень короткая	И	КР=очень малое	И	СПО=правильно	ТО	СНР=Увеличить на два уровня
ЕСЛИ	ОТ=очень высокий	И	ДО=короткая	И	КР=очень малое	И	СПО=правильно	ТО	СНР=Увеличить на два уровня
ЕСЛИ	ОТ=очень высокий	И	ДО=очень короткая	И	КР=очень малое	И	СПО=почти правильно	ТО	СНР=Увеличить на два уровня
ЕСЛИ	ОТ=высокий	И	ДО=короткая	И	КР=очень малое	И	СПО=почти правильно	ТО	СНР=Увеличить на два уровня
ЕСЛИ	ОТ=высокий	И	ДО=короткая	И	КР=малое	И	СПО=почти правильно	ТО	СНР=Увеличить на один уровень
ЕСЛИ	ОТ=высокий	И	ДО=короткая	И	КР=среднее	И	СПО=правильно	ТО	СНР=Увеличить на один уровень
ЕСЛИ	ОТ=высокий	И	ДО=короткая	И	КР=малое	И	СПО=почти правильно	ТО	СНР=Увеличить на один уровень
ЕСЛИ	ОТ=высокий	И	ДО=очень короткая	И	КР=малое	И	СПО=почти правильно	ТО	СНР=Увеличить на один уровень
ЕСЛИ	ОТ=выше среднего	И	ДО=средняя	И	КР=среднее	И	СПО=почти правильно	ТО	СНР=Оставить прежний уровень
ЕСЛИ	ОТ=выше среднего	И	ДО=средняя	И	КР=малое	И	СПО=почти правильно	ТО	СНР=Оставить прежний уровень
ЕСЛИ	ОТ=выше среднего	И	ДО=короткая	И	КР=среднее	И	СПО=почти правильно	ТО	СНР=Оставить прежний уровень
ЕСЛИ	ОТ=низкий	И	ДО=длинная	И	КР=большое	И	СПО=не правильно	ТО	СНР=Уменьшить на один уровень
ЕСЛИ	ОТ=низкий	И	ДО=очень длинная	И	КР=очень малое	И	СПО=неточно	ТО	СНР=Уменьшить на один уровень
ЕСЛИ	ОТ=низкий	И	ДО=длинная	И	КР=малое	И	СПО=не правильно	ТО	СНР=Уменьшить на один уровень
ЕСЛИ	ОТ=очень низкий	И	ДО=очень длинная	И	КР=большое	И	СПО=не правильно	ТО	СНР=Уменьшить на два уровня
ЕСЛИ	ОТ=очень низкий	И	ДО=очень длинная	И	КР=большое	И	СПО=неточно	ТО	СНР=Уменьшить на два уровня
ЕСЛИ	ОТ=очень низкий	И	ДО=длинная	И	КР=большое	И	СПО=не правильно	ТО	СНР=Уменьшить на два уровня

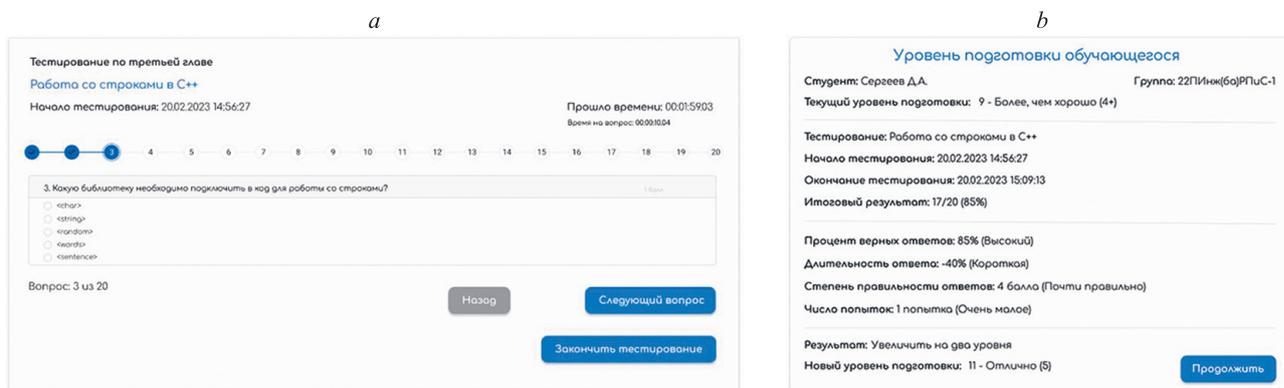


Рис. 5. Вид окон прохождения тестирования (а) и результатов определения текущего уровня подготовки обучающегося (б)

Fig. 5. View of the windows for passing the test (a) and the results of determining the current level of training of the trainee (b)

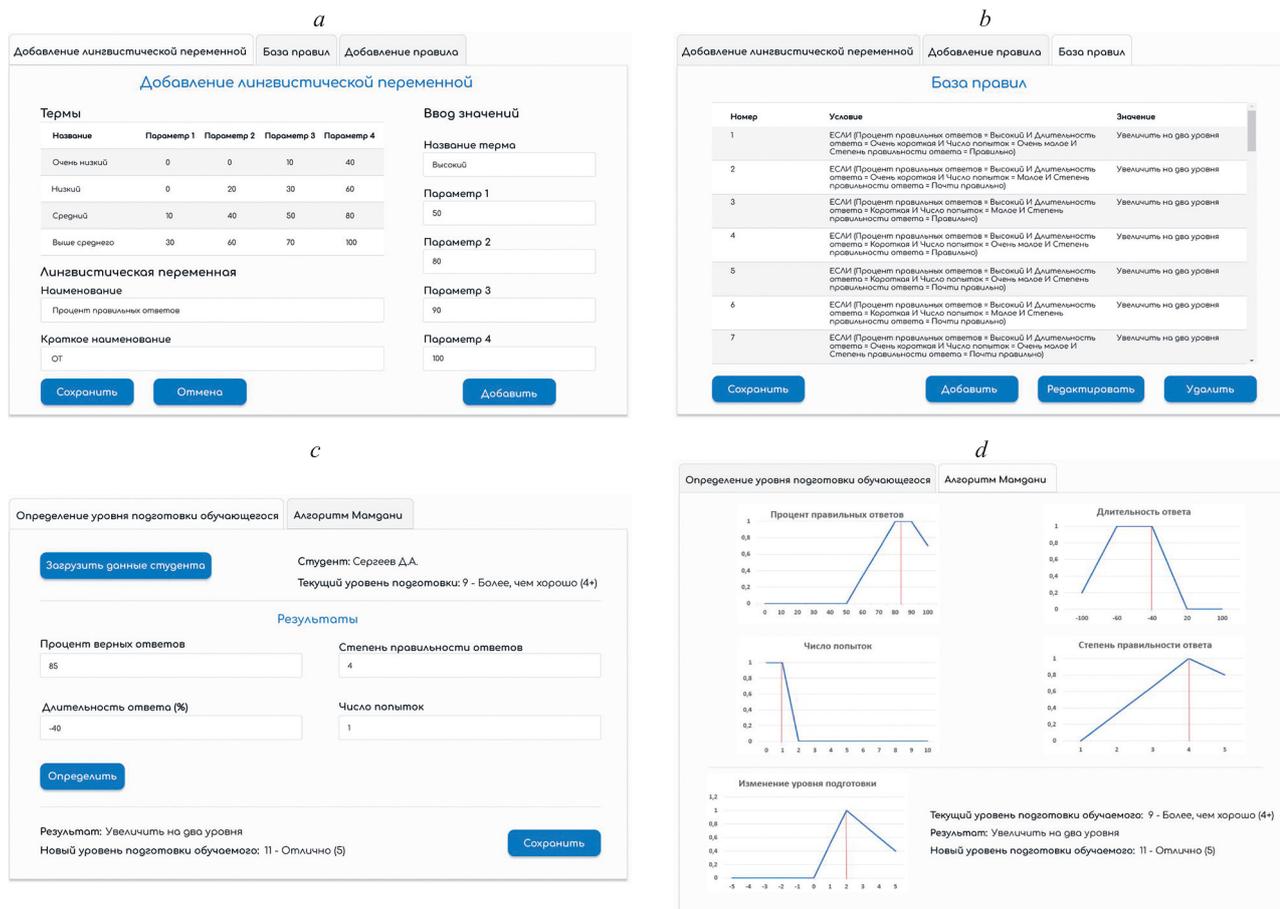


Рис. 6. Вид окон: работы с лингвистическими переменными (а); базы правил нечеткой экспертной системы (б); определения уровня подготовки для роли «Эксперт» (с) и работы алгоритма Мамдани (д)

Fig. 6. View of windows: working with linguistic variables (a); the rule bases of the fuzzy expert system (b); determining the level of training for the role of “Expert” (c) and the operation of the Mamdani algorithm (d)

Отметим, что обучающемуся могут быть определены темы повторного изучения для повышения уровня подготовки.

При работе с интеллектуальной системой под уровнем доступа «Эксперт» предоставлены более широкие функциональные возможности.

Преподавателю доступны следующие функции: заполнение базы данных, которая включает теоретический материал; задания для практических и лабораторных работ; тестовый материал; настройка параметров сложности каждого теста (вопросы, которые имеют больший вес, объединяются в более сложный тест (предусмотрено двенадцать уровней сложности)); правильность ответов и рекомендованное время для ответа на вопрос. Также возможно установить число попыток прохождения теста, при этом время между попытками произвольное, но должно укладываться в рамки учебного процесса. Преподаватель заполняет и настраивает базу правил — задает лингвистические переменные, на основании которых формируется база правил экспертной системы.

На рис. 6, а приведен пример работы с лингвистической переменной «Процент правильных ответов», где задаются термы, параметры и их значения. Для работы с остальными характеристиками выполняются аналогичные действия.

При настройке экспертной системы значения функций принадлежности всех нечетких множеств могут быть изменены «Экспертом». Количество уровней измерения (детализация) каждого параметра также может быть изменено, это повлечет за собой изменение правил, использующих данный параметр.

После описание всех лингвистических переменных «Эксперт» имеет возможность добавления новых правил. В результате формируется база правил нечеткой экспертной системы (рис. 6, б).

Также «Эксперту» доступна функция просмотра текущего уровня подготовки выбранного обучающегося (рис. 6, в) и возможность выбора способа получения данного результата во вкладке «Алгоритм Мамдани» (рис. 6, д).

На рис. 6, д показана реализация этапа дефазификации, и как сработало правило. В данной реализации алгоритма использован метод дефазификации левого модального значения.

## Заключение

Разработанная интеллектуальная система адаптивного тестирования позволила на основании анализа результатов тестирования адекватно оценить уровень подготовки обучающихся. Использование технологии

нечетких экспертных систем дало возможность осуществить автоматизированный и интеллектуальный контроль знаний обучающихся.

Подтверждено, что интеллектуальная система адаптивного тестирования эффективно собирает персонализированную диагностическую информацию о каждом обучаемом, на основании которой предлагается тестовый материал соответствующей сложности.

В ходе проведенной работы:

- исследованы информационные потоки, протекающие при прохождении оценки знаний обучающихся и предоставления адаптивных тестовых заданий. Определена структура и взаимосвязь компонентов интеллектуальной системы адаптивного тестирования;

- выявлены основные параметры, на основании которых реализована оценка текущего уровня подготовки обучаемого: процент верных ответов на тест, степень правильности ответа, длительность ответа, число попыток;

- разработана экспертная система, которая позволяет на основе анализа входных данных определять уровень подготовки обучающихся. В качестве алгоритма нечеткого вывода использован алгоритм Мамдани.

Применение систем адаптивного тестирования с внедрением методов искусственного интеллекта позволяет реализовать управление процессом оценки знаний и процессом обучения в целом, что соответствует современным требованиям и задачам, которые стоят перед высшей школой подготовки специалистов.

### Литература

1. Семенова Н.Г., Томина И.П. Разработка и применение электронных образовательных ресурсов в условиях цифровой трансформации образования: монография. Оренбург: ОГУ, 2022. 139 с.
2. Григорьев А.П., Бурлуцкий С.Г. Нейросетевая навигационная тренажерно-обучающая система // Информационно-управляющие системы. 2017. № 3. С. 89–98. <https://doi.org/10.15217/issnl684-8853.2017.3.89>
3. Чумакова Е.В., Корнеев Д.Г., Гаспарян М.С. Разработка метода адаптивного тестирования на основе нейротехнологий // Открытое образование. 2022. Т. 26. № 2. С. 4–13. <https://doi.org/10.21686/1818-4243-2022-2-4-13>
4. Гусятников В.Н., Соколова Т.Н., Безруков А.И., Каюкова И.В. Адаптивная модель тестирования нескольких компетенций на основе алгоритма Байеса // Современные наукоемкие технологии. 2022. № 1. С. 40–46. <https://doi.org/10.17513/snt.39007>
5. Choi Y., McClenen C. Development of adaptive formative assessment system using computerized adaptive testing and dynamic Bayesian networks // Applied Sciences. 2020. V. 10. N 22. P. 8196. <https://doi.org/10.3390/app10228196>
6. Герасимова Т.Н., Гусева Н.В. Использование профессионально адаптивного тестирования в процессе обучения иностранному языку курсантов военных вузов // Мир науки, культуры, образования. 2019. № 3(76). С. 224.
7. Перегудова И.П. Динамическое адаптивное тестирование учебной деятельности студентов при изучении времен английского языка // Международный научно-исследовательский журнал. 2020. № 10-2(100). С. 40–45. <https://doi.org/10.23670/IRJ.2020.100.10.042>
8. Бровка Н.В., Дьячук П.П., Носков М.В., Перегудова И.П. Марковская математическая модель динамического адаптивного тестирования активного агента // Информатика и образование. 2018. № 10(299). С. 29–35. <https://doi.org/10.32517/0234-0453-2018-33-10-29-35>
9. Частикова В.А., Колесник Н.М. Система адаптивного тестирования на основе критикующего подхода к построению экспертных систем // Электронный сетевой политематический журнал «Научные труды КубГТУ». 2018. № 3. С. 506–517.
10. Дуплик С.В. Модель адаптивного тестирования на нечеткой математике // Открытое и дистанционное образование. 2004. № 4(16). С. 78–88.
11. Oppl S., Reisinger F., Eckmaier A., Helm C. A flexible online platform for computerized adaptive testing // International Journal of Educational Technology in Higher Education. 2017. V. 14. N 1. P. 2. <https://doi.org/10.1186/s41239-017-0039-0>
12. Yang A.C.M., Flanagan B., Ogata H. Adaptive formative assessment system based on computerized adaptive testing and the learning memory cycle for personalized learning // Computers and Education: Artificial Intelligence. 2022. V. 3. P. 100104. <https://doi.org/10.1016/j.caeai.2022.100104>
13. Huang H.T.D., Hung Sh.T.A., Chao H.-Y., Chen J.-H., Lin T.-P., Shih C.-L. Developing and validating a computerized adaptive testing

### References

1. Semenova N.G., Tomina I.P. *Development and Application of Electronic Educational Resources in the Context of Education Digital Transformation*. Orenburg, OSU, 2022, 139 p. (in Russian)
2. Grigoryev A.P., Burlutsky S.G. A neuronet navigational training system. *Information and Control Systems*, 2017, no. 3, pp. 89–98. (in Russian), <https://doi.org/10.15217/issnl684-8853.2017.3.89>
3. Chumakova E.V., Korneev D.G., Gasparian M.S. Development of adaptive testing method based on neurotechnologies. *Open Education*, 2022, vol. 26, no. 2, pp. 4–13. (in Russian). <https://doi.org/10.21686/1818-4243-2022-2-4-13>
4. Gusyatinikov V.N., Sokolova T.N., Bezrukov A.I., Kayukova I.V. Adaptive model for testing several competencies based on the Bayes algorithm. *Modern high technologies*, 2022, no. 1, pp. 40–46. (in <https://doi.org/10.17513/snt.39007>)
5. Choi Y., McClenen C. Development of adaptive formative assessment system using computerized adaptive testing and dynamic Bayesian networks. *Applied Sciences*, 2020, vol. 10, no. 22, pp. 8196. <https://doi.org/10.3390/app10228196>
6. Gerasimova T.N., Guseva N.V. The use of professional computer-based testing systems in the process of foreign language teaching in military universities. *The world of science, culture and education*, 2019, no. 3(76), pp. 224. (in Russian)
7. Peregudova I.P. Dynamic adaptive testing of educational activities of students studying tenses of english language. *International Research Journal*, 2020, no. 10-2(100), pp. 40–45. (in Russian). <https://doi.org/10.23670/IRJ.2020.100.10.042>
8. Dyachuk P.P., Brovka N.V., Noskov M.V., Peregudova I.P. Markov mathematical model of dynamic adaptive testing of an active agent. *Informatics and Education*, 2018, no. 10(299), pp. 29–35. (in Russian). <https://doi.org/10.32517/0234-0453-2018-33-10-29-35>
9. Chastikova V.A., Kolesnik N.M. Adaptive testing system based on the critical approach to building expert systems. *Scientific Works of the Kuban State Technological University*, 2018, no. 3, pp. 506–517. (in Russian)
10. Duplik C.V. Model of adaptive testing on fuzzy mathematics. *Open and Distance Education*, 2004, no. 4(16), pp. 78–88. (in Russian)
11. Oppl S., Reisinger F., Eckmaier A., Helm C. A flexible online platform for computerized adaptive testing. *International Journal of Educational Technology in Higher Education*, 2017, vol. 14, no. 1, pp. 2. <https://doi.org/10.1186/s41239-017-0039-0>
12. Yang A.C.M., Flanagan B., Ogata H. Adaptive formative assessment system based on computerized adaptive testing and the learning memory cycle for personalized learning. *Computers and Education: Artificial Intelligence*, 2022, vol. 3, pp. 100104. <https://doi.org/10.1016/j.caeai.2022.100104>
13. Huang H.T.D., Hung Sh.T.A., Chao H.-Y., Chen J.-H., Lin T.-P., Shih C.-L. Developing and validating a computerized adaptive testing system for measuring the English proficiency of Taiwanese EFL university students. *Language Assessment Quarterly*, 2022, vol. 19, no. 2, pp. 162–188. <https://doi.org/10.1080/15434303.2021.1984490>

- system for measuring the English proficiency of Taiwanese EFL university students // *Language Assessment Quarterly*. 2022. V. 19. N 2. P. 162–188. <https://doi.org/10.1080/15434303.2021.1984490>
14. Lin G.-H., Huang Y.-J., Lee Sh.-Ch., Huang Sh.-L., Hsieh Ch.-L. Development of a computerized adaptive testing system of the functional assessment of stroke // *Archives of Physical Medicine and Rehabilitation*. 2018. V. 99. N 4. P. 676–683. <https://doi.org/10.1016/j.apmr.2017.09.116>
  15. Istiyono E., Dwandaru W., Setiawan R., Megawati I. Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use // *European Journal of Educational Research*. 2020. V. 9. N 1. P. 91–101. <https://doi.org/10.12973/eu-jer.9.1.91>
  16. Горбаченко В.И., Ахметов Б.С., Кузнецова О.Ю. Интеллектуальные системы: нечеткие системы и сети: учебное пособие для вузов / 2-е изд., испр. и доп. М.: Юрайт, 2023. 105 с.
  17. Тагирова Л.Ф., Черняков А.А. Программная система адаптивного тестирования студентов дистанционного обучения. Свидетельство о государственной регистрации программы для ЭВМ RU2022681117. Бюл. 2022. № 11.
  18. Тагирова Л.Ф., Бачковская Ю.С. Интеллектуальная система адаптивного тестирования. Свидетельство о государственной регистрации программы для ЭВМ RU2023662375. Бюл. 2023. № 6.
  14. Lin G.-H., Huang Y.-J., Lee Sh.-Ch., Huang Sh.-L., Hsieh Ch.-L. Development of a computerized adaptive testing system of the functional assessment of stroke. *Archives of Physical Medicine and Rehabilitation*, 2018, vol. 99, no. 4, pp. 676–683. <https://doi.org/10.1016/j.apmr.2017.09.116>
  15. Istiyono E., Dwandaru W., Setiawan R., Megawati I. Developing of computerized adaptive testing to measure physics higher order thinking skills of senior high school students and its feasibility of use. *European Journal of Educational Research*, 2020, vol. 9, no. 1, pp. 91–101. <https://doi.org/10.12973/eu-jer.9.1.91>
  16. Gorbachenko V.I., Akhmetov B.S., Kuznetcova O.Iu. *Intelligence Systems: Fuzzy Systems and Networks*. Moscow, Jurajt Publ., 2023, 105 p. (in Russian)
  17. Tagirova L.F., Cherniakov A.A. Program system for adaptive testing of the distance learning students. *Certificate of the state computer program registration RU2022681117*, 2022. (in Russian)
  18. Tagirova L.F., Bachkovskaia Iu.S. Intelligence system of adaptive testing. *Certificate of the state computer program registration RU2023662375*, 2023. (in Russian)

### Авторы

**Тагирова Лилия Фаритовна** — кандидат педагогических наук, доцент, доцент, Оренбургский государственный университет, Оренбург, 460018, Российская Федерация, [sc 57210446135](https://orcid.org/0000-0002-3388-9462), <https://orcid.org/0000-0002-3388-9462>, [LG-77@mail.ru](mailto:LG-77@mail.ru)

**Зубкова Татьяна Михайловна** — доктор технических наук, профессор, профессор, Оренбургский государственный университет, Оренбург, 460018, Российская Федерация, [sc 57202282917](https://orcid.org/0000-0001-6831-1006), <https://orcid.org/0000-0001-6831-1006>, [bars87@mail.ru](mailto:bars87@mail.ru)

Статья поступила в редакцию 19.04.2023  
Одобрена после рецензирования 08.06.2023  
Принята к печати 26.07.2023

### Authors

**Liliya F. Tagirova** — PhD (Education), Associate Professor, Associate Professor, Orenburg State University, Orenburg, 460018, Russian Federation, [sc 57210446135](https://orcid.org/0000-0002-3388-9462), <https://orcid.org/0000-0002-3388-9462>, [LG-77@mail.ru](mailto:LG-77@mail.ru)

**Tatyana M. Zubkova** — D.Sc., Full Professor, Orenburg State University, Orenburg, 460018, Russian Federation, [sc 57202282917](https://orcid.org/0000-0001-6831-1006), <https://orcid.org/0000-0001-6831-1006>, [bars87@mail.ru](mailto:bars87@mail.ru)

Received 19.04.2023  
Approved after reviewing 08.06.2023  
Accepted 26.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И КОГНИТИВНЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ  
ARTIFICIAL INTELLIGENCE AND COGNITIVE INFORMATION TECHNOLOGIES

doi: 10.17586/2226-1494-2023-23-4-767-775

УДК 004.855.5

**Нейросетевой метод визуального распознавания голосовых команд водителя с использованием механизма внимания****Александр Александрович Аксёнов<sup>1</sup>✉, Елена Витальевна Рюмина<sup>2</sup>,  
Дмитрий Александрович Рюмин<sup>3</sup>, Денис Викторович Иванько<sup>4</sup>,  
Алексей Анатольевич Карпов<sup>5</sup>**<sup>1,2,3,4,5</sup> Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, 199178, Российская Федерация<sup>1</sup> [axyonov.a@iias.spb.su](mailto:axyonov.a@iias.spb.su)✉, <https://orcid.org/0000-0002-7479-2851><sup>2</sup> [ryumina.e@iias.spb.su](mailto:ryumina.e@iias.spb.su), <https://orcid.org/0000-0002-4135-6949><sup>3</sup> [ryumin.d@iias.spb.su](mailto:ryumin.d@iias.spb.su), <https://orcid.org/0000-0002-7935-0569><sup>4</sup> [ivanko.d@iias.spb.su](mailto:ivanko.d@iias.spb.su), <https://orcid.org/0000-0003-0412-7765><sup>5</sup> [karpov@iias.spb.su](mailto:karpov@iias.spb.su), <https://orcid.org/0000-0003-3424-652X>**Аннотация**

**Введение.** Визуальное распознавание речи или автоматическое чтение речи по губам все чаще применяется для преобразования речи в текст. Видеоданные доказывают свою необходимость в системах мультимодального распознавания речи, особенно когда использование акустических данных затруднено в виду сильных аудиозумов или недоступно. Основная цель исследования заключается в повышении эффективности распознавания команд водителя путем анализа визуальной информации для снижения тактильного взаимодействия с различными автомобильными системами (мультимедийными и навигационными, телефонными звонками и др.) во время вождения. **Метод.** Предложен метод автоматического чтения речи водителя по губам в процессе управления транспортным средством на основе глубокой нейронной сети архитектуры 3DResNet18. Выполнен анализ динамической информации о движении губ диктора с помощью 3D-сверточных слоев нейросети. Использование нейросетевой архитектуры с двунаправленной моделью Long Short-Term Memory и механизмом внимания позволяет добиться более высокой точности распознавания при незначительном снижении скорости работы.

**Основные результаты.** Предложены и исследованы два варианта нейросетевых архитектур для визуального распознавания речи. При использовании первой нейросетевой архитектуры результат распознавания голосовых команд водителя составил 77,68 %, что ниже на 5,78 %, по сравнению со второй. Скорость работы системы определена показателем реального времени (Real-Time Factor, RTF), значение которого для первой нейросетевой архитектуры равен 0,076, а второй — 0,183, что выше более чем в два раза. Предложенный метод апробирован на данных дикторов многомодального корпуса RUSAVIC, записанных в автомобиле. **Обсуждение.** Результаты исследования могут найти применение в системах аудиовизуального распознавания речи. Подобные системы могут быть рекомендованы для применения в сильно зашумленных условиях, например, в процессе управления транспортным средством. Проведенный анализ позволил выбрать оптимальную нейросетевую модель визуального распознавания речи для последующего встраивания в ассистивную систему на базе мобильного устройства.

**Ключевые слова**

голосовые команды водителя, визуальное распознавание речи, автоматическое чтение речи по губам, машинное обучение, CNN, LSTM, механизм внимания

**Благодарности**

Исследование выполнено при поддержке РФФИ (проект № 19-29-09081-мк), ведущей научной школы Российской Федерации (грант № НШ-17.2022.1.6) и за счет средств государственного финансирования, тема FFZF-2022-0005.

**Ссылка для цитирования:** Аксёнов А.А., Рюмина Е.В., Рюмин Д.А., Иванько Д.В., Карпов А.А. Нейросетевой метод визуального распознавания голосовых команд водителя с использованием механизма внимания // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 767–775. doi: 10.17586/2226-1494-2023-23-4-767-775

© Аксёнов А.А., Рюмина Е.В., Рюмин Д.А., Иванько Д.В., Карпов А.А., 2023

## Neural network-based method for visual recognition of driver's voice commands using attention mechanism

Alexandr A. Axyonov<sup>1</sup>✉, Elena V. Ryumina<sup>2</sup>, Dmitry A. Ryumin<sup>3</sup>, Denis V. Ivanko<sup>4</sup>, Alexey A. Karpov<sup>5</sup>

St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation

<sup>1</sup> axyonov.a@iias.spb.su✉, <https://orcid.org/0000-0002-7479-2851>

<sup>2</sup> ryumina.e@iias.spb.su, <https://orcid.org/0000-0002-4135-6949>

<sup>3</sup> ryumin.d@iias.spb.su, <https://orcid.org/0000-0002-7935-0569>

<sup>4</sup> ivanko.d@iias.spb.su, <https://orcid.org/0000-0003-0412-7765>

<sup>5</sup> karpov@iias.spb.su, <https://orcid.org/0000-0003-3424-652X>

### Abstract

Visual speech recognition or automated lip-reading systems actively apply to speech-to-text translation. Video data proves to be useful in multimodal speech recognition systems, particularly when using acoustic data is difficult or not available at all. The main purpose of this study is to improve driver command recognition by analyzing visual information to reduce touch interaction with various vehicle systems (multimedia and navigation systems, phone calls, etc.) while driving. We propose a method of automated lip-reading the driver's speech while driving based on a deep neural network of 3DResNet18 architecture. Using neural network architecture with bi-directional LSTM model and attention mechanism allows achieving higher recognition accuracy with a slight decrease in performance. Two different variants of neural network architectures for visual speech recognition are proposed and investigated. When using the first neural network architecture, the result of voice recognition of the driver was 77.68 %, which was lower by 5.78 % than when using the second one the accuracy of which was 83.46 %. Performance of the system which is determined by a real-time indicator RTF in the case of the first neural network architecture is equal to 0.076, and the second — RTF is 0.183 which is more than two times higher. The proposed method was tested on the data of multimodal corpus RUSAVIC recorded in the car. Results of the study can be used in systems of audio-visual speech recognition which is recommended in high noise conditions, for example, when driving a vehicle. In addition, the analysis performed allows us to choose the optimal neural network model of visual speech recognition for subsequent incorporation into the assistive system based on a mobile device.

### Keywords

driver's voice commands, visual speech recognition, automatic lip reading, machine learning, CNN, LSTM, attention mechanisms

### Acknowledgements

The study was supported by the Russian Foundation for Basic Research (project no. 19-29-09081-mk), the leading scientific school of the Russian Federation (grant no. NSH-17.2022.1.6) and at the expense of state funding, topic FFZF-2022-0005.

**For citation:** Axyonov A.A., Ryumina E.V., Ryumin D.A., Ivanko D.V., Karpov A.A. Neural network-based method for visual recognition of driver's voice commands using attention mechanism. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 767–775 (in Russian). doi: 10.17586/2226-1494-2023-23-4-767-775

### Введение

На сегодняшний день не существует надежных систем автоматического распознавания речи, устойчивых к динамическим акустическим шумам, которые могли бы использоваться в реальных условиях вождения транспортного средства. Управление навигационной системой, кондиционером/смартфоном с применением сенсорного интерфейса может привести к отвлечению внимания водителя и стать причиной дорожно-транспортных происшествий (ДТП). Практически акустический шум является актуальной проблемой в данной области [1]. Фоновый шум оказывает влияние не только на микрофон, но и вынуждает говорящего повышать громкость голоса для того, чтобы компенсировать уровень шума в ушах (так называемый «эффект Ломбарда»). В реальных условиях применение изменения речевой активности, вызываемое шумовым воздействием на органы слуха, может повлиять на эффективность работы системы распознавания больше, чем акустический шум [2].

Современные технологии автоматического визуального распознавания речи (Visual Speech Recognition, VSR) позволяют распознавать речь людей, анализируя движения губ и лица. Системы на основе этих технологий имеют множество применений, таких как идентификация говорящего, преобразование речи в текст, а также голосовое управление техническими устройствами. Применение технологий автоматического «чтения речи по губам» в автомобильной отрасли может способствовать развитию ассистивных систем помощи водителю [3].

В последние годы для повышения эффективности распознавания речи некоторые исследователи используют визуальную информацию и анализируют, как чтение речи по губам может способствовать повышению эффективности в распознавании речи на основе аудиоинформации [4]. Визуальные сигналы содержат достаточный объем информации, позволяющий распознавать произносимые фразы [5, 6], и совместное использование этих модальностей является шагом к созданию робастной системы распознавания речи в сложных акустических условиях [7].

В связи с этим использование систем на основе анализа визуальной информации позволяет взаимодействовать с автомобильными информационно-развлекательными системами в режиме «свободных рук», снижая когнитивную нагрузку и повышая безопасность дорожного движения [8, 9].

Отметим, что внедрение систем анализа визуальной информации в автомобильной отрасли сопряжено с определенными трудностями из-за вариативности уровня освещения и расположения камер, которые, в свою очередь, влияют на точность обнаружения и отслеживания лица и рта говорящего [10]. Несмотря на эти проблемы, недавние исследования демонстрируют перспективность применения таких систем в условиях вождения [11]. Подобные системы могут быть эффективным инструментом для распознавания и отслеживания состояний сонливости и усталости водителя, что может способствовать снижению числа ДТП на дорогах [12, 13].

Основная цель настоящей работы — повышение эффективности распознавания команд водителя путем анализа визуальной информации для снижения физического взаимодействия с различными автомобильными системами (мультимедийными и навигационными системами, звонками и др.) во время вождения.

### Краткий обзор предметной области

В последние годы системы VSR используют технологии глубокого обучения для извлечения информативных признаков и машинной классификации. Различные типы нейросетей, такие как сверточные нейронные сети (Convolutional Neural Network, CNN), сети прямого распространения и автоэнкодеры, используются в системах чтения речи по губам [14, 15]. Одним из перспективных подходов в VSR являются 3D CNN, которые применяются на входе системы, поскольку они отлично справляются с анализом пространственной и временной информации, а также извлечением информативных признаков [16–18]. Некоторые исследователи в дополнение к этому интегрируют механизмы пространственного внимания для обработки извлеченных признаков [19, 20]. Механизм пространственного внимания позволяет модели фокусироваться на наиболее информативных областях во входных видеокдрах.

Для распознавания слитной речи (словосочетания или фразы) также используются нейросети обработки последовательностей — рекуррентные нейронные сети (Recurrent Neural Network, RNN) [21]. В последнее время в качестве альтернативы RNN для классификации стали применять модели трансформеров с использованием механизмов внимания и временные сверточные сети [22, 23]. Однако для обучения таких моделей требуются большие вычислительные мощности, а также значительный объем обучающих данных. Потому одним из популярных подходов в таких случаях является трансферное обучение [24] — метод использования предварительно обученной модели для улучшения прогнозов в рамках другой, но схожей задачи, позволяющий сократить время обучения и уменьшить

потребность в данных, а также повысить производительность нейронной сети [25].

Существуют подходы, в которых распознавание речи основывается на моделях глубоких машин Больцмана (Deep Boltzmann Machine, DBM) [26]. DBM используются для извлечения признаков из визуальной модальности, которые потом объединяются с вектором признаков дискретного косинусного преобразования [27]. Далее проводится линейный дискриминантный анализ [28], используемый для анализа корреляции признакового пространства и уменьшения его размерности. Для моделирования и распознавания речи применяются модели смеси гауссовских распределений (Gaussian Mixture Model, GMM) [29] и скрытые марковские модели.

Кроме одномодального визуального распознавания речи широко используется многомодальное распознавание речи — аудиовизуальное распознавание речи (Audio Visual Speech Recognition, AVSR). В последние годы такие подходы демонстрируют многообещающие результаты [30, 31]. В работах [32, 33] предложено использовать иерархическую многомодальную нейросетевую архитектуру для AVSR. В работе [34] рассмотрен метод, основанный на нейросетевой архитектуре типа «последовательность-последовательность» (sequence-to-sequence) для автоматического обучения объединенному представлению из аудио- и видеомодальностей на основе его информативности, тем самым повышая устойчивость к шумам при распознавании речи за счет использования дополнительной визуальной информации.

Методы, предложенные в перечисленных работах, демонстрируют потенциал систем визуального распознавания речи, которые также могут найти применение в автомобильной сфере.

### Многомодальный корпус аудиовизуальной русской речи RUSAVIC

В настоящей работе использован собственный корпус данных Russian Audio-Visual Speech in Cars (RUSAVIC) [35], созданный по разработанной методологии [36]. Корпус данных RUSAVIC разработан специально для решения задач распознавания речи на основе наиболее часто встречающихся голосовых команд водителей и предназначен для использования в системах помощи водителю. Основные характеристики многодикторного мультимодального корпуса представлены в табл. 1.

Каждый диктор произнес 62 управляющие голосовые команды (например, «Найти заправку», «Набрать номер», «Карта», «Включить свет в салоне») не менее 10 раз в течение нескольких сеансов записи. В корпусе содержатся данные, записанные в реальных условиях вождения, а также в транспортном средстве, припаркованном около оживленного перекрестка.

Корпус данных RUSAVIC обрабатывал информацию с помощью извлечения аудиосигнала из мультимедийной информации, далее применялся детектор голосовой активности (Voice Activity Detection, VAD). В работе использована специально обученная модель

Таблица 1. Основные характеристики корпуса данных RUSAVIC

Table 1. RUSAVIC corpus parameters

Характеристика	Значение
Количество дикторов	20
Разрешение видеоданных, пиксел	1920 × 1080
Частота кадров, кадр/с	60
Количество команд в словаре	62
Количество сеансов записи	10 (для каждого диктора)
Общее число фраз	около 12 400

обнаружения голосовой активности Vosk<sup>1</sup>, которая способна достаточно точно обнаружить границы речи даже в зашумленных условиях. Результат применения VAD — получение мультимедийного файла с произнесенной фразой из словаря, в котором отсутствует лишняя информация (тишина до начала речи и после).

### Метод визуального распознавания голосовых команд водителя

Метод визуального распознавания голосовых команд водителя включал в себя два этапа (рис. 1). На этапе 1 выполнена сегментация входного видеосигнала. На каждом кадре из входной последовательности определена графическая область интереса (Region of Interest, ROI), которой является область рта диктора. Детектирование ROI выполнено с помощью программной системы MediaPipe FaceMesh [37], которая определила 468 трехмерных лицевых ориентиров-маркеров. Извлеченные области губ подвергались преобразованию в градации серого, нормализации до изображения 88 × 88 пикселов. Далее области губ разбивались на сегменты, состоящие из 120 кадров с частичным перекрытием в 48 кадров (40 %). При недостатке кадров в конце последовательности недостающие кадры заполнялись последним кадром последовательности.

На этапе 2 проведено распознавание голосовой команды с помощью предварительно обученной нейросетевой модели. Визуальная речь на выходе нейросети декодирована в понятный формат и сопоставлена с метками в словаре. Распознанная голосовая команда водителя определена максимальным значением в выходном векторе вероятностных значений.

### Обучение нейросетевой модели

Процесс обучения нейросетевой модели содержал параметры, которые могут влиять на эффективность распознавания обученной модели. В процессе обучения после каждой эпохи выполняется определение точности на обучающем и валидационном наборах данных. Часто точность распознавания на обучающем

наборе может достигать высоких значений, вплоть до 100 %. Это значит, что модель переобучается и ее эффективность на тестовом и валидационном наборах будет низкой. Для генерации дополнительного набора обучающих данных применяются различные методы аугментации данных.

В настоящей работе на 40 % изображений и меток, которые выбраны случайным образом, применен метод аугментации данных MixUP [38]. Коэффициент слияния двух изображений и их меток (классов команд водителя) варьировался от 30 до 70 % (при сумме равной 100 % соблюдается условие нулевой прозрачности для генерируемого изображения). Для оставшихся 60 % бинарных векторов меток после применения MixUP использован метод сглаживания Label Smoothing [39].

Рассмотрим две наиболее эффективные для задачи исследования архитектуры нейросетевых моделей (рис. 2).

**Архитектура 3DResNet18 + SA.** Сегментированные последовательности поступают на вход обучаемой нейросетевой модели (рис. 2, а), далее слой 3D Conv свертывает их в пространстве и времени. Для извлечения наиболее значимых признаков из последовательностей использован слой подвыборки MaxPooling 3D, способный уменьшать размерности входного тензора. Отметим, что слой MaxPooling 3D позволяет ускорить обучение, а также уменьшить количество параметров модели, при этом сохраняя ее эффективность. Модифицированные остаточные блоки Residual Blocks модели ResNet18 [40] извлекают карты признаков размерностью 4 × 3 × 3 × 512 для всей входной последовательности. В последнем остаточном блоке присутствует механизм внимания (Squeeze-and-Attention, SA) [41], использование которого улучшает результаты распознавания. В SA сначала выполняется операция «сжатия» (squeeze), которая сокращает размерность последовательности данных, а операция «внимания» (attention), позволяет модели сосредоточиться на наиболее важных аспектах данных, игнорируя несущественные детали.

Признаки, поступающие с выхода сверточных слоев, преобразуются в вектор фиксированной длины за счет использования слоя Global Average Pooling 3D, суть которого заключается в том, чтобы усреднить значения признаков по всему объему 3D-пространства, полученные с последнего сверточного слоя в сети. Результатом является один признак для каждого фильтра сверточного слоя. Эти признаки затем используются для классификации. В архитектуре также применяется техника регуляризации нейросетей Dropout, которая помогает предотвратить переобучение модели.

Полученные одномерные вектора размерностью 512 признаков поступают в полносвязный слой с функцией активации ReLU, выполняющей нелинейное преобразование входных данных. Размерность этого слоя увеличивается до 1024, тем самым улучшая представление признаков и повышая точность модели. В конце этот слой соединяется с последним полносвязным слоем (Fully connected, FC), содержащим 62 нейрона с функцией активации Softmax. Это означает, что на выходе формируется вероятность принадлежности входных данных к 62 классам, которые интерпретируются

<sup>1</sup> Vosk Api [Электронный ресурс]. Режим доступа: <https://github.com/alphacep/vosk-api>, свободный (дата обращения: 23.04.2023).

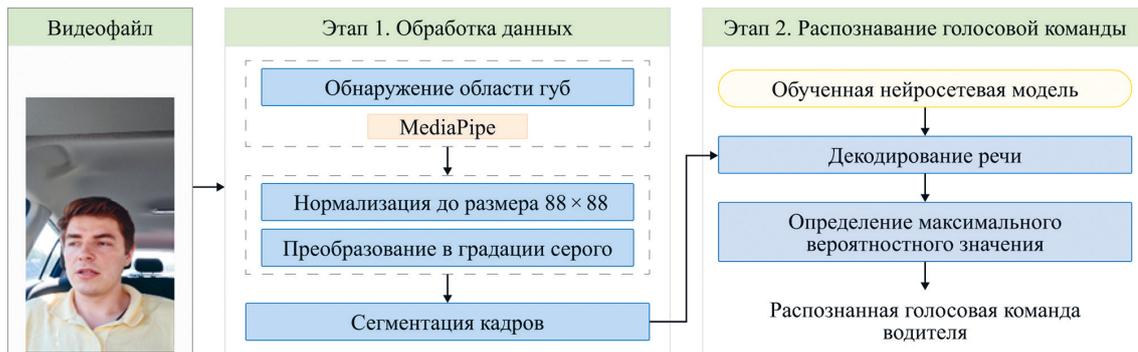


Рис. 1. Метод визуального распознавания команд водителя  
 Fig. 1. Method for visual recognition a driver's commands

как гипотезы предсказания модели для каждого из этих классов (62 голосовые команды).

Функция активации Softmax — одна из возможных функций активации, широко используемая в выходном слое нейронных сетей для многоклассовой классификации. Функция Softmax преобразует линейный выход полносвязного слоя в вероятность принадлежности каждого входа к определенному классу.

**Архитектура 3DResNet18 + SA + BiLSTM.** Вторая нейросетевая архитектура отличается от первой тем, что в ней 3D Conv обрабатывает только пространственное представление признаков. Для обработки временного пространства признаков использовано два слоя двунаправленных Long short-term memory

(LSTM) (рис. 2, b). Длина последовательности изображений в 120 кадров поступают в 3D сверточный слой и подвергаются операции подвыборки MaxPooling 3D. Из каждого изображения входной последовательности извлекаются карты признаков размерностью  $120 \times 3 \times 3 \times 512$  благодаря модифицированным остаточным блокам. Слой подвыборки Average Pooling преобразует полученные карты признаков в одномерные вектора размерностью  $120 \times 512$ , которые поступают на последующих два слоя двунаправленной LSTM. Первый слой является sequence-to-sequence, на входе и выходе которого 120 векторов признаков, а второй sequence-to-one — на выходе один вектор признаков размерностью 512. В конце полносвязный слой FC

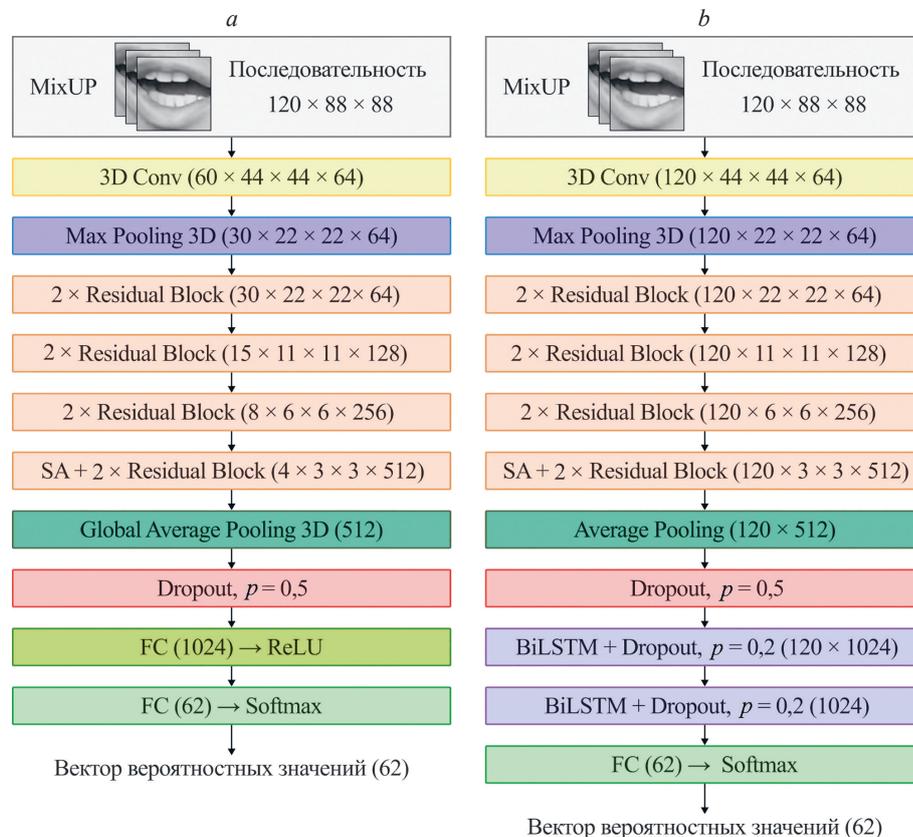


Рис. 2. Нейросетевые архитектуры: 3DResNet18 + SA (a); 3DResNet18 + SA+ BiLSTM (b)  
 Fig. 2. Neural network architectures: 3DResNet18 + SA (a); 3DResNet18 + SA+ BiLSTM (b)

Таблица 2. Сравнение эффективности архитектур нейросетевых моделей  
 Table 2. Comparison of neural network architectures performance

Архитектура нейросетевой модели	Время инициализации модели, с	Точность распознавания команд, %		RTF
		Val	Test	
3DResNet18	2,3	86,53	76,74	0,071
3DResNet18 + SA (рис. 2, а)	2,7	88,66	77,68	0,076
3DResNet18 + BiLSTM	7,9	82,94	75,89	0,153
3DResNet18 + SA+ BiLSTM (рис. 2, б)	8,2	85,55	83,46	0,183

с количеством нейронов 62 также формирует вектор вероятностных значений.

### Экспериментальные результаты

Приведем результаты экспериментальных исследований с различными архитектурами нейросетей. Апробация предложенного метода выполнена на корпусе данных RUSAVIC. Набор данных для обучения и оценивания нейросетевых моделей разбивался следующим образом: обучающий набор (train) включал 14 389 видеозаписей 13 дикторов, тестовый набор (test) — 4241 видеозапись трех дикторов и валидационный (val) — 2480 видеозаписей двух дикторов. В качестве планировщика скорости обучения применялась техника косинусного отжига (Cosine Annealing), значения которого варьировались от 0,0001 до 0,001. Изначальное количество эпох устанавливалось равным 100. После каждой эпохи выполняется промежуточное оценивание модели на валидационном наборе данных. В случае если в течение 10 эпох обучаемая модель не прогрессировала, то процесс обучения останавливался, а лучшие веса модели определялись по максимальной точности, достигнутой на валидационном наборе. Валидационная выборка использована для оптимизации модели, тогда как экспериментальные результаты представлены для тестовой выборки. В качестве оптимизатора применен Adam.

Программная реализация системы визуального распознавания голосовых команд выполнена на языке программирования Python v.3.9. Для машинного обучения моделей использован фреймворк с открытым исходным кодом TensorFlow v.2.8.0 в связке с расширением модуля TensorFlow-GPU v.2.8.0. Экспериментальные исследования осуществлены на персональном компьютере под управлением операционной системы Microsoft Windows 10 Pro со следующими техническими характеристиками: CPU — AMD Ryzen 9 5950X; GPU — GeForce RTX 3090 TI 24Gb; хранилище данных — твердотельный накопитель SSD M.2.

Проведено сравнение архитектур нейросетевых моделей, основанных на архитектуре ResNet18. Выбранные модели анализируют пространственно-временные зависимости последовательности входных изображений. В табл. 2 представлены две принципиально разные архитектуры и их модернизированные версии за счет добавления механизмов внимания (рис. 2).

Параметры обучения нейросетевых архитектур, представленных в сравнении, подбирались эмпириче-

ским путем и являются для каждой архитектуры оптимальными. Неизменные параметры: набор обучающих данных, длина последовательности и разрешение изображений, которые подаются на вход.

Исследуемые архитектуры сравнивались не только по показателю точности распознавания команд водителя, но и по таким показателям как: среднее время активации модели, размер модели и скорость распознавания по показателю реального времени (Real-Time Factor, RTF). Сравнение проводилось в одинаковых условиях. Для определения скорости работы системе предлагалось распознать 15 случайно выбранных фраз различной длительности из корпуса данных RUSAVIC. Для расчета времени инициализации модели и усредненного значения RTF описанная процедура производилась 10 раз.

По результатам анализа табл. 2 можно сделать вывод, что нейросетевая архитектура 3DResNet18 + SA + BiLSTM оказалась точнее архитектуры без добавления слоев BiLSTM. Однако эта архитектура немного проигрывает в скорости работы. Из результатов сравнения также следует, что при добавлении механизма внимания SA независимо от типа архитектуры точность распознавания голосовых команд повышается. Добавление SA в архитектуры нейросетевых моделей оказывает незначительное влияние на скорость работы системы. Также добавление различных методов аугментации, таких как Cosine WR, MixUP и Label Smoothig, дает прирост точности [42].

### Заключение

Предложен метод чтения речи по губам водителя во время управления транспортным средством с использованием модифицированной нейросетевой архитектуры ResNet18. Метод может быть встроен в системы распознавания речи, функционирующие в условиях сильного акустического шума, вызванного различными факторами, такими как скорость движения, покрытие дороги, степень открытия окон, наличие источников звука и др.

Результаты экспериментов показали, что метод позволяет эффективно распознавать произносимые водителем команды из словаря, состоящего из 62 русскоязычных фраз с точностью до 83,46 % и со скоростью распознавания RTF порядка 0,183. Метод можно использовать в режиме реального времени.

Для улучшения качества распознавания речи в дальнейшем планируется исследовать другие архитектуры глубоких нейросетей, а также расширять обучающую

базу данных. Предложенный метод визуального распознавания планируется использовать в рамках разработки системы аудиовизуального распознавания речи,

что может значительно повысить точность и устойчивость к шумам существующих систем распознавания речи.

### Литература

1. Lin S.C., Hsu C.H., Talamonti W., Zhang Y., Oney S., Mars J., Tang L. Adasa: A conversational in-vehicle digital assistant for advanced driver assistance features // Proc. of the 31<sup>st</sup> Annual ACM Symposium on User Interface Software and Technology. 2018. P. 531–542. <https://doi.org/10.1145/3242587.3242593>
2. Lee B., Hasegawa-Johnson M., Goudeseune C., Kamdar S., Borys S., Liu M., Huang T. AVICAR: Audio-visual speech corpus in a car environment // Proc. of the 8<sup>th</sup> International Conference on Spoken Language Processing. 2004. P. 2489–2492. <https://doi.org/10.21437/Interspeech.2004-424>
3. Ivanko D., Ryumin D., Kashevnik A., Axyonov A., Karpov A. Visual speech recognition in a driver assistance system // Proc. of the 30<sup>th</sup> European Signal Processing Conference (EUSIPCO). 2022. P. 1131–1135. <https://doi.org/10.23919/EUSIPCO55093.2022.9909819>
4. Xu B., Wang J., Lu C., Guo Y. Watch to listen clearly: Visual speech enhancement driven multi-modality speech recognition // Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). 2020. P. 1637–1646. <https://doi.org/10.1109/wacv45572.2020.9093314>
5. Afouras T., Chung, J.S., Senior A., Vinyals O., Zisserman A. Deep audio-visual speech recognition // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022. V. 44. N 12. P. 8717–8727. <https://doi.org/10.1109/TPAMI.2018.2889052>
6. Кухарев Г.А., Матвеев Ю.Н., Олейник А.Л. Алгоритмы взаимной трансформации изображений для систем обработки и поиска визуальной информации // Научно-технический вестник информационных технологий, механики и оптики. 2017. Т. 17. № 1. С. 62–74. <https://doi.org/10.17586/2226-1494-2017-17-1-62-74>
7. Shi B., Hsu W.N., Mohamed A. Robust self-supervised audio-visual speech recognition // Proc. of the International Conference INTERSPEECH. 2022. P. 2118–2122. <https://doi.org/10.21437/interspeech.2022-99>
8. Chand H.V., Karthikeyan J. CNN based driver drowsiness detection system using emotion analysis // Intelligent Automation & Soft Computing. 2022. V. 31. N 2. P. 717–728. <https://doi.org/10.32604/iasc.2022.020008>
9. Ivanko D., Kashevnik A., Ryumin D., Kitenko A., Axyonov A., Lashkov I., Karpov A. MIDriveSafely: Multimodal interaction for drive safely // Proc. of the 2022 International Conference on Multimodal Interaction (ICMI). 2022. P. 733–735. <https://doi.org/10.1145/3536221.3557037>
10. Biswas A., Sahu P.K., Chandra M. Multiple cameras audio visual speech recognition using active appearance model visual features in car environment // International Journal of Speech Technology. 2016. V. 19. N 1. P. 159–171. <https://doi.org/10.1007/s10772-016-9332-x>
11. Nambi A.U., Bannur S., Mehta I., Kalra H., Virmani A., Padmanabhan V.N., Bhandari R., Raman B. HAMS: Driver and driving monitoring using a smartphone // Proc. of the 24<sup>th</sup> Annual International Conference on Mobile Computing and Networking. 2018. P. 840–842. <https://doi.org/10.1145/3241539.3267723>
12. Kashevnik A., Lashkov I., Gurtov A. Methodology and mobile application for driver behavior analysis and accident prevention // IEEE Transactions on Intelligent Transportation Systems. 2020. V. 21. N 6. P. 2427–2436. <https://doi.org/10.1109/TITS.2019.2918328>
13. Jang S.W., Ahn B. Implementation of detection system for drowsy driving prevention using image recognition and IoT // Sustainability. 2020. V. 12. N 7. P. 3037. <https://doi.org/10.3390/su12073037>
14. Mishra R.K., Urolagin S., Jothi J.A.A., Gaur P. Deep hybrid learning for facial expression binary classifications and predictions // Image and Vision Computing. 2022. V. 128. P. 104573. <https://doi.org/10.1016/j.imavis.2022.104573>
15. Sunitha G., Geetha K., Neelakandan S., Pundir A.K.S., Hemalatha S., Kumar V. Intelligent deep learning based ethnicity recognition and classification using facial images // Image and Vision Computing. 2022. V. 121. P. 104404. <https://doi.org/10.1016/j.imavis.2022.104404>
16. Yuan Y., Tian C., Lu X. Auxiliary loss multimodal GRU model in audio-visual speech recognition // IEEE Access. 2018. V. 6. P. 5573–5583. <https://doi.org/10.1109/ACCESS.2018.2796118>

### References

1. Lin S.C., Hsu C.H., Talamonti W., Zhang Y., Oney S., Mars J., Tang L. Adasa: A conversational in-vehicle digital assistant for advanced driver assistance features. *Proc. of the 31<sup>st</sup> Annual ACM Symposium on User Interface Software and Technology*, 2018, pp. 531–542. <https://doi.org/10.1145/3242587.3242593>
2. Lee B., Hasegawa-Johnson M., Goudeseune C., Kamdar S., Borys S., Liu M., Huang T. AVICAR: Audio-visual speech corpus in a car environment. *Proc. of the 8<sup>th</sup> International Conference on Spoken Language Processing*, 2004, pp. 2489–2492. <https://doi.org/10.21437/Interspeech.2004-424>
3. Ivanko D., Ryumin D., Kashevnik A., Axyonov A., Karpov A. Visual speech recognition in a driver assistance system. *Proc. of the 30<sup>th</sup> European Signal Processing Conference (EUSIPCO)*, 2022, pp. 1131–1135. <https://doi.org/10.23919/EUSIPCO55093.2022.9909819>
4. Xu B., Wang J., Lu C., Guo Y. Watch to listen clearly: Visual speech enhancement driven multi-modality speech recognition. *Proc. of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2020, pp. 1637–1646. <https://doi.org/10.1109/wacv45572.2020.9093314>
5. Afouras T., Chung, J.S., Senior A., Vinyals O., Zisserman A. Deep audio-visual speech recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, vol. 44, no. 12, pp. 8717–8727. <https://doi.org/10.1109/TPAMI.2018.2889052>
6. Kухарев G.A., Matveev Yu.N., Oleinik A.L. Mutual image transformation algorithms for visual information processing and retrieval. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2017, vol. 17, no. 1, pp. 62–74. (in Russian). <https://doi.org/10.17586/2226-1494-2017-17-1-62-74>
7. Shi B., Hsu W.N., Mohamed A. Robust self-supervised audio-visual speech recognition. *Proc. of the International Conference INTERSPEECH*, 2022, pp. 2118–2122. <https://doi.org/10.21437/interspeech.2022-99>
8. Chand H.V., Karthikeyan J. CNN based driver drowsiness detection system using emotion analysis. *Intelligent Automation & Soft Computing*, 2022, vol. 31, no. 2, pp. 717–728. <https://doi.org/10.32604/iasc.2022.020008>
9. Ivanko D., Kashevnik A., Ryumin D., Kitenko A., Axyonov A., Lashkov I., Karpov A. MIDriveSafely: Multimodal interaction for drive safely. *Proc. of the 2022 International Conference on Multimodal Interaction (ICMI)*, 2022, pp. 733–735. <https://doi.org/10.1145/3536221.3557037>
10. Biswas A., Sahu P.K., Chandra M. Multiple cameras audio visual speech recognition using active appearance model visual features in car environment. *International Journal of Speech Technology*, 2016, vol. 19, no. 1, pp. 159–171. <https://doi.org/10.1007/s10772-016-9332-x>
11. Nambi A.U., Bannur S., Mehta I., Kalra H., Virmani A., Padmanabhan V.N., Bhandari R., Raman B. HAMS: Driver and driving monitoring using a smartphone. *Proc. of the 24<sup>th</sup> Annual International Conference on Mobile Computing and Networking*, 2018, pp. 840–842. <https://doi.org/10.1145/3241539.3267723>
12. Kashevnik A., Lashkov I., Gurtov A. Methodology and mobile application for driver behavior analysis and accident prevention. *IEEE Transactions on Intelligent Transportation Systems*, 2020, vol. 21, no. 6, pp. 2427–2436. <https://doi.org/10.1109/TITS.2019.2918328>
13. Jang S.W., Ahn B. Implementation of detection system for drowsy driving prevention using image recognition and IoT. *Sustainability*, 2020, vol. 12, no. 7, pp. 3037. <https://doi.org/10.3390/su12073037>
14. Mishra R.K., Urolagin S., Jothi J.A.A., Gaur P. Deep hybrid learning for facial expression binary classifications and predictions. *Image and Vision Computing*, 2022, vol. 128, pp. 104573. <https://doi.org/10.1016/j.imavis.2022.104573>
15. Sunitha G., Geetha K., Neelakandan S., Pundir A.K.S., Hemalatha S., Kumar V. Intelligent deep learning based ethnicity recognition and classification using facial images. *Image and Vision Computing*, 2022, vol. 121, pp. 104404. <https://doi.org/10.1016/j.imavis.2022.104404>
16. Yuan Y., Tian C., Lu X. Auxiliary loss multimodal GRU model in audio-visual speech recognition // *IEEE Access*. 2018, V. 6. P. 5573–5583. <https://doi.org/10.1109/ACCESS.2018.2796118>

17. Hou J.C., Wang S.S., Lai Y.H., Tsao Y., Chang H.W., Wang H.M. Audio-visual speech enhancement using multimodal deep convolutional neural networks // *IEEE Transactions on Emerging Topics in Computational Intelligence*. 2018. V. 2. N 2. P. 117–128. <https://doi.org/10.1109/TETCI.2017.2784878>
18. Chan Z.M., Lau C.Y., Thang K.F. Visual speech recognition of lips images using convolutional neural network in VGG-M model // *Journal of Information Hiding and Multimedia Signal Processing*. 2020. V. 11. N 3. P. 116–125.
19. Zhu X., Cheng D., Zhang Z., Lin S., Dai J. An empirical study of spatial attention mechanisms in deep networks // *Proc. of the IEEE/CVF International Conference on Computer Vision*. 2019. P. 6688–6697. <https://doi.org/10.1109/iccv.2019.00679>
20. Bhaskar S., Thasleema T.M. LSTM model for visual speech recognition through facial expressions // *Multimedia Tools and Applications*. 2023. V. 82. N 4. P. 5455–5472. <https://doi.org/10.1007/s11042-022-12796-1>
21. Hori T., Cho J., Watanabe S. End-to-end Speech recognition with word-based RNN language models // *Proc. of the 2018 IEEE Spoken Language Technology Workshop (SLT)*. 2018. P. 389–396. <https://doi.org/10.1109/SLT.2018.8639693>
22. Serdyuk D.D., Braga O.P.F., Siohan O. Transformer-based video front-ends for audio-visual speech recognition for single and multi-person video // *Proc. of the INTERSPEECH*. 2022. P. 2833–2837. <https://doi.org/10.21437/interspeech.2022-10920>
23. Chen C.F.R., Fan Q., Panda R. CrossViT: Cross-attention multi-scale vision transformer for image classification // *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2021. P. 347–356. <https://doi.org/10.1109/iccv48922.2021.00041>
24. Pan S.J., Yang Q. A survey on transfer learning // *IEEE Transactions on Knowledge and Data Engineering*. 2010. V. 22. N 10. P. 1345–1359. <https://doi.org/10.1109/tkde.2009.191>
25. Романенко А.Н., Матвеев Ю.Н., Минкер В. Перенос знаний в задаче автоматического распознавания русской речи в телефонных переговорах // *Научно-технический вестник информационных технологий, механики и оптики*. 2018. Т. 18. № 2. С. 236–242. <https://doi.org/10.17586/2226-1494-2018-18-2-236-242>
26. Sui C., Bennamoun M., Togneri R. Listening with your eyes: towards a practical visual speech recognition system using deep boltzmann machines // *Proc. of the IEEE International Conference on Computer Vision (ICCV)*. 2015. P. 154–162. <https://doi.org/10.1109/iccv.2015.26>
27. Ahmed N., Natarajan T., Rao K.R. Discrete cosine transform // *IEEE Transactions on Computers*. 1974. V. C-23. N 1. P. 90–93. <https://doi.org/10.1109/T-C.1974.223784>
28. Xanthopoulos P., Pardalos P.M., Trafalis T.B. Linear discriminant analysis // *Robust Data Mining*. Springer New York, 2013. P. 27–33. [https://doi.org/10.1007/978-1-4419-9878-1\\_4](https://doi.org/10.1007/978-1-4419-9878-1_4)
29. Томашенко Н.А., Хохлов Ю.Ю., Ларшер Э., Эстев Я., Матвеев Ю.Н. Использование в системах автоматического распознавания речи GMM-моделей для адаптации акустических моделей, построенных на основе искусственных нейронных сетей // *Научно-технический вестник информационных технологий, механики и оптики*. 2016. Т. 16. № 6. С. 1063–1072. <https://doi.org/10.17586/2226-1494-2016-16-6-1063-1072>
30. Ma P., Petridis S., Pantic M. End-to-end audio-visual speech recognition with conformers // *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2021. P. 7613–7617. <https://doi.org/10.1109/ICASSP39728.2021.9414567>
31. Ryumin D., Ivanko D., Ryumina E. Audio-visual speech and gesture recognition by sensors of mobile devices // *Sensors*. 2023. V. 23. N 4. P. 2284. <https://doi.org/10.3390/s23042284>
32. Huang J., Kingsbury B. Audio-visual deep learning for noise robust speech recognition // *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*. 2013. P. 7596–7599. <https://doi.org/10.1109/ICASSP.2013.6639140>
33. Ivanko D., Ryumin D., Kashevnik A., Axyonov A., Kitenko A., Lashkov I., Karpov A. DAVIS: Driver's audio-visual speech recognition // *Proc. of the International Conference INTERSPEECH*. 2022. P. 1141–1142.
34. Zhou P., Yang W., Chen W., Wang Y., Jia J. Modality attention for end-to-end audio-visual speech recognition // *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2019. P. 6565–6569. <https://doi.org/10.1109/ICASSP.2019.8683733>
16. Yuan Y., Tian C., Lu X. Auxiliary loss multimodal GRU model in audio-visual speech recognition. *IEEE Access*, 2018, vol. 6, pp. 5573–5583. <https://doi.org/10.1109/ACCESS.2018.2796118>
17. Hou J.C., Wang S.S., Lai Y.H., Tsao Y., Chang H.W., Wang H.M. Audio-visual speech enhancement using multimodal deep convolutional neural networks. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2018, vol. 2, no. 2, pp. 117–128. <https://doi.org/10.1109/TETCI.2017.2784878>
18. Chan Z.M., Lau C.Y., Thang K.F. Visual speech recognition of lips images using convolutional neural network in VGG-M model. *Journal of Information Hiding and Multimedia Signal Processing*, 2020, vol. 11, no. 3, pp. 116–125.
19. Zhu X., Cheng D., Zhang Z., Lin S., Dai J. An empirical study of spatial attention mechanisms in deep networks. *Proc. of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 6688–6697. <https://doi.org/10.1109/iccv.2019.00679>
20. Bhaskar S., Thasleema T.M. LSTM model for visual speech recognition through facial expressions. *Multimedia Tools and Applications*, 2023, vol. 82, no. 4, pp. 5455–5472. <https://doi.org/10.1007/s11042-022-12796-1>
21. Hori T., Cho J., Watanabe S. End-to-end Speech recognition with word-based RNN language models. *Proc. of the 2018 IEEE Spoken Language Technology Workshop (SLT)*, 2018, pp. 389–396. <https://doi.org/10.1109/SLT.2018.8639693>
22. Serdyuk D.D., Braga O.P.F., Siohan O. Transformer-based video front-ends for audio-visual speech recognition for single and multi-person video. *Proc. of the INTERSPEECH*, 2022, pp. 2833–2837. <https://doi.org/10.21437/interspeech.2022-10920>
23. Chen C.F.R., Fan Q., Panda R. CrossViT: Cross-attention multi-scale vision transformer for image classification. *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 347–356. <https://doi.org/10.1109/iccv48922.2021.00041>
24. Pan S.J., Yang Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, 2010, vol. 22, no. 10, pp. 1345–1359. <https://doi.org/10.1109/tkde.2009.191>
25. Romanenko A.N., Matveev Yu.N., Minker W. Knowledge transfer for Russian conversational telephone automatic speech recognition. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2018, vol. 18, no. 2, pp. 236–242 (in Russian). <https://doi.org/10.17586/2226-1494-2018-18-2-236-242>
26. Sui C., Bennamoun M., Togneri R. Listening with your eyes: towards a practical visual speech recognition system using deep boltzmann machines. *Proc. of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 154–162. <https://doi.org/10.1109/iccv.2015.26>
27. Ahmed N., Natarajan T., Rao K.R. Discrete cosine transform. *IEEE Transactions on Computers*, 1974, vol. C-23, no. 1, pp. 90–93. <https://doi.org/10.1109/T-C.1974.223784>
28. Xanthopoulos P., Pardalos P.M., Trafalis T.B. Linear discriminant analysis. *Robust Data Mining*, Springer New York, 2013, pp. 27–33. [https://doi.org/10.1007/978-1-4419-9878-1\\_4](https://doi.org/10.1007/978-1-4419-9878-1_4)
29. Tomashenko N.A., Khokhlov Yu. Yu., Larcher A., Estève Ya., Matveev Yu.N. Gaussian mixture models for adaptation of deep neural network acoustic models in automatic speech recognition systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2016, vol. 16, no. 6, pp. 1063–1072. (in Russian). <https://doi.org/10.17586/2226-1494-2016-16-6-1063-1072>
30. Ma P., Petridis S., Pantic M. End-to-end audio-visual speech recognition with conformers. *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 7613–7617. <https://doi.org/10.1109/ICASSP39728.2021.9414567>
31. Ryumin D., Ivanko D., Ryumina E. Audio-visual speech and gesture recognition by sensors of mobile devices. *Sensors*, 2023, vol. 23, no. 4, pp. 2284. <https://doi.org/10.3390/s23042284>
32. Huang J., Kingsbury B. Audio-visual deep learning for noise robust speech recognition. *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 7596–7599. <https://doi.org/10.1109/ICASSP.2013.6639140>
33. Ivanko D., Ryumin D., Kashevnik A., Axyonov A., Kitenko A., Lashkov I., Karpov A. DAVIS: Driver's audio-visual speech recognition. *Proc. of the International Conference INTERSPEECH*, 2022, pp. 1141–1142.
34. Zhou P., Yang W., Chen W., Wang Y., Jia J. Modality attention for end-to-end audio-visual speech recognition. *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*

35. Ivanko D., Axyonov A., Ryumin D., Kashevnik A., Karpov A. RUSAVIC Corpus: Russian audio-visual speech in cars // Proc. of the 13<sup>th</sup> Language Resources and Evaluation Conference (LREC). 2022. P. 1555–1559.
36. Kashevnik A., Lashkov I., Axyonov A., Ivanko D., Ryumin D., Kolchin A., Karpov A. Multimodal corpus design for audio-visual speech recognition in vehicle cabin // IEEE Access. 2021. V. 9. P. 34986–35003. <https://doi.org/10.1109/ACCESS.2021.3062752>
37. Lugaresi C., Tang J., Nash H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.-L., Yong M., Lee J., Chang W.-T., Hua W., Georg M., Grundmann M. MediaPipe: A framework for perceiving and processing reality // Proc. of the 3<sup>rd</sup> Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR). 2019. V. 2019. P. 1–4.
38. Zhang H., Cisse M., Dauphin Y.N., Lopez-Paz D. MixUp: Beyond empirical risk minimization // Proc. of the ICLR Conference. 2018. P. 1–13.
39. Feng D., Yang S., Shan S. An efficient software for building LIP reading models without pains // Proc. of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW). 2021. P. 1–2. <https://doi.org/10.1109/ICMEW53276.2021.9456014>
40. Kim M., Hong J., Park S.J., Ro Y.M. Multi-modality associative bridging through memory: speech sound recollected from face video // Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV). 2021. P. 296–306. <https://doi.org/10.1109/iccv48922.2021.00036>
41. Zhong Z., Lin Z.Q., Bidart R., Hu X., Daya I.B., Li Z., Zheng W., Li J., Wong A. Squeeze-and-attention networks for semantic segmentation // Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. P. 13065–13074. <https://doi.org/10.1109/cvpr42600.2020.01308>
42. Аксёнов А.А., Рюмин Д.А., Кашевник А.М., Иванько Д.В., Карпов А.А. Метод визуального анализа лица водителя для автоматического чтения речи по губам при управлении транспортным средством // Компьютерная оптика. 2022. Т. 46. № 6. С. 955–962. <https://doi.org/10.18287/2412-6179-CO-1092>
- (ICASSP), 2019, pp. 6565–6569. <https://doi.org/10.1109/ICASSP.2019.8683733>
35. Ivanko D., Axyonov A., Ryumin D., Kashevnik A., Karpov A. RUSAVIC Corpus: Russian audio-visual speech in cars. Proc. of the 13<sup>th</sup> Language Resources and Evaluation Conference (LREC), 2022, pp. 1555–1559.
36. Kashevnik A., Lashkov I., Axyonov A., Ivanko D., Ryumin D., Kolchin A., Karpov A. Multimodal corpus design for audio-visual speech recognition in vehicle cabin. IEEE Access, 2021, vol. 9, pp. 34986–35003. <https://doi.org/10.1109/ACCESS.2021.3062752>
37. Lugaresi C., Tang J., Nash H., McClanahan C., Uboweja E., Hays M., Zhang F., Chang C.-L., Yong M., Lee J., Chang W.-T., Hua W., Georg M., Grundmann M. MediaPipe: A framework for perceiving and processing reality. Proc. of the 3<sup>rd</sup> Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR), 2019, vol. 2019, pp. 1–4.
38. Zhang H., Cisse M., Dauphin Y.N., Lopez-Paz D. MixUp: Beyond empirical risk minimization. Proc. of the ICLR Conference, 2018, pp. 1–13.
39. Feng D., Yang S., Shan S. An efficient software for building LIP reading models without pains. Proc. of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW), 2021, pp. 1–2. <https://doi.org/10.1109/ICMEW53276.2021.9456014>
40. Kim M., Hong J., Park S.J., Ro Y.M. Multi-modality associative bridging through memory: speech sound recollected from face video. Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 296–306. <https://doi.org/10.1109/iccv48922.2021.00036>
41. Zhong Z., Lin Z.Q., Bidart R., Hu X., Daya I.B., Li Z., Zheng W., Li J., Wong A. Squeeze-and-attention networks for semantic segmentation. Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 13065–13074. <https://doi.org/10.1109/cvpr42600.2020.01308>
42. Axyonov A.A., Ryumin D.A., Kashevnik A.M., Ivanko D.V., Karpov A.A. Method for visual analysis of driver's face for automatic lip-reading in the wild. Computer Optic, 2022, vol. 46, no. 6, pp. 955–962. (in Russian). <https://doi.org/10.18287/2412-6179-CO-1092>

#### Авторы

**Аксёнов Александр Александрович** — младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57203963345](https://orcid.org/0000-0002-7479-2851), <https://orcid.org/0000-0002-7479-2851>, [axyonov.a@iias.spb.su](mailto:axyonov.a@iias.spb.su)

**Рюмина Елена Витальевна** — младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57220572427](https://orcid.org/0000-0002-4135-6949), <https://orcid.org/0000-0002-4135-6949>, [ryumina.e@iias.spb.su](mailto:ryumina.e@iias.spb.su)

**Рюмин Дмитрий Александрович** — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57191960214](https://orcid.org/0000-0002-7935-0569), <https://orcid.org/0000-0002-7935-0569>, [dl\\_03.03.1991@mail.ru](mailto:dl_03.03.1991@mail.ru)

**Иванько Денис Викторович** — кандидат технических наук, старший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57190967993](https://orcid.org/0000-0003-0412-7765), <https://orcid.org/0000-0003-0412-7765>, [ivanko.d@iias.spb.su](mailto:ivanko.d@iias.spb.su)

**Карпов Алексей Анатольевич** — доктор технических наук, профессор, заведующий лабораторией, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57219469958](https://orcid.org/0000-0003-3424-652X), <https://orcid.org/0000-0003-3424-652X>, [karpov@iias.spb.su](mailto:karpov@iias.spb.su)

Статья поступила в редакцию 12.04.2023  
Одобрена после рецензирования 17.05.2023  
Принята к печати 24.07.2023

#### Authors

**Alexandr A. Axyonov** — Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation, [sc 57203963345](https://orcid.org/0000-0002-7479-2851), <https://orcid.org/0000-0002-7479-2851>, [axyonov.a@iias.spb.su](mailto:axyonov.a@iias.spb.su)

**Elena V. Ryumina** — Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation, [sc 57220572427](https://orcid.org/0000-0002-4135-6949), <https://orcid.org/0000-0002-4135-6949>, [ryumina.e@iias.spb.su](mailto:ryumina.e@iias.spb.su)

**Dmitry A. Ryumin** — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation, [sc 57191960214](https://orcid.org/0000-0002-7935-0569), <https://orcid.org/0000-0002-7935-0569>, [dl\\_03.03.1991@mail.ru](mailto:dl_03.03.1991@mail.ru)

**Denis V. Ivanko** — PhD, Senior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation, [sc 57190967993](https://orcid.org/0000-0003-0412-7765), <https://orcid.org/0000-0003-0412-7765>, [ivanko.d@iias.spb.su](mailto:ivanko.d@iias.spb.su)

**Alexey A. Karpov** — D.Sc., Professor, Head of Laboratory, St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), Saint Petersburg, 199178, Russian Federation, [sc 57219469958](https://orcid.org/0000-0003-3424-652X), <https://orcid.org/0000-0003-3424-652X>, [karpov@iias.spb.su](mailto:karpov@iias.spb.su)

Received 12.04.2023  
Approved after reviewing 17.05.2023  
Accepted 24.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-776-785

## Brain tumour segmentation in MRI using fuzzy deformable fusion model with Dolphin-SCA

Anjali Hemant Tiple<sup>1</sup>✉, Anandrao B. Kakade<sup>2</sup>

<sup>1</sup> Shivaji University, Kolhapur, Maharashtra, 416004, India

<sup>2</sup> Rajarambapu Institute of Technology, Uran Islampur, Maharashtra, 415414, India

<sup>1</sup> [anjalitiple6@gmail.com](mailto:anjalitiple6@gmail.com)✉, <https://orcid.org/0000-0003-4411-3878>

<sup>2</sup> [abkakade22@gmail.com](mailto:abkakade22@gmail.com), <https://orcid.org/0000-0002-1882-8901>

### Abstract

It is evident that when the human brain stops functioning for a small period of time, it will lead to death. As a result, dealing with brain disorders should be done early and properly. A brain tumour is one of the most serious brain illnesses. The development of tumours can be detected using Magnetic Resonance Imaging (MRI). However, because an MRI image has loud noise, it can be hard to diagnose a tumour. The diagnosis process is slow, yet illness necessitates prompt and accurate medical attention in order for patients to survive. One of the solutions for tumour diagnosis is to employ MRI brain picture segmentation. In this designed model, MRI of the brain is collected and pre-processed with Non-Local Means (NLM) to reduce noise from captured raw data. This pre-processed image is first segmented with Region of Interest (ROI) for identifying regions of interest and then with a fusion deformable fuzzy system, which combines fuzzy C-means (FCM) and deformable systems. By analyzing the fitness value of  $\alpha$  and  $\beta$  constants, segmented pictures from models are fused using the Dolphin Sine Cosine Algorithm (SCA) method to combine the model results. The integrated output from the algorithm is classified with the deep Convolutional Neural Network (CNN) classifier. The created model experimental findings are analyzed and compared to current methodologies. The proposed model performance measures are 0.90, 0.89, 0.88, and 0.10 in terms of selectivity, precision, accuracy and errors. As a result, when compared to previous strategies, the proposed approach outperforms them.

### Keywords

Dolphin-SCA, FCM, deformable model, ROI, NLM, tumour segmentation

**For citation:** Tiple A.H., Kakade A.B. Brain tumour segmentation in MRI using fuzzy deformable fusion model with Dolphin-SCA. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 776–785. doi: 10.17586/2226-1494-2023-23-4-776-785

УДК 004.89

## Сегментация опухоли головного мозга на магнитно-резонансной томографии с использованием нечеткого деформируемого слияния и алгоритма Dolphin-SCA

Анджали Хемант Типле<sup>1</sup>✉, Анандрао Б. Какаде<sup>2</sup>

<sup>1</sup> Университет Шиваджи, Колхатур, Махараштра, 416004, Индия

<sup>2</sup> Технологический институт Раджарамбапу, Уран Ислампур, Махараштра, 415414, Индия

<sup>1</sup> [anjalitiple6@gmail.com](mailto:anjalitiple6@gmail.com)✉, <https://orcid.org/0000-0003-4411-3878>

<sup>2</sup> [abkakade22@gmail.com](mailto:abkakade22@gmail.com), <https://orcid.org/0000-0002-1882-8901>

### Аннотация

Прекращение функционирования мозга человека на небольшой промежуток времени приводит к смерти. Лечение нарушений головного мозга должно проводиться на ранней стадии и до появления клинических симптомов. Опухоль головного мозга является одним из самых серьезных заболеваний. Развитие опухоли можно обнаружить с помощью магнитно-резонансной томографии (МРТ). В связи с наличием шумов на изображении МРТ опухоль сложно точно и быстро диагностировать. Одно из решений в диагностике опухолей — использование

© Tiple A.H., Kakade A.B., 2023

сегментации изображений головного мозга на МРТ. В работе представлена модель томограммы головного мозга обработанная с помощью нелокальных средств (Non-Local Means, NLM) для уменьшения шума от захваченных необработанных данных. Полученное изображение сегментировано с помощью определения областей интереса (ROI) и деформируемой нечеткой системы слияния. Система слияния сочела в себе метод нечеткой кластеризации С-средних (Fuzzy C-Means, FCM) и деформируемых систем. Выполнен анализ значений пригодности констант  $\alpha$  и  $\beta$  сегментированных изображений моделей, объединенных с использованием алгоритма синус-косинуса на основе эхолокации Dolphin-SCA. Интегрированный вывод алгоритма классифицирован с помощью глубокого классификатора (Convolutional Neural Network, CNN). Проведен анализ и сравнение экспериментальных данных созданной модели с текущими методологиями. Значения показателей эффективности предлагаемой модели для селективности, прецизионности, правильности и ошибок составили 0,90, 0,89, 0,88 и 0,10 соответственно. Таким образом, по сравнению с предыдущими стратегиями, предлагаемый подход превосходит ранее применяемые методы.

#### Ключевые слова

Dolphin-SCA, метод нечеткой кластеризации С-средних, FCM, деформируемая модель, определение областей интереса, ROI, нелокальные средства, NLM, сегментация опухоли

**Ссылка для цитирования:** Типле А.Х., Какаде А.Б. Сегментация опухоли головного мозга на магнитно-резонансной томографии с использованием нечеткого деформируемого слияния и алгоритма Dolphin-SCA // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 776–785 (на англ. яз.). doi: 10.17586/2226-1494-2023-23-4-776-785

## Introduction

Brain tumours develop as a result of aberrant cell division and uncontrolled growth in the brain. The formation of tissues is classified as benign or malicious, and it takes into various parts of the brain that aren't affected by tumours [1]. As a consequence, tumours must be detected as soon as possible to avoid their development into an uncontrollable state. The ability to accurately segment tumours using medical pictures is critical because it gives information that is needed for cancer assessment and diagnosis, and also planning out therapy choices and tracking disease progression [2]. Brain tumours are among the most lethal malignancies in the world, and they are divided into primary and secondary tumour types based on their origin [3]. Glioma, which arises from brain glial cells and accounts for 80 % of all malignant brain tumours, is the most frequent histological form of primary brain cancer [4]. Gliomas could be of two types: slow-progressing low-grade which has a good prognosis, or aggressive and infiltrative high-grade glioma or glioblastoma which requires prompt treatment.

The approach used to detect brain tumours is Magnetic Resonance Imaging (MRI). It is a complicated technique that provides information on the human soft tissue structure which is then expanded to examine the anatomy of the region and aids in the creation of elaborate representations in each direction [5]. MRI is a sort of medical imaging that may identify a variety of changes in soft tissues all over the body. Fluid Attenuated Inversion Recovery (FLAIR), T2-weighted, T1-weighted, and post-contrast T1-weighted (T1ce) are examples of complementing 3D MRI techniques [6]. The tumours highlighted sub-regions throughout the varying intensity of these patterns, like the entire tumour, are far more evident in T2 modalities and FLAIR [7]. T1 and T1ce scans, on either hand, reveal the tumour centre without any peritumoral oedema. It allows for the detection of various tumour sub-regions using such scans in conjunction with extra information they supply.

Nowadays, the clinical procedures used in the medical profession for detecting tumours utilizing MRI images of the brain are not very effective or stable [8]. Various

techniques for image segmentation have been developed, and some of the most commonly used methods include clustering-based methods, histogram-based methods, thresholding, region-growing methods, edge detection, and so on [9]. Clustering algorithms attempt to find patterns in a dataset in order to extract additional information. There are several clustering methods available today, including clustering based on the distribution, hierarchical clustering, k-median, k-means and many others [10]. The MRI is segmented across this proposed approach utilizing a deformable model and the Fuzzy C-Means (FCM) system. For properly segmenting brain tumours from MRI, the segmented image of two models is fused with the fusion rule on the basis of weighted variables represented as  $\alpha$  and  $\beta$ . The proposed model primary contributions are as follows:

- Development of a fusion-based image segmentation technique for brain tumour cells using MRI.
- To filter the noise from the raw data, the Non-Local Means (NLM) filter is used as a pre-processing technique.
- The Region of Interest (ROI) and Fuzzy deformable fusion models based on fitness evaluation of and with Dolphin — Sine Cosine Algorithm (SCA) are employed to segment pre-processed picture.
- Deep Convolutional Neural Network (CNN) was employed to forecast the performance of the proposed segmentation technique.

## Literature survey

Many studies used a variety of ways to segment a brain tumour using MRI images. Most of the existing techniques are studied, and among those techniques a few of them are reviewed below.

Sheela C.J.J. & Suganthi G. [11] had performed MRI-based segmentation of brain tumours automatically. The approximated ROI is established first in every developed model by removing the non-tumour portion of the picture by utilizing two-level morphological reconstruction methods like dilation and erosion. A mask is constructed by thresholding the reconstructed image but then degraded

to improve segmentation accuracy in the Greedy Snake technique. By employing mask boundary as the snake initial contour, the greedy snake method computes new tumour borders. These limits are more accurate in areas with sharp edges and less accurate in areas with ramp edges. Lastly, a region with the biggest circumference is picked to eliminate areas that were incorrectly divided.

Hrosik R.C. et al. [12] had developed k-means clustering, and the firefly method was used to segment brain images. Metastatic adenocarcinoma, sarcoma, glioma, and metastatic bronchogenic carcinoma are four distinct primary brain tumours that can be detected using MRI, Single-Photon Emission Computed Tomography, Positron Emission Tomography pictures, and methods for brain image segmentation have been formed with the aim of showcasing them. The created picture segmentation technique depends on the firefly algorithm which results in enhanced by the k-means clustering method, and the fitness function is Otsu's criterion. This method enhances segmentation when traditional segmentation quality measurements like normalized root square mean error, peak signal to noise, and structural similarity index is utilized.

Huang H. et al. [13] had performed segmentation of brain images using the FCM clustering algorithm. The image is first segmented into several tiny portions depending on discernible attribute relationships, and then an attribute value table is built based on FCM segmentation findings for various clustering numbers. The weights of every characteristic are obtained by the decreased amount and user base for calculating difference across regions, following which similarity of every region is determined utilizing equivalence link supplied by difference degree. Finally, regions are merged, and picture segmentation is completed using the final equivalence relation provided by similarity. This technique is used to confirm the segmentation of artificially induced images, MRI images and brain Computerised Tomography scans.

Srinivas K. & Kantapalli B. [14] had designed segmentation of MRI using Unified Iterative Partitioned Fuzzy Clustering (U-IPFC). In medical sector applications, detecting tissues from the MRI brain is a difficult problem. To effectively identify the tissues from the MRI, the segmentation technique is used. This model is comprised of a pre-processing method for detecting several tissues in MR brain images and calculating tissue area. The efficiency of the designed U-IPFC technique has been demonstrated by the extensive simulated study. The major objective is to enhance the accuracy of existing FCM and

k-means clustering methods for the identification of multi-tissues.

Ali M. et al. [15] had performed Deep Neural Networks for brain tumour picture segmentation. In this paper, 3D CNN and U-Net are used to create an aggregation of two segmentation networks, which results in more accurate and effective forecasts. These two classifiers are developed separately on BraTS-19 benchmark datasets and examined to produce segmentation maps which were considerably diverse in terms of segmented tumour sub-regions and overall variance in order to get the final forecast. For augmenting tumour, total tumour, and tumour core, the developed ensemble got maximum dice values.

### Proposed methodology

Brain tumours can affect in a variety of sizes and forms as well as varied picture intensity. Currently, clinical approaches utilized in the medical field are not very accurate or robust in identifying tumours using MRI scans of the brain. As a result, the proposed method is meant for precise brain tumour segmentation and identification.

The proposed model process flow is depicted in Fig. 1. In this designed model, an MRI of the brain is collected and pre-processed with NLM reduce noise from captured raw data. This pre-processed image is first segmented with ROI for segmenting interested regions and then again segmented with a fusion deformable fuzzy model that consists of an FCM system and deformable system. Through measuring the fitness value of  $\alpha$  and  $\beta$  constants, segmented images from classifiers are merged using the Dolphin-SCA method to combine the model results. The integrated output from the algorithm is classified with a deep CNN classifier.

**Pre-processing.** Collection of data is the process of gathering the necessary data, such as text or images, to be predicted. These collected data consist of raw data that cannot be segmented or classified with high accuracy. Hence, specific pre-processing procedures are employed to achieve high accuracy.

**NLM.** A weighted filter is a linear coefficient of the resemblance image patches in an image [16] and it is the foundation of NLM denoising method. The pixel indexes are  $i$  and  $j$ , and the following equation describes NLM:

$$y(i) = \sum w(i, j)v(j), j \in s_i, \quad (1)$$

where  $v(j)$  represents the centered position at  $j$ ;  $w(i, j)$  is weight function among image patches around  $i$  and  $j$

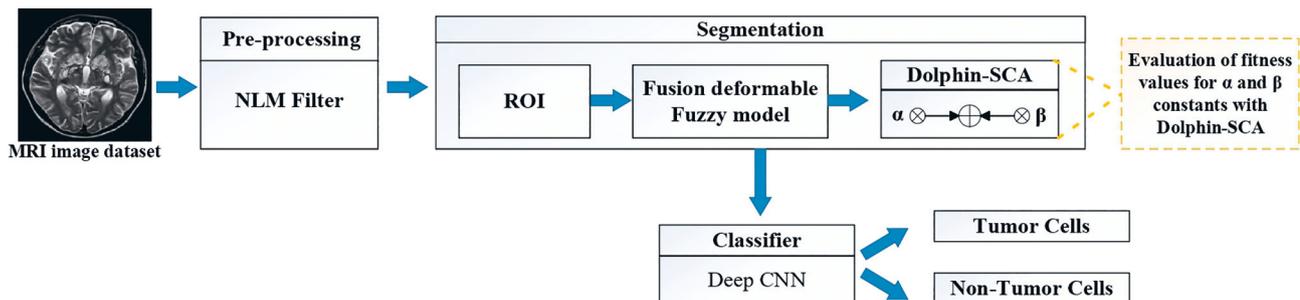


Fig. 1. The proposed model process flow

described by  $y(i)$ ; and  $j$  is indeed searching window  $s \times s$  of about  $i$ . NLM denoising is an excellent image denoising technique that reduces the computational complexity by using the accelerating mechanisms. The majority of representational approaches have implemented dimensionality reduction, although this procedure comes at a high expense of output operations.

**Segmentation.** Segmentation is mainly useful for eliminating the unwanted portions of the image to reduce the computational time and wrong predictions. In this designed model, the Grab cut algorithm is used for segmentation.

**ROI align.** ROI is used to identify the particular portion of a 2D/3D/4D image. The pixels are classified under two categories,  $X$  and  $Y$  (objects and backgrounds, or conversely) when the region is identified [17]. Pixels at levels  $[1, \dots, K]$  make up  $X$ , while pixels at levels  $[K + 1, \dots, L]$  make up  $Y$ . The variance between the two classes is expressed in equation

$$\sigma^2(k) = \frac{[\lambda_T w(k) - \lambda(k)]^2}{w(k)[1 - w(k)]}, \quad (2)$$

where  $\sigma(k)$  is the variance of threshold level,  $w(k)$  is the weight of the threshold level,  $\lambda(k) = \sum_{i=1}^k ip_i$ , and  $\lambda_T(k) = \sum_{i=1}^k ip_i$ .  $\lambda_T$  is the arithmetic mean for the entire image and  $\lambda_T = w_X \lambda_X + w_Y \lambda_Y$ . The maximum variance value will be considered as the ROI value. The feature aligned from the ROI is given to the fully connected layer and convolution layer.

**Fusion deformable Fuzzy model.** In this model, the segmentation of the MRI is done using the deformable and FCM model. The models result is combined with Dolphin-SCA algorithm-generated optimal constants. Give FCM and deformable method ROI-segmented images. The deformable model outputs and the FCM resulting images are merged with constants acquired from the Dolphin-SCA method leading to changes in the outputs of both pictures. The output of fuzzy deformable fusion technique for brain image  $O$  following segmentation is denoted by

$$O = \alpha D_m + \beta F_m,$$

where  $D_m$  is deformable model outputs after segmentation and  $F_m$  is the FCM model segmented outputs. Furthermore, their constants  $\alpha$  and  $\beta$  indicate the Dolphin-SCA method optimal constants.

**Deformable model.** In image segmentation, active contours are the most extensively employed deformable model [18]. The snake is an active contour method which is constrained by extraneous factors and guided by image pressures that tug it toward specific characteristics like edges and lines. A snake system is described as a direct parametric curve with the following formula:

$$n(s) = [a(s), b(s)], s \in [0, 1], \quad (3)$$

where  $s$  is normalized arc lengths within range of  $0 < s < 1$  and  $n(s)$  is a set of nodes like a snake.

**FCM model.** For such a clustering procedure, FCM creates fuzzy matrices by calculating Euclidean distance

[19]. Clustering yields a matrix depending on FCM which is expressed as

$$M = \sum_{c=1}^i \sum_{d=1}^j i_{cd} r_{cd}, 1 \leq \infty, \quad (4)$$

where  $r_{cd}$  signifies Euclidean distance,  $i_{cd}$  means membership function matrix,  $c$  signifies pixel location, and  $d$  is the overall number of pixels representing fuzziness parameter.

**Determination of segmentation constants using Dolphin-SCA algorithm.** The optimal constants are obtained utilizing the output of deformable and FCM methods [20] and the Dolphin-SCA method. The characteristics of the Dolphin Echolocation (DE) method and Sine Cosine Algorithm (SCA) are combined in this algorithm. The SCA technique uses features of sine and cosine functions to discover ideal outputs, whereas the DE method is inspired by Dolphin behaviour.

**Initialization.** The Dolphin-SCA first sets up solution space in order to get the best values for constants. A response area is provided by Dolphin-SCA which is described as

$$Y = \{Y_1, Y_2, \dots, Y_a, \dots, Y_b\},$$

where  $Y_a$  denotes  $a^{\text{th}}$  option of dimensions  $1 \times 2$ , and  $b$  denotes overall possibilities.

1) *Evaluation of the fitness.*

To determine the best solution, we calculate the fitness of each unique solution. The fitness is calculated using the segmented image centre and pixel covering the segmented image centre pixel. The ideal option is one that provides the least amount of fitness value. The Dolphin-fitness SCA algorithm for calculating segmentation constants is as follows:

$$fitness = \sum_{L=1}^j \sum_{\substack{F=1 \\ F \in k}}^T (R^k - M_F).$$

Parameter  $T$  denotes the diameter of the brain picture, while  $R^k$  specifies the image segment centre. The  $F^{\text{th}}$  pixel in the  $k^{\text{th}}$  segment is represented by  $P^k$  and  $M_F$ . Assume the  $k^{\text{th}}$  segment of the image that has  $T$  pixels and is indicated by  $M_F$  for every pixel. The segmented image centre is calculated as

$$R^k = \frac{1}{T} \sum_{F=1}^T M_F \in P^k.$$

2) *Update solution using Dolphin-SCA*

SCA provides two services depending on sine and cosine update formulas accordingly. The equations are updated by the Dolphin-SCA algorithm utilizing response updates calculated by the DE algorithm. For sine and cosine functions, the updated answer are:

$$X(y+1) = \frac{1 - z_1 - z_2}{1 - z_1 - z_2 - 1 + \beta_1 \sin(\beta_2)} \times \left[ \frac{\beta_1 \sin(\beta_2)}{1 - z_1 - z_2} [u(y) + z_1 K + z_2 M] - \frac{z_1 K}{1 - z_1 - z_2} - \right]$$

$$-\frac{z_2 M}{1-z_1-z_2} + \beta_1 \beta_3 \sin(\beta_2) M \Big]; \beta_4 \leq 0.5.$$

$$X(y+1) = \frac{1-z_1-z_2}{1-z_1-z_2-1+\beta_1 \cos(\beta_2)} \times$$

$$\times \left[ \frac{\beta_1 \cos(\beta_2)}{1-z_1-z_2} [u(y) + z_1 K + z_2 M] - \frac{z_1 K}{1-z_1-z_2} - \right.$$

$$\left. - \frac{z_2 M}{1-z_1-z_2} + \beta_1 \beta_3 \cos(\beta_2) M \Big]; \beta_4 > 0.5.$$

The update solution for the Dolphin-SCA algorithm is provided by two reactions within the preceding equation.

### 3) Finding the best solution using fitness.

To determine whether a new approach is feasible, the fitness function is applied to update the solution. The best solution for delivering less fitness value is to provide fewer fitness values.

### 4) Termination.

The modified answer is replicated till  $Y_{\max}$ , the highest number of iterations, is reached. After multiple repetitions, the best segmentation number is discovered.

**Classification.** Classification is a method for predicting the outcome of supplied data sets. Classes are described using terminology like targets, labels, and categories. To categorize tumour and non-tumour cells, a deep CNN approach is designed.

**Deep CNN.** Finally, the Deep CNN classifier was used as a training example and the feature vectors are built using obtained characteristics. Deep CNN is utilized to distinguish between tumour and non-tumour regions. It is made up of three layers, such as convolutional layer, pooling layer, and fully connected layer. The greatest significant layer is a convolutional layer which became liable for intensive processing and hoisting control. In addition, this layer is in charge of collecting numerous features from input photos. The pooling layer, which is sandwiched among convolutional layers, is the following layer. The feature maps are subsampled by pooling layers and lower the spatial scale of depiction to decrease variables and network calculation. The fully connected layer manages the established links by previous layers and serves as a classifier for categorizing inputs. Consider the following source to Deep CNN:

$$(B_{\delta}^l)_{u,y} = (E_{\delta}^l)_u + \sum_{\beta_1=1}^{v_1^{\beta_1-1}} \sum_{\beta_1=v_1}^{v_1} \sum_{\beta_1=v_2}^{v_1} (K_{\delta,\beta_1}^l)_{\beta_3,\beta_2} * (B_{\beta_1}^{l-1})_{u+\beta_2,y+\beta_3},$$

where  $(B_{\delta}^l)_u$  symbolizes created output from  $l^{\text{th}}$  conventional layer centred as  $(u, y)$ ,  $K_{\delta,\beta_1}^l$  signifies convolutional operators for producing local designs from conventional layers, symbol \* indicates convolutional operators for producing local features from conventional layers,  $\beta_1$ ,  $\beta_2$  and  $\beta_3$  describe feature maps produced from every conventional filter via reading its input image, and  $E_{\delta}^l$  represents a bias of  $l^{\text{th}}$  conventional layer. An activation function is employed to streamline calculations or to eliminate undesirable results. The activation function of  $(l-1)^{\text{th}}$  layer generates outputs from  $p^{\text{th}}$  conventional layer which is written as

$$B_{\delta}^l = AFN(B_{\delta}^{l-1}).$$

The pool layers were modified to perform a defined function, and their characteristics were passed onto fully linked layers. The image is converted to a vector which is then classified. Fully connected layers yield the following output:

$$G_{\delta}^l = \eta(B_{\delta}^l) \text{With}(B_{\delta}^l) =$$

$$= \sum_{\beta_1=1}^{v_1^{\beta_1-1}} \sum_{\beta_1=v_1}^{v_1} \sum_{\beta_1=v_2}^{v_1} (K_{\delta,\beta_1}^l)_{\beta_3,\beta_2} * (B_{\beta_1}^{l-1})_{u+\beta_2,y+\beta_3}, \quad (5)$$

where  $h$  denotes the normalization factor.

Thus, the fused images from the fusion fuzzy deformable model are classified using the deep CNN model.

**Algorithm 1:** pseudocode for classification and segmentation of brain tumours

**Input:** Dataset of MRI ( $D=x_1, x_2, x_3 \dots x_n$ )

**#Pre-processing**

NLM = Non-linear filter( $D$ ) // by using equation (1)

ROI = Region of interest(NLM) // by using equation (2)

**#Segmentation**

DM = Deformable model(ROI) // by using equation (3)

FM = FCM(ROI) // by using equation (4)

**#Fusion with Dolphin-SCA algorithm**

If  $(\alpha + \beta = 1)$

O = DM + FM

Else

Search for other optimal values for fusing

End if

**#classification**

DCNN = Deep CNN(O) // by using equation (5)

If (class = 0)

Output = tumour cell

Else

Output = non-tumour cell

End if

**Output:** classified output of brain tumour cell

## Result and discussion

The proposed approach for brain tumour segmentation is implemented using an MRI brain imaging dataset and MATLAB software to test its functionality and performance. The 8.0 GB Memory (RAM), CPU @ 2.80 GHz, Intel(R) Core(TM) i5-3450S processor, and system type of 64-bit operating system are used in testing. The tumour is divided and categorized using MRI brain pictures in this proposed system. To remove noise in collected original data, an MRI of the brain is gathered and pre-processed utilizing NLM. This pre-processed

Sample No	Original	Pre-processing	Segmentation			Fusion
		NLM	ROI	Deformable model	FCM model	Deformable +FCM
1						
2						
3						
4						
5						

Note: Box (Yellow Color) in ROI image is the selected region of interest for the segmentation

Fig. 2. Conversion of sample images using pre-processing and Segmenting techniques

image is first segmented with ROI to isolate areas of interest and then with a fusion deformable fuzzy system, which combines FCM and deformable methods. Through analyzing the fitness value of  $\alpha$  and  $\beta$  constants, segmented images from models are merged using the dolphin SCA method to incorporate the output of the model. The integrated output from the algorithm is classified with a deep CNN classifier.

**Dataset.** In this dataset [21], the Glioma type of brain tumour is collected, which can affect your brain function and life-threatening depending on its location and rate of growth. Gliomas are one of the most common types of primary brain tumours. In this dataset, 224 pictures are collected for the categorization of tumour and non-tumour cells. In this case, 80 % of the data is utilized for learning, and 20 % is employed for evaluation.

Fig. 2 illustrates the conversion of sample images using pre-processing, segmentation and fusion techniques. Based on this converted images, the DCNN classifier is trained and tested for evaluating the performance of the model. Dolphin-SCA + FNB, PFCM + KNN, Deformable model + FNB, FCM + SVM are previous methodologies that are compared with the proposed technique using performance metrics like F1-score, Specificity, kappa, Sensitivity, Matthews Correlation Coefficient (MCC), Accuracy, False Positive Rate (FPR), False Negative Rate (FNR), Precision and Error.

The proposed method confusion matrix (error matrix) is depicted in Fig. 3. This error matrix is used to evaluate the actual and predicted data from the given dataset. The

prediction of data for ‘no’ and ‘yes’ classes are 64 and 93. The total data used for testing is 179. 157 of them are predicted according to the actual class, and the rest of the 22 are predicted wrongly.

Fig. 4 explains the comparison performance metrics for proposed and existing methods. Performance of the proposed model, such as Accuracy, Sensitivity, Specificity, Precision, Error, Kappa, FPR, FNR, F1-score and MCC, are 0.90, 0.89, 0.88, 0.89, 0.10, 0.72, 0.10, 0.12, 0.85 and 0.73. These evaluated proposed model values are greater than the existing techniques. Thus, the designed Dolphin-SCA+DCNN model accurately segments and predicts the brain tumour.

True Class	no	64	10
	yes	12	93
		no	yes
		Predicted Class	

Fig. 3. Error matrix for the proposed model

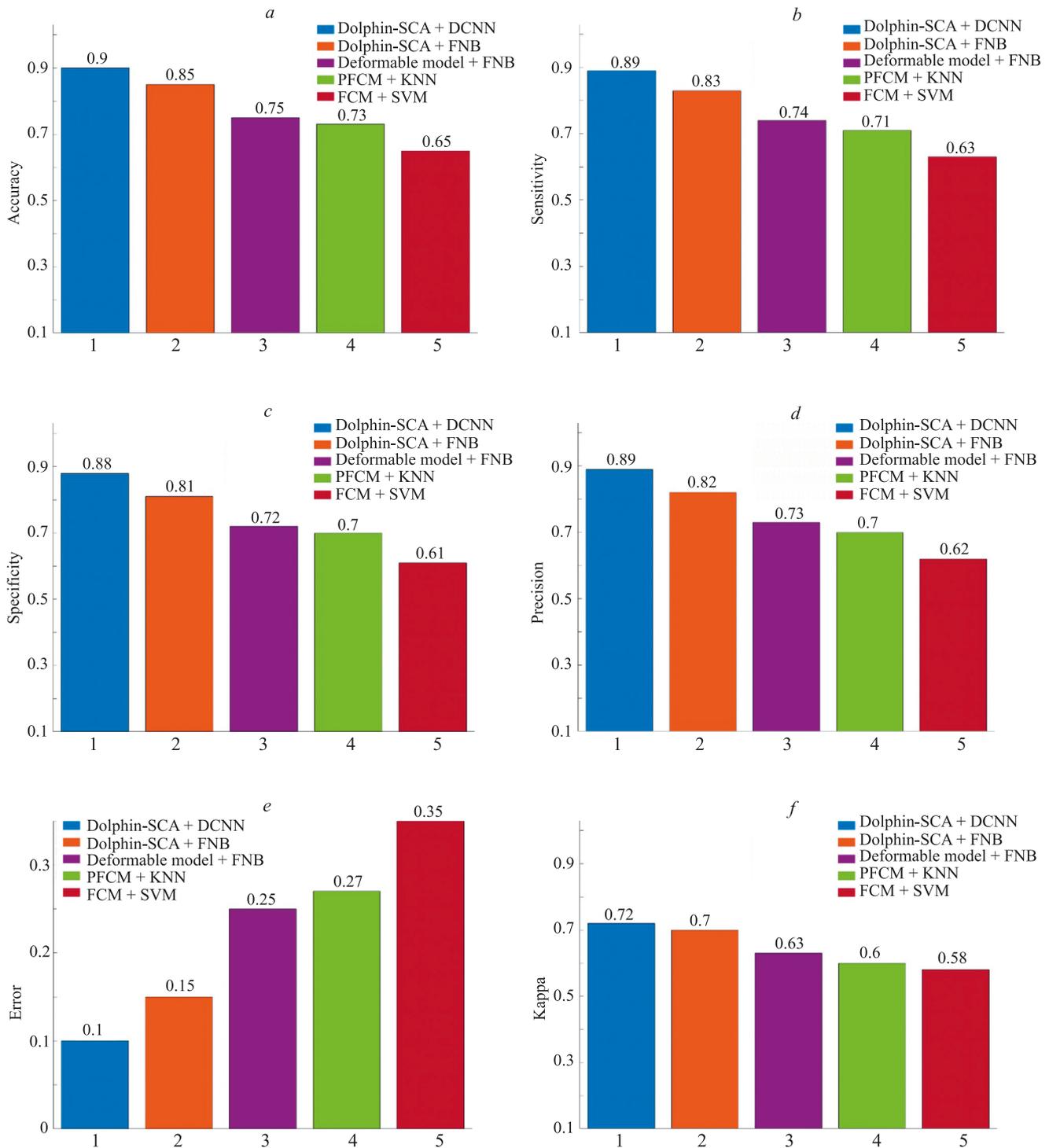


Fig. 4. Performance metrics: Accuracy (a); Sensitivity (b); Specificity (c); Precision (d); Error (e); Kappa (f); FPR (g); FNR (h); F1-score (i); MCC (j)

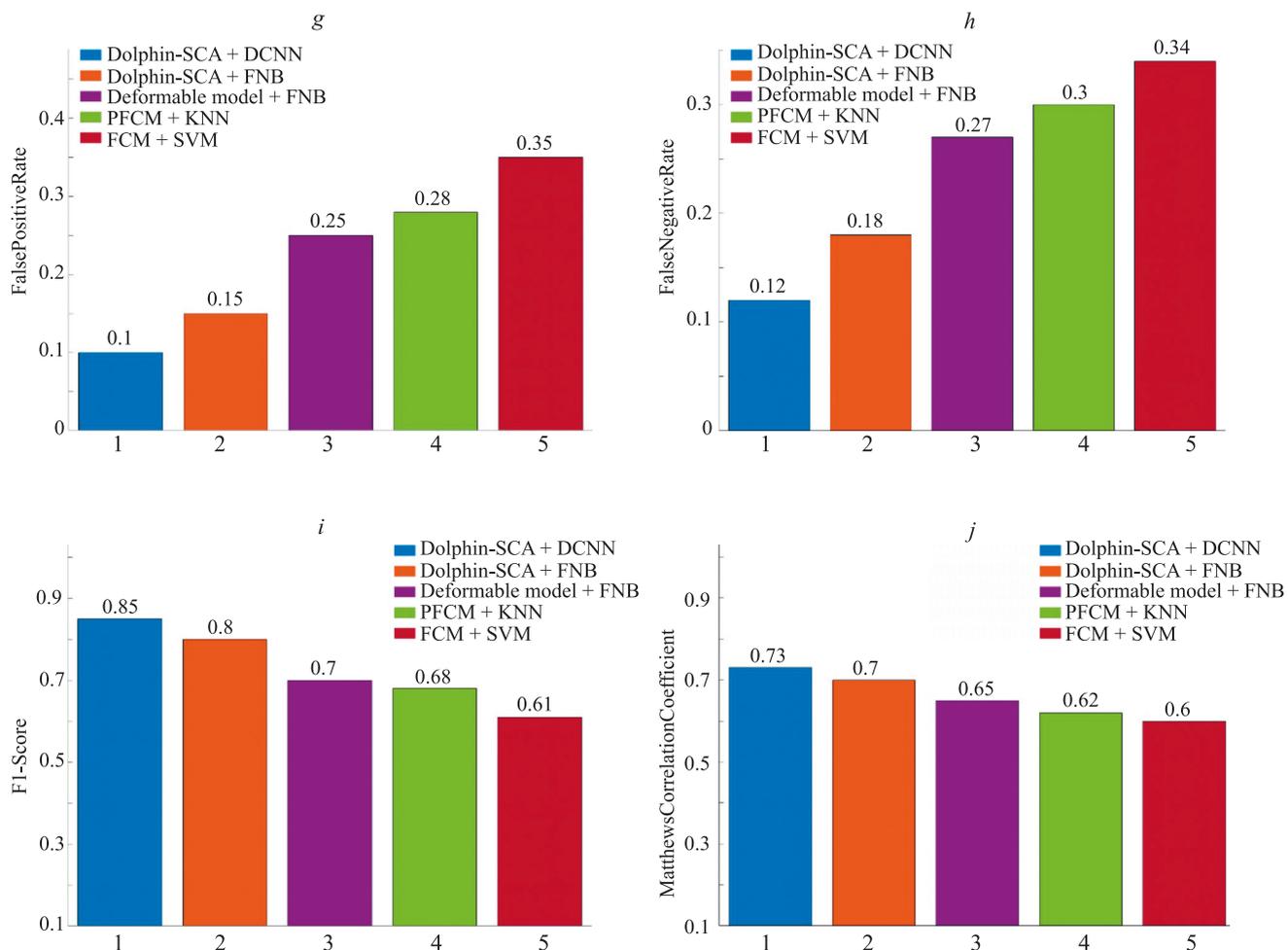


Fig. 4. Continued

### Conclusion

In able to locate and predict brain tumours at an early point, segmentation is critical. In clinical practice, segmenting MRI images takes time since it requires physical evaluation. Additionally, boundary inadequacies like missing borders and loss of textural difference among ROI and backdrop make segmentation of regions difficult. As a result, for speedy and reliable diagnosis of cancers, appropriate segmentation approaches are necessary. In this designed model, the fusion deformable fuzzy model is introduced for segmenting tumour cells from the MRI of the brain. The ROI is employed to partition the region

of the tumour with the bounding box, and the NLM filter is used to remove noise from brain MRI. Finally, output from the fusion deformable fuzzy model is classified using deep CNN. The experimental results of the proposed model are evaluated and compared with existing techniques. Performance metrics of the proposed model, such as Accuracy, Precision, Specificity, Error, Sensitivity etc., are 0.90, 0.89, 0.88, 0.10 and 0.89. Thus, the performance metrics of the proposed model are better compared to existing techniques. In the coming decade, the proposed algorithm could be used to forecast tumour cells from a big collection of brain MRI images.

## References

## Литература

- Amin J., Sharif M., Yasmin M., Saba T., Anjum M.A., Fernandes S.L. A new approach for brain tumor segmentation and classification based on score level fusion using transfer learning. *Journal of Medical Systems*, 2019, vol. 43, no. 11, pp. 326. <https://doi.org/10.1007/s10916-019-1453-8>
- Sultan H.H., Salem N.M., Al-Atabany W. Multi-classification of brain tumor images using deep neural network. *IEEE Access*, 2019, vol. 7, pp. 69215–69225. <https://doi.org/10.1109/access.2019.2919122>
- Noreen N., Palaniappan S., Qayyum A., Ahmad I., Imran M., Shoaib M. A deep learning model based on concatenation approach for the diagnosis of brain tumor. *IEEE Access*, 2020, vol. 8, pp. 55135–55144. <https://doi.org/10.1109/access.2020.2978629>
- Swati Z.N.K., Zhao Q., Kabir M., Ali F., Ali Z., Ahmed S., Lu J. Brain tumor classification for MR images using transfer learning and fine-tuning. *Computerized Medical Imaging and Graphics*, 2019, vol. 75, pp. 34–46. <https://doi.org/10.1016/j.compmedimag.2019.05.001>
- Vaibhavi P., Rupal K. Brain tumor segmentation using K-means-FCM hybrid technique. *Advances in Intelligent Systems and Computing*, 2018, vol. 696, pp. 341–352. [https://doi.org/10.1007/978-981-10-7386-1\\_30](https://doi.org/10.1007/978-981-10-7386-1_30)
- Sandhya S., Chidambararajan B., Kumar M.S. FCM-based segmentation and neural network classification of tumor in brain MRI images. *Advances in Intelligent Systems and Computing*, 2021, vol. 1167, pp. 371–378. [https://doi.org/10.1007/978-981-15-5285-4\\_37](https://doi.org/10.1007/978-981-15-5285-4_37)
- Daimary D., Bora M.B., Amitab K., Kandar D. Brain tumor segmentation from MRI images using hybrid convolutional neural networks. *Procedia Computer Science*, 2020, vol. 167, pp. 2419–2428. <https://doi.org/10.1016/j.procs.2020.03.295>
- Khairandish M.O., Sharma M., Jain V., Chatterjee J.M., Jhanjhi N.Z. A Hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images. *IRBM*, 2022, vol. 43, no. 4, pp. 290–299. <https://doi.org/10.1016/j.irbm.2021.06.003>
- Aygün M., Şahin Y.H., Ünal G. Multi modal convolutional neural networks for brain tumor segmentation. *arXiv*, 2018, arXiv:1809.06191. <https://doi.org/10.48550/arXiv.1809.06191>
- Badža M.M., Barjaktarović M.Č. Classification of brain tumors from MRI images using a convolutional neural network. *Applied Sciences*, 2020, vol. 10, no. 6, pp. 1999. <https://doi.org/10.3390/app10061999>
- Sheela C.J.J., Suganthi G. Automatic brain tumor segmentation from MRI using greedy snake model and fuzzy C-means optimization. *Journal of King Saud University — Computer and Information Sciences*, 2022, vol. 34, no. 3, pp. 557–566. <https://doi.org/10.1016/j.jksuci.2019.04.006>
- Hrosik R.C., Tuba E., Dolicanin E., Jovanovic R., Tuba M. Brain image segmentation based on firefly algorithm combined with k-means clustering. *Studies in Informatics and Control*, 2019, vol. 28, no. 2, pp. 167–176. <https://doi.org/10.24846/v28i2y201905>
- Huang H., Meng F., Zhou S., Jiang F., Manogaran G. Brain image segmentation based on FCM clustering algorithm and rough set. *IEEE Access*, 2019, vol. 7, pp. 12386–12396. <https://doi.org/10.1109/access.2019.2893063>
- Srinivas K., Kantapalli B. Segmentation of MR brain images using unified iterative partitioned fuzzy clustering. *International Journal of Recent Technology and Engineering*, 2019, vol. 8, no. 1, pp. 2755–2758.
- Ali M., Gilani S.O., Waris A., Zafar K., Jamil M. Brain tumour image segmentation using deep networks. *IEEE Access*, 2020, vol. 8, pp. 153589–153598. <https://doi.org/10.1109/access.2020.3018160>
- Kanoun B., Ambrosanio, M., Baselice F., Ferraioli G., Pascasio V., Gómez L. Anisotropic weighted KS-NLM filter for noise reduction in MRI. *IEEE Access*, 2020, vol. 8, pp. 184866–184884. <https://doi.org/10.1109/access.2020.3029297>
- Sun L., Shao W., Zhang D., Liu M. Anatomical attention guided deep networks for ROI segmentation of brain MR images. *IEEE Transactions on Medical Imaging*, 2020, vol. 39, no. 6, pp. 2000–2012. <https://doi.org/10.1109/tmi.2019.2962792>
- Ismael M.R. Edge enhancement based active contour model for segmentation of brain tumor in MRI images. *ECTI Transactions on Electrical Engineering, Electronics, and Communications*, 2021, vol. 19, no. 3. <https://doi.org/10.37936/ecti-ec.2021193.244942>
- Ghosh P., Mali K., Das S.K. Chaotic firefly algorithm-based fuzzy C-means algorithm for segmentation of brain tissues in magnetic
- Amin J., Sharif M., Yasmin M., Saba T., Anjum M.A., Fernandes S.L. A new approach for brain tumor segmentation and classification based on score level fusion using transfer learning // *Journal of Medical Systems*. 2019. V. 43. N 11. P. 326. <https://doi.org/10.1007/s10916-019-1453-8>
- Sultan H.H., Salem N.M., Al-Atabany W. Multi-classification of brain tumor images using deep neural network // *IEEE Access*. 2019. V. 7. P. 69215–69225. <https://doi.org/10.1109/access.2019.2919122>
- Noreen N., Palaniappan S., Qayyum A., Ahmad I., Imran M., Shoaib M. A deep learning model based on concatenation approach for the diagnosis of brain tumor // *IEEE Access*. 2020. V. 8. P. 55135–55144. <https://doi.org/10.1109/access.2020.2978629>
- Swati Z.N.K., Zhao Q., Kabir M., Ali F., Ali Z., Ahmed S., Lu J. Brain tumor classification for MR images using transfer learning and fine-tuning // *Computerized Medical Imaging and Graphics*. 2019. V. 75. P. 34–46. <https://doi.org/10.1016/j.compmedimag.2019.05.001>
- Vaibhavi P., Rupal K. Brain tumor segmentation using K-means-FCM hybrid technique // *Advances in Intelligent Systems and Computing*. 2018. V. 696. P. 341–352. [https://doi.org/10.1007/978-981-10-7386-1\\_30](https://doi.org/10.1007/978-981-10-7386-1_30)
- Sandhya S., Chidambararajan B., Kumar M.S. FCM-based segmentation and neural network classification of tumor in brain MRI images // *Advances in Intelligent Systems and Computing*. 2021. V. 1167. P. 371–378. [https://doi.org/10.1007/978-981-15-5285-4\\_37](https://doi.org/10.1007/978-981-15-5285-4_37)
- Daimary D., Bora M.B., Amitab K., Kandar D. Brain tumor segmentation from MRI images using hybrid convolutional neural networks // *Procedia Computer Science*. 2020. V. 167. P. 2419–2428. <https://doi.org/10.1016/j.procs.2020.03.295>
- Khairandish M.O., Sharma M., Jain V., Chatterjee J.M., Jhanjhi N.Z. A Hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images // *IRBM*. 2022. V. 43. N 4. P. 290–299. <https://doi.org/10.1016/j.irbm.2021.06.003>
- Aygün M., Şahin Y.H., Ünal G. Multi modal convolutional neural networks for brain tumor segmentation // *arXiv*. 2018. arXiv:1809.06191. <https://doi.org/10.48550/arXiv.1809.06191>
- Badža M.M., Barjaktarović M.Č. Classification of brain tumors from MRI images using a convolutional neural network // *Applied Sciences*. 2020. V. 10. N 6. P. 1999. <https://doi.org/10.3390/app10061999>
- Sheela C.J.J., Suganthi G. Automatic brain tumor segmentation from MRI using greedy snake model and fuzzy C-means optimization // *Journal of King Saud University — Computer and Information Sciences*. 2022. V. 34. N 3. P. 557–566. <https://doi.org/10.1016/j.jksuci.2019.04.006>
- Hrosik R.C., Tuba E., Dolicanin E., Jovanovic R., Tuba M. Brain image segmentation based on firefly algorithm combined with k-means clustering // *Studies in Informatics and Control*. 2019. V. 28. N 2. P. 167–176. <https://doi.org/10.24846/v28i2y201905>
- Huang H., Meng F., Zhou S., Jiang F., Manogaran G. Brain image segmentation based on FCM clustering algorithm and rough set // *IEEE Access*. 2019. V. 7. P. 12386–12396. <https://doi.org/10.1109/access.2019.2893063>
- Srinivas K., Kantapalli B. Segmentation of MR brain images using unified iterative partitioned fuzzy clustering // *International Journal of Recent Technology and Engineering*. 2019. V. 8. N 1. P. 2755–2758.
- Ali M., Gilani S.O., Waris A., Zafar K., Jamil M. Brain tumour image segmentation using deep networks // *IEEE Access*. 2020. V. 8. P. 153589–153598. <https://doi.org/10.1109/access.2020.3018160>
- Kanoun B., Ambrosanio, M., Baselice F., Ferraioli G., Pascasio V., Gómez L. Anisotropic weighted KS-NLM filter for noise reduction in MRI // *IEEE Access*. 2020. V. 8. P. 184866–184884. <https://doi.org/10.1109/access.2020.3029297>
- Sun L., Shao W., Zhang D., Liu M. Anatomical attention guided deep networks for ROI segmentation of brain MR images // *IEEE Transactions on Medical Imaging*. 2020. V. 39. N 6. P. 2000–2012. <https://doi.org/10.1109/tmi.2019.2962792>
- Ismael M.R. Edge enhancement based active contour model for segmentation of brain tumor in MRI images // *ECTI Transactions on Electrical Engineering, Electronics, and Communications*. 2021. V. 19. N 3. <https://doi.org/10.37936/ecti-ec.2021193.244942>
- Ghosh P., Mali K., Das S.K. Chaotic firefly algorithm-based fuzzy C-means algorithm for segmentation of brain tissues in magnetic resonance images // *Journal of Visual Communication and Image*

resonance images. *Journal of Visual Communication and Image Representation*, 2018, vol. 54, pp. 63–79. <https://doi.org/10.1016/j.jvcir.2018.04.007>

20. Kumar S., Mankame D.P. Optimization driven deep convolution neural network for brain tumor classification. *Biocybernetics and Biomedical Engineering*, 2020, vol. 40, no. 3, pp. 1190–1204. <https://doi.org/10.1016/j.bbe.2020.05.009>

Representation. 2018. V. 54. P. 63–79. <https://doi.org/10.1016/j.jvcir.2018.04.007>

20. Kumar S., Mankame D.P. Optimization driven deep convolution neural network for brain tumor classification // *Biocybernetics and Biomedical Engineering*. 2020. V. 40. N 3. P. 1190–1204. <https://doi.org/10.1016/j.bbe.2020.05.009>

#### Authors

**Anjali Hemant Tiple** — Associate Professor, Shivaji University, Kolhapur, Maharashtra, 416004, India, <https://orcid.org/0000-0003-4411-3878>, [anjalitiple6@gmail.com](mailto:anjalitiple6@gmail.com)

**Anandrao B. Kakade** — PhD, Associate Professor, Rajarambapu Institute of Technology, Uran Islampur, Maharashtra, 415414, India, [sc 24171107900](https://orcid.org/0000-0002-1882-8901), <https://orcid.org/0000-0002-1882-8901>, [abkakade22@gmail.com](mailto:abkakade22@gmail.com)

#### Авторы

**Хемант Типле Анджали** — доцент, Университет Шиваджи, Колхатур, Махараштра, 416004, Индия, <https://orcid.org/0000-0003-4411-3878>, [anjalitiple6@gmail.com](mailto:anjalitiple6@gmail.com)

**Какаде Анандрао Б.** — PhD, доцент, Технологический институт Раджарамбапу, Уран Ислампур, Махараштра, 415414, Индия, [sc 24171107900](https://orcid.org/0000-0002-1882-8901), <https://orcid.org/0000-0002-1882-8901>, [abkakade22@gmail.com](mailto:abkakade22@gmail.com)

*Received 03.08.2022*

*Approved after reviewing 17.05.2023*

*Accepted 18.07.2023*

*Статья поступила в редакцию 03.08.2022*

*Одобрена после рецензирования 17.05.2023*

*Принята к печати 18.07.2023*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-786-794

УДК 004.9

## Оптимизация систем отслеживания человека в виртуальной реальности на основе нейросетевого подхода

Артём Дмитриевич Обухов<sup>1</sup>, Даниил Вячеславович Теселкин<sup>2</sup>

<sup>1,2</sup> Тамбовский государственный технический университет, Тамбов, 392000, Российская Федерация

<sup>1</sup> [obuhov.art@gmail.com](mailto:obuhov.art@gmail.com), <https://orcid.org/0000-0002-3450-5213>

<sup>2</sup> [teselkin.dv@tstu.ru](mailto:teselkin.dv@tstu.ru), <https://orcid.org/0000-0002-1304-9490>

### Аннотация

**Введение.** Рассмотрена проблема определения оптимального количества и расположения точек отслеживания на теле человека для обеспечения необходимой точности реконструкции кинематических параметров движений человека в виртуальном пространстве. Выполнена оптимизация системы отслеживания человека в виртуальной реальности для снижения объема передаваемой информации, вычислительной нагрузки и стоимости систем захвата движений за счет снижения количества физических датчиков. **Метод.** Поставлена задача оптимизации количества и расположения точек отслеживания на теле человека, необходимых для реконструкции виртуальной модели тела из ограниченного набора входных точек, с использованием численной аппроксимации функции регрессии. Разработан алгоритм сбора большого количества данных с модели тела человека в виртуальной сцене и с костюма захвата движений в реальном мире. **Основные результаты.** Получено наименьшее количество точек отслеживания тела человека и их расположение с использованием предложенного алгоритма. Обучены и протестированы различные топологии нейронных сетей, позволяющие аппроксимировать регрессионную зависимость между ограниченным по размеру вектором точек отслеживания (от 3 до 13) и вектором 18 виртуальных точек, используемых для полной реконструкции модели тела человека. Необходимая точность реконструкции кинематических параметров движений человека обеспечено при 5 и 7 входных точках.

**Обсуждение.** Предложенный подход позволил использовать 5 или 7 физических датчиков для построения модели тела человека и восстановления кинематических параметров его движений в виртуальной реальности. Подход может быть применен при решении задач инверсной кинематики с целью снижения количества физических датчиков, размещенных на поверхности исследуемого объекта, для упрощения процессов обработки и передачи информации. За счет объединения данных как с костюма захвата движений, так и с виртуального аватара значительно ускорен процесс сбора информации, расширен объем обучающей выборки и смоделированы различные паттерны движений тел пользователей.

### Ключевые слова

виртуальная реальность, отслеживание движений человека, инверсная кинематика, оптимизация систем отслеживания и захвата движений, цифровое представление человека

### Благодарности

Исследование выполнено за счет гранта Российского научного фонда (№ 22-71-10057, <https://rscf.ru/project/22-71-10057/>).

**Ссылка для цитирования:** Обухов А.Д., Теселкин Д.В. Оптимизация систем отслеживания человека в виртуальной реальности на основе нейросетевого подхода // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 786–794. doi: 10.17586/2226-1494-2023-23-4-786-794

## Optimization of human tracking systems in virtual reality based on a neural network approach

Artem D. Obukhov<sup>1</sup>, Daniil V. Teselkin<sup>2</sup>

<sup>1,2</sup> Tambov State Technical University, Tambov, 392000, Russian Federation

<sup>1</sup> [obuhov.art@gmail.com](mailto:obuhov.art@gmail.com), <https://orcid.org/0000-0002-3450-5213>

<sup>2</sup> [teselkin.dv@tstu.ru](mailto:teselkin.dv@tstu.ru), <https://orcid.org/0000-0002-1304-9490>

© Обухов А.Д., Теселкин Д.В., 2023

**Abstract**

The problem of determining the optimal number and location of tracking points on the human body to ensure the necessary accuracy of reconstruction of kinematic parameters of human movements in virtual space is considered. Optimization of the human tracking system in virtual reality has been performed to reduce the amount of transmitted information, computational load and cost of motion capture systems by reducing the number of physical sensors. The task of optimizing the number and location of tracking points on the human body necessary for the reconstruction of a virtual body model from a limited set of input points using numerical approximation of the regression function is set. An algorithm has been developed for collecting a large amount of data from a human body model in a virtual scene and from a motion capture suit in the real world. The smallest number of human body tracking points and their location were obtained using the proposed algorithm. Various neural network topologies have been trained and tested to approximate the regression relationship between a vector of tracking points limited in size (from 3 to 13) and a vector of 18 virtual points used for the complete reconstruction of the human body model. The necessary accuracy of reconstruction of kinematic parameters of human movements is provided at 5 and 7 input points. The proposed approach made it possible to use 5 or 7 physical sensors to build a model of the human body and restore the kinetic parameters of its movements in virtual reality. The approach can be applied to solving inverse kinematics problems in order to reduce the number of physical sensors placed on the surface of the object under study, to simplify the processing and transmission of information. By combining data from both the motion capture suit and the virtual avatar, the process of collecting information has been significantly accelerated, the volume of the training sample has been expanded and various patterns of user body movements have been modeled.

**Keywords**

virtual reality, human movement tracking, inverse kinematics, optimization of motion tracking and capture systems, digital representation of a person

**Acknowledgements**

The research was carried out at the expense of the grant of the Russian Science Foundation (no. 22-71-10057, <https://rscf.ru/project/22-71-10057/>).

**For citation:** Obukhov A.D., Teselkin D.V. Optimization of human tracking systems in virtual reality based on a neural network approach. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 786–794 (in Russian). doi: 10.17586/2226-1494-2023-23-4-786-794

**Введение**

На текущем этапе развития систем виртуальной реальности (VR) одной из актуальных проблем является точное воспроизведение тела пользователя в виртуальном пространстве [1]. Под виртуальным пространством подразумевается смоделированная среда, имитирующая реальное или вымышленное окружение посредством трехмерных моделей, анимации, звуковых и визуальных эффектов, а также обеспечивающее пользователю возможность взаимодействия с виртуальными объектами с помощью специального оборудования, например систем VR [2]. Базовый комплект поставки большинства систем VR включает шлем и два контроллера для перемещения и организации взаимодействия с объектами. Отслеживание действий может производиться по камерам, установленным в шлеме, с использованием внешних датчиков и базовых станций (технология Lighthouse) или с применением костюмов захвата движений. В первом случае возможно использование технологий компьютерного зрения, однако они обеспечивают качественное воспроизведение только рук, в частности ладоней и пальцев, распознавание остальных частей тела из-за специфики расположения камер затруднено. Использование технологии Lighthouse позволяет точно позиционировать шлем, контроллеры и дополнительные датчики (трекеры). Недостатком данного подхода является неудобство одновременного закрепления на поверхности тела человека большого количества датчиков. Третий подход основан на использовании дорогостоящих и зачастую чувствительных к внешним электромагнитным помехам костюмов с множеством инерциальных датчиков, что значительно ограничивает их использование в массовых решениях.

Разработчики программного обеспечения для VR осуществляют решение проблемы визуализации тела человека в виртуальном пространстве двумя основными методами [3]: с использованием прямой и инверсной кинематики. В первом случае требуется достаточный объем информации обо всех ключевых точках тела пользователя, так как кинематика моделируется сверху вниз, начиная, например от плеча, далее к локтю и заканчивая кистью. В инверсной кинематике, наоборот, перемещение дочерних сегментов приводит к изменению положения родительских, т. е. алгоритм рассчитывает положение и ориентацию плеча и локтя, исходя из положения и ориентации кисти. Первый подход используется в костюмах захвата движений, второй — в условиях ограниченного объема информации о точках тела человека, например, когда разработчику доступно только положение кистей пользователя, по которому восстанавливаются координаты точек рук, туловища и ног (учитывая текущую высоту шлема).

В настоящей работе рассмотрена проблема отслеживания положения тела человека для последующей реконструкции в VR кинематических параметров его движений с учетом минимизации количества отслеживаемых точек.

**Обзор существующих исследований по отслеживанию человека в VR**

На первом этапе рассмотрены существующие направления исследований в данной области.

В работе [4] использовано пять дополнительных закрепленных на спине, пятках и локтях человека трекеров, что в совокупности с двумя контроллерами и шлемом составляет 8 ключевых точек, на основе которых

разработчики решили задачу инверсной кинематики. Полученная система обладает достаточно низкой задержкой и высокой точностью по субъективной оценке участников контрольной группы.

Работа [5] основана на анализе походки человека и сопоставлении положения головы в шлеме VR с фазами шага, что позволило использовать только одну точку для восстановления кинематики ног с использованием различных подходов (Threshold, метода на основе корреляции Пирсона, Support Vector Machine, Bidirectional Long Short-Term Memory (BLSTM)). Архитектура BLSTM обеспечила наименьшую погрешность, однако недостаточно производительна для применения в режиме реального времени. По этой причине наиболее предпочтителен метод, основанный на расчете корреляции Пирсона между ускорением головы и ног. Данный подход не является универсальным и ограничен сценарием равномерной ходьбы.

Ряд научных работ сконцентрирован на более точной реконструкции только отдельных частей тела, например рук и пальцев. Для этого могут использоваться комбинации из перчаток VR и дополнительных акселерометров для каждого пальца [6] или сочетание камер с датчиком глубины и нейронных сетей для более точного восстановления трехмерных координат [7]. Нейронные сети также применены для успешной реконструкции трехмерной модели тела человека по двумерному изображению, на котором сначала распознаются основные точки скелета, а потом осуществляется их преобразование в трехмерное представление модели тела с коррекцией углов поворота сегментов [8].

Применение методов машинного обучения для аппроксимации функций регрессии при решении задач инверсной кинематики осуществляется достаточно давно, но ранние работы [9] в этой области отмечали низкую точность и производительность данных подходов. Развитие программного и аппаратного обеспечения в сфере машинного обучения позволило обеспечить прогресс по данному направлению: сочетания различных нейронных сетей используются для реализации инверсной кинематики роботов [10], прогнозирования положения тела с учетом окружающей среды на основе виртуального скелета [11] или набора датчиков [12]. В [13] рассмотрено применение нейронных сетей для восстановления ключевых точек силуэта человека при удалении от 10 до 30 % данных.

Таким образом, проведенный анализ показал, что для решения задачи инверсной кинематики оправдано использование небольшого количества датчиков и методов численной аппроксимации функций регрессии (например, нейронных сетей различной топологии) в качестве инструмента восстановления недостающих точек виртуальной модели тела человека. Определение минимального количества точек отслеживания позволяет снизить объем передаваемой информации, вычислительную нагрузку, а в случае использования трекеров VR — стоимость реализации системы отслеживания движений.

### Постановка задачи оптимизации количества и расположения точек отслеживания тела человека в VR

Для достижения цели работы на первом этапе выполнен анализ предметной области. Далее формально определены основные объекты и поставлена задача определения оптимального количества и расположения точек отслеживания на теле человека.

Пусть задан набор исходных точек (физических трекеров VR, датчиков, виртуальных маркеров и др.), соответствующих ключевым зонам человеческого тела и позволяющих осуществить его последующую реконструкцию в виртуальном пространстве. Обозначим количество датчиков (трекеров) через  $TK$ . Каждый трекер  $tr_i \in TR$  определим набором реальных точек с координатами по трем осям:

$$TR = \{tr_i\},$$

$$tr_i = \langle x_i, y_i, z_i \rangle.$$

Обозначим через  $\mathbf{R}$  вектор множества значений координат всех трекеров:

$$\mathbf{R} = \{x_1, y_1, z_1, \dots, x_{TK}, y_{TK}, z_{TK}\}.$$

Размер  $\mathbf{R}$  обозначим как  $N$ . Так как отслеживается положение трекеров по трем осям, то  $N = 3TK$ . Пусть  $\mathbf{R}$  с размером  $N$  однозначно определяет цифровое представление тела человека (виртуальный аватар) с необходимой точностью (т. е. позволяет реализовать все необходимые части модели виртуального тела, их соединение и воспроизвести кинематические параметры процесса движения тела человека).

Оптимизация системы отслеживания тела человека состоит в определении минимального размера  $N$  вектора  $\mathbf{R}$ , т. е. нахождения количества трекеров, позволяющих обеспечить необходимую точность реконструкции цифрового представления человека. Так как уменьшение  $\mathbf{R}$  при реализации инверсной кинематики модели тела человека приведет к увеличению погрешности при определении координат частей тела виртуального аватара, необходимо восстановить по наблюдениям регрессионную зависимость  $\mathbf{F}$  вида:

$$\mathbf{F}(\mathbf{R}) = \mathbf{VR},$$

которая отображает пространство  $\mathbb{R}^N$  реализации  $\mathbf{R}$  в некоторое пространство  $\mathbb{R}^K$ , в котором реализованы величины  $\mathbf{VR}$  (вектора виртуальных точек), определенным образом вычисляемые на основании величин  $\mathbf{R}$ . В качестве аппроксимации регрессии  $\mathbf{F}$  использована некоторая  $NN$  определенной топологии. Топология указанной нейронной сети зависит от длины  $N$  вектора  $\mathbf{R}$ , которая определяет размерность входного слоя нейронной сети, скрытые слои для различных  $NN$  могут быть одинаковыми.

В формализованном виде задача оптимизации заключается в следующем: необходимо найти минимальный размер  $N$  входного вектора  $\mathbf{R}$ , а также определить параметры и топологию  $NN_N$  для каждого  $N$ , при которых среднеквадратичное отклонение восстановленных

элементов  $\mathbf{VR}$  от истинных значений будет минимальным.

Для решения поставленной задачи необходимо сформировать обучающую выборку достаточных размеров и обучить набор  $NN_N$  для каждого выбранного  $N$ .

### Алгоритм формирования данных для обучения нейронных сетей

Для формирования достаточного объема обучающей выборки необходим сбор информации, например, с костюмов захвата движений [14], которые позволяют сформировать полное представление о положении и кинематических параметрах модели тела человека в виртуальном пространстве. Отметим, что такой подход отличается значительными временными затратами и необходимостью привлечения большого количества участников с различными физическими параметрами. Главным недостатком является невозможность смоделировать в реальном мире ряд труднодоступных паттернов перемещения и небезопасных движений.

Решение данной проблемы — объединение данных от двух источников: систем захвата движений и виртуального аватара из игровой сцены [15]. Так как процедура построения цифрового представления модели тела человека в виртуальном пространстве осуществляется на основе набора точек в метрической системе координат, то для конечной реализации виртуальной модели тела человека источник данных не будет иметь принципиальной разницы. Предложенный подход дает возможность значительно ускорить процесс сбора данных за счет записи с повышенной частотой кадров и параллельно на нескольких виртуальных сценах.

Проведенный анализ виртуальных моделей тела человека показал, что для его реконструкции достаточно 18 ключевых точек. Это соответствует многим коммерческим решениям, например, Perception Neuron, основанном на 18 сенсорах [16]. Таким образом, при записи данных необходимо фиксировать 54 значения  $\mathbf{R}$ . Расположение этих точек должно соответствовать узлам соединения основных сегментов тела человека: кисть, локоть, плечо, крайние точки стопы, колено, тазобедренный сустав, крайние и центральная точка спины, голова.

На следующем этапе осуществлен выбор сочетаний трекеров на основе эмпирического подхода в порядке их приоритетности и удобства закрепления на поверхности тела человека: начиная с крайних точек на руках и ногах, заканчивая промежуточными положениями (на коленях, локтях, спине и т. д.). В итоге сформулировано 6 предпочтительных наборов, начиная от простейшего с 3 точками и заканчивая максимальным с 13 точками:

- 3 точки: голова, правая и левая кисть;
- 5 точек: 3 точки с добавлением правой и левой стопы (зона пяты);
- 7 точек: 5 точек с добавлением правого и левого локтя;
- 9 точек: 7 точек с добавлением правого и левого колена;
- 11 точек: 9 точек с добавлением правого и левого тазобедренного сустава;

— 13 точек: 11 точек с добавлением верхней и нижней точек спины.

После обучения нейронных сетей на собранной базе данных осуществляется их тестирование на контрольном наборе и визуализация модели тела человека для оценки величины отклонения положения восстановленных точек от реальных. Дополнительно рассчитывается среднеквадратичное отклонение координат восстановленных точек позы от ожидаемых (среднее по всем точкам и максимальное по всему телу человека).

### Результаты экспериментальных исследований

На первом этапе экспериментальных исследований осуществлен сбор данных с различных источников: типовые движения тела нескольких людей с разными физическими параметрами и воспроизведенные анимации в виртуальной сцене (источник анимаций — CMU Graphics Lab Motion Capture Database<sup>1</sup>). В совокупности это позволило собрать 367 млн записей 18 трекеров (54 точки). В качестве обучающей выборки использовано 3583 тыс. записей, в качестве контрольной выборки для тестирования — 80 тыс. записей.

На предварительном этапе рассмотрено несколько топологий нейронных сетей, часть из которых из-за недостаточного или избыточного количества нейронов и слоев не достигла необходимых значений точности. Для выбора оптимальной топологии нейронной сети проведен анализ следующих гиперпараметров:

- наличие слоев разреживания (dropout): {да, нет};
- увеличение количества нейронов на каждом последующем слое: {да, нет};
- тип слоев: {Dense, Conv1D, RNN, LSTM};
- количество скрытых слоев: {1, 2, 3, 4, 5, 6, 7, 8, 9, 10};
- количество нейронов в скрытом слое: {20, 50, 100, 200, 400, 800, 1600};
- количество эпох {5, 10, 15, 25, 50}.

Результаты выбора наилучших гиперпараметров представлены в табл. 1. Для выбора параметров в целях ускорения процесса обучения использована ограниченная выборка из 5000 элементов тренировочного набора и фиксированный вход из трех точек модели человека. В качестве функции потерь используется среднеквадратичная ошибка — Mean Square Error (MSE). Для выбора каждого гиперпараметра выполнена фиксация остальных элементов и полный перебор значений текущего гиперпараметра. Далее осуществлен переход к следующему параметру с фиксацией наилучшего значения предыдущих параметров.

Наилучшие показатели получены с применением многослойной сети из пяти полносвязных (Dense) слоев с количеством нейронов от 400 до 1600, а также двумя слоями разреживания. Для финальной архитектуры создано 6 моделей, размер выходного слоя каждой — 54 нейрона, входного — варьируется от 9 (для трех то-

<sup>1</sup> B. Hahne. The daz-friendly bvh release of CMU motion capture database, 2010 [Электронный ресурс]. URL: <https://sites.google.com/a/cgspeed.com/cgspeed/motion-capture/daz-friendly-release> (дата обращения: 01.09.2023).

Таблица 1. Сравнение различных архитектур нейронных сетей  
 Table 1. Comparison of different neural network architectures

Гиперпараметр	Выбранное значение	Значение ошибки MSE	Анализ альтернативных вариантов
Наличие слоев разреживания (dropout)	да	0,008	При выборе «нет» MSE выше (0,010)
Увеличение количества нейронов на каждом последующем слое	да	0,009	При выборе «нет» MSE выше (0,014)
Тип слоев	Dense	0,009	С учетом структуры данных все альтернативные типы слоев показали сравнимые результаты (от 0,010 до 0,012), не превосходящие Dense, поэтому выбран наиболее производительный и простой тип слоев
Количество скрытых слоев	5	0,008	MSE выше при меньшем количестве слоев. При увеличении количества слоев растет сложность модели без значительного улучшения MSE. После 7–8 слоев начинается рост MSE
Количество нейронов в скрытом слое	400, 400, 800, 800, 1600	0,006	Получены сравнимые результаты для всех вариантов до 400 нейронов включительно, дальше наблюдается ухудшение. Меньшее количество нейронов не выбрано для универсальности архитектуры на большом наборе данных и при различном количестве входных точек
Количество эпох	10	0,005	После 10-й эпохи процесс обучения замедлился, скорость уменьшения MSE значительно упала

чек) до 39 (для 13 точек). Процесс обучения выполнен в течение 10 эпох и представлен на рис. 1. В результате можно сделать вывод, что  $NN_3$  обладает наименьшей точностью, но, так как в процессе обучения точность обученных нейронных сетей соизмерима, необходимы дополнительные исследования для оценки полученных результатов. Для проверки качества нейронных сетей и величины отклонения восстановленных точек от реальных проведена реконструкция виртуальной модели тела человека. Ее результаты для всех нейронных сетей показаны на рис. 2. Красные линии соответствуют

эталонным значениям точкам модели тела человека, а синие — восстановленным с применением соответствующей  $i$ -ой нейронной сети.

Для контрольных данных рассчитано среднее отклонение по всем точкам полученных виртуальных моделей тела пользователя (в метрах), максимальное отклонение среди всех точек тела, а также средние отклонения этих величин в выборке. Результаты представлены в табл. 2.

Анализ производительности обученных нейронных сетей показал, что среднее время реконструкции

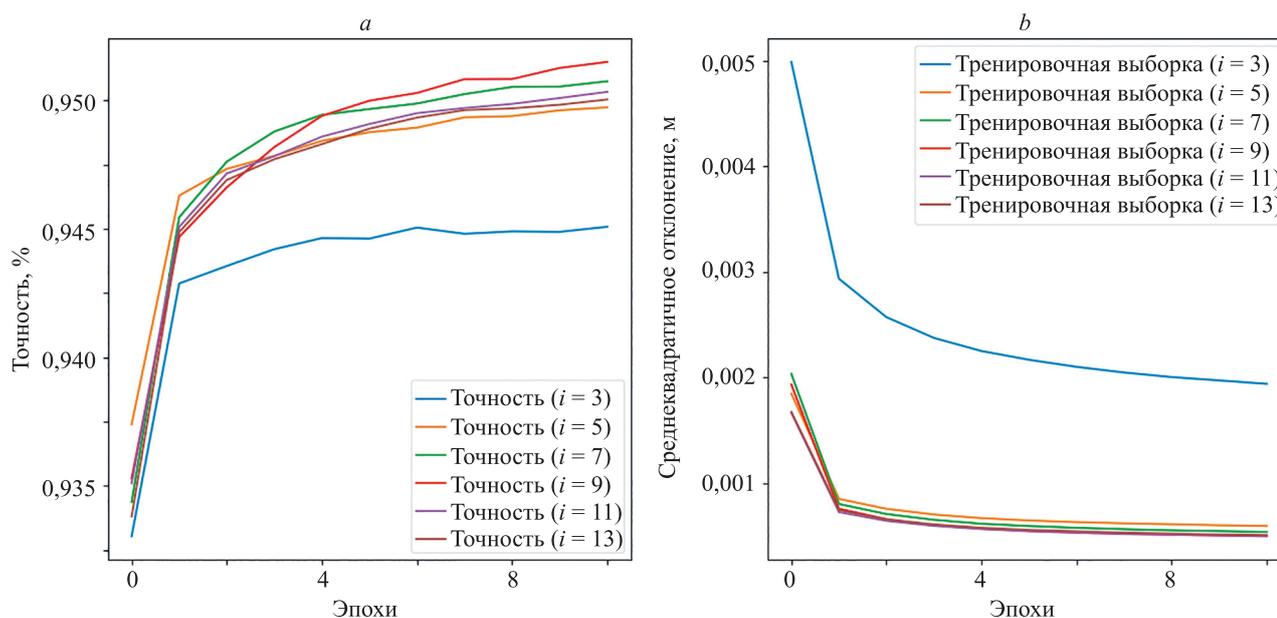


Рис. 1. Процесс обучения нейронной сети: точность нейронной сети (a); значения функций потерь (MSE) (b)  
 Fig. 1. Neural network training process: neural network accuracy (a); loss function values (MSE) (b)

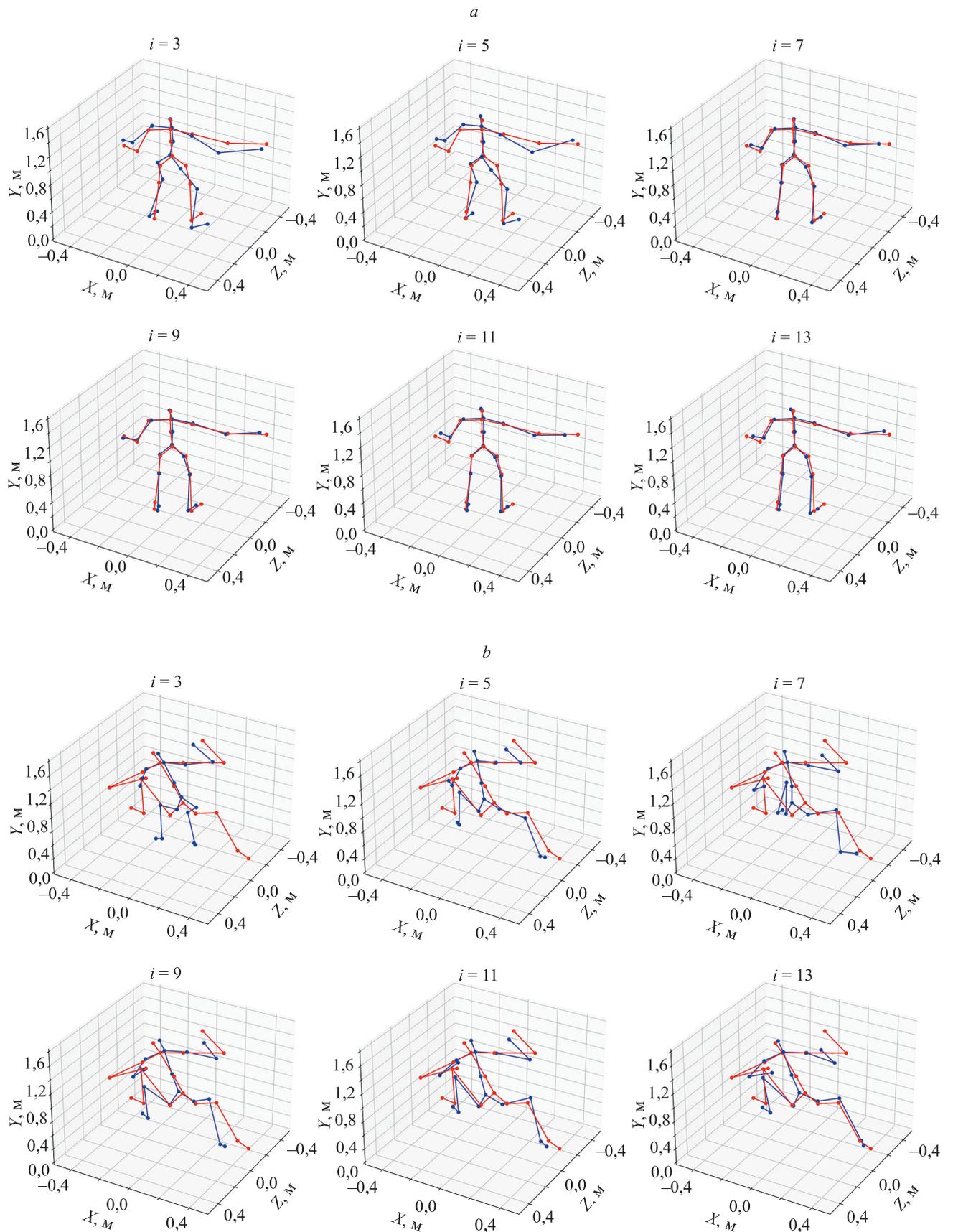


Рис. 2. Реконструкции виртуальных моделей тела для различных нейронных сетей: первая (а) и вторая (б) позы  
 Fig. 2. Reconstructions of virtual body models for various neural networks: the first (a) and second (b) poses

Таблица 2. Сравнение погрешности нейронных сетей на тестовой выборке  
 Table 2. Comparison of the error of neural networks on a test sample

$NN_i$	Среднее отклонение по всем точкам модели тела человека (среднее по выборке $\pm$ отклонение), м	Максимальное отклонение среди всех точек модели тела человека (среднее по выборке $\pm$ отклонение), м
3	0,0771 $\pm$ 0,0459	0,2358 $\pm$ 0,1290
5	0,0420 $\pm$ 0,0236	0,1021 $\pm$ 0,0559
7	0,0306 $\pm$ 0,0186	0,0761 $\pm$ 0,0485
9	0,0298 $\pm$ 0,0198	0,0723 $\pm$ 0,0474
11	0,0295 $\pm$ 0,0186	0,0737 $\pm$ 0,0406
13	0,0299 $\pm$ 0,0168	0,0745 $\pm$ 0,0376

модели тела человека составляет от 1,6 до 2,8 мс, что обеспечивает необходимое быстродействие (до 300 вызовов в секунду) и не повлияет на работоспособность систем VR.

Проведенные исследования показали, что использование трех точек не является достаточным, так как обученная таким образом  $NN_3$  не может восстанавливать корректное положение ног.  $NN_5$  точнее позиционирует нижние конечности. Использование 7 точек повышает точность реконструкции при сложных положениях рук, особенно в локтевой зоне. Далее применение 9 и более точек оказывает некоторое положительное влияние, но оно не является значительным и визуально значимым.

Полученные нейронные сети могут использоваться для решения следующих задач:

- $NN_3$ : вариант для упрощенных систем захвата движений, в которых не требуется отслеживание ног и точное позиционирование рук;
- $NN_5$ : оптимальное решение для большинства систем VR благодаря возможности достаточно точно позиционировать все конечности пользователя;
- $NN_7$ : решение для систем, требующих высокой точности позиционирования верхних конечностей, обеспечивающее сравнимую с более сложными моделями погрешность восстановления.

С учетом необходимости минимизации количества трекеров для широкого спектра систем VR оптимальная точность реконструкции тела человека достигается при 5 и 7 входных точках.

Интеграция обученных нейронных сетей в среды разработки систем VR требует изменения их архитектуры путем добавления дополнительных модулей.

Рассмотрим их специфику.

**Модуль формирования цифрового представления** человека используется для взаимодействия с нейронной сетью. В ходе запуска модуля указывается количество и расположение входных трекеров в зависимости от текущих настроек системы. Изначально подмодуль настроен для работы с игровым движком Unity и библиотекой SteamVR для взаимодействия с трекерами VR, но может быть модифицирован, так как предоставляет интерфейс для получения точек трекеров извне и передачу восстановленного тела человека в произвольную систему.

**Модуль инверсной кинематики** осуществляет достраивание виртуального аватара на основе совокупности точек, добавляя необходимые связи между

ними. Также модуль учитывает некоторые физические ограничения, что позволяет корректировать положение точек, если они выходят за границы допустимых пределов.

**Модуль визуализации** формирует виртуальную сцену и аватар пользователя, построенный на основе данных от модуля инверсной кинематики.

Функционирование системы тогда основано не на прямом использовании точек от трекеров VR, а на их отправке в модуль формирования цифрового представления, получение восстановленного набора точек модели тела человека, формирование виртуального аватара и изменение положения итоговой модели человека. Сформированный таким образом аватар будет максимально приближен к реальному положению тела с учетом минимального количества отслеживаемых точек.

Обученные нейронные сети и проект, реализующий интеграцию предложенной концепции по реконструкции модели тела человека, размещен в репозитории GitHub<sup>1</sup>.

### Заключение

Рассмотрена задача минимизации точек для обеспечения необходимой точности реконструкции кинематических параметров движений человека в виртуальном пространстве. Использование специализированных костюмов позволяет эффективно решать данный класс задач, однако подобный подход связан с большими затратами, а оборудование достаточно чувствительно к помехам и неудобно для повседневного использования. Использование же небольшого количества датчиков (трекеров виртуальной реальности, контроллеров) для отслеживания ключевых точек тела человека является перспективным направлением с точки зрения материальных и вычислительных затрат. Для точной реконструкции тела человека необходимо использование 18 точек отслеживания (датчиков), что не является оптимальным решением с экономической и эргономической точки зрения (датчики имеют достаточно большой размер).

Предложен подход, основанный на численной аппроксимации регрессионной зависимости между огра-

<sup>1</sup> Репозиторий GitHub DigitalShadowInMotion [Электронный ресурс]. URL: <https://github.com/Obukhov-Artem/DigitalShadowInMotion> (дата обращения: 22.05.2023).

ническим по размеру вектором точек отслеживания и вектором 18 виртуальных точек, используемых для полной реконструкции модели тела человека. В рамках подхода сформулирована регрессионная зависимость между реальными и виртуальными точками, поставлена задача оптимизации количества трекеров. Выполнены экспериментальные исследования и обучены топологии нейронных сетей для различных наборов входных точек (от 3 до 13). Оптимальная точность восстановления точек получена при 5 и 7 входных точках, что позво-

ляет упростить реализацию систем захвата движений, построение виртуальной модели тела человека и восстановление кинетических параметров его движений в виртуальной реальности. Новизна исследования также включает подход к сбору данных не только с реального человека, но и с виртуальной модели, что позволило ускорить процесс сбора информации для обучения нейронных сетей и увеличить объем данных за счет моделирования различных паттернов движения тела человека в виртуальном пространстве.

## Литература

1. Obukhov A.D., Volkov A.A., Vekhteva N.A., Teselkin D.V., Arkhipov A.E. Human motion capture algorithm for creating digital shadows of the movement process // *Journal of Physics: Conference Series*. 2022. V. 2388. N 1. P. 012033. <https://doi.org/10.1088/1742-6596/2388/1/012033>
2. Azarby S., Rice A. Understanding the effects of virtual reality system usage on spatial perception: The potential impacts of immersive virtual reality on spatial design decisions // *Sustainability*. 2022. V. 14. N 16. P. 10326. <https://doi.org/10.3390/su141610326>
3. Parger M., Mueller J.H., Schmalstieg D., Steinberger M. Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality // *Proc. of the 24th ACM Symposium on Virtual Reality Software and Technology*. 2018. P. 1–10. <https://doi.org/10.1145/3281505.3281529>
4. Caserman P., Garcia-Agundez A., Konrad R., Göbel S., Steinmetz R. Real-time body tracking in virtual reality using a Vive tracker // *Virtual Reality*. 2019. V. 23. N 2. P. 155–168. <https://doi.org/10.1007/s10055-018-0374-z>
5. Feigl T., Gruner L., Mutschler C., Roth D. Real-time gait reconstruction for virtual reality using a single sensor // *Proc. of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 2020. P. 84–89. <https://doi.org/10.1109/ismar-adjunct51615.2020.00037>
6. Liu H., Zhang Z., Xie X., Zhu Y., Liu Y., Wang Y., Zhu S.-C. High-fidelity grasping in virtual reality using a glove-based system // *Proc. of the 2019 International Conference on Robotics and Automation (ICRA)*. 2019. P. 5180–5186. <https://doi.org/10.1109/icra.2019.8794230>
7. Liu R., Liu C. Human motion prediction using adaptable recurrent neural networks and inverse kinematics // *IEEE Control Systems Letters*. 2021. V. 5. N 5. P. 1651–1656. <https://doi.org/10.1109/lcsys.2020.3042609>
8. Li J., Xu C., Chen Z., Bian S., Yang L., Lu C. Hybrik: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021. P. 3383–3392. <https://doi.org/10.1109/cvpr46437.2021.00339>
9. Oyama E., Agah A., MacDorman K.F., Maeda T., Tachi S. A modular neural network architecture for inverse kinematics model learning // *Neurocomputing*. 2001. V. 38–40. P. 797–805. [https://doi.org/10.1016/S0925-2312\(01\)00416-7](https://doi.org/10.1016/S0925-2312(01)00416-7)
10. Bai Y., Luo M., Pang F. An algorithm for solving robot inverse kinematics based on FOA optimized BP neural network // *Applied Sciences*. 2021. V. 11. N 15. P. 7129. <https://doi.org/10.3390/app11157129>
11. Kratzer P., Toussaint M., Mainprice J. Prediction of human full-body movements with motion optimization and recurrent neural networks // *Proc. of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*. 2020. P. 1792–1798. <https://doi.org/10.1109/icra40945.2020.9197290>
12. Bataineh M., Timothy M., Karim A.-M., Jasbir A. Neural network for dynamic human motion prediction // *Expert Systems with Applications*. 2016. V. 48. P. 26–34. <https://doi.org/10.1016/j.eswa.2015.11.020>
13. Kucherenko T., Beskow J., Kjellström H. A neural network approach to missing marker reconstruction in human motion capture // *arXiv*. 2018. arXiv:1803.02665. <https://doi.org/10.48550/arXiv.1803.02665>

## References

1. Obukhov A.D., Volkov A.A., Vekhteva N.A., Teselkin D.V., Arkhipov A.E. Human motion capture algorithm for creating digital shadows of the movement process. *Journal of Physics: Conference Series*. 2022, vol. 2388, no. 1, pp. 012033. <https://doi.org/10.1088/1742-6596/2388/1/012033>
2. Azarby S., Rice A. Understanding the effects of virtual reality system usage on spatial perception: The potential impacts of immersive virtual reality on spatial design decisions. *Sustainability*, 2022, vol. 14, no. 16, pp. 10326. <https://doi.org/10.3390/su141610326>
3. Parger M., Mueller J.H., Schmalstieg D., Steinberger M. Human upper-body inverse kinematics for increased embodiment in consumer-grade virtual reality. *Proc. of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–10. <https://doi.org/10.1145/3281505.3281529>
4. Caserman P., Garcia-Agundez A., Konrad R., Göbel S., Steinmetz R. Real-time body tracking in virtual reality using a Vive tracker. *Virtual Reality*, 2019, vol. 23, no. 2, pp. 155–168. <https://doi.org/10.1007/s10055-018-0374-z>
5. Feigl T., Gruner L., Mutschler C., Roth D. Real-time gait reconstruction for virtual reality using a single sensor. *Proc. of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, 2020, pp. 84–89. <https://doi.org/10.1109/ismar-adjunct51615.2020.00037>
6. Liu H., Zhang Z., Xie X., Zhu Y., Liu Y., Wang Y., Zhu S.-C. High-fidelity grasping in virtual reality using a glove-based system. *Proc. of the 2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 5180–5186. <https://doi.org/10.1109/icra.2019.8794230>
7. Liu R., Liu C. Human motion prediction using adaptable recurrent neural networks and inverse kinematics. *IEEE Control Systems Letters*, 2021, vol. 5, no. 5, pp. 1651–1656. <https://doi.org/10.1109/lcsys.2020.3042609>
8. Li J., Xu C., Chen Z., Bian S., Yang L., Lu C. Hybrik: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 3383–3392. <https://doi.org/10.1109/cvpr46437.2021.00339>
9. Oyama E., Agah A., MacDorman K.F., Maeda T., Tachi S. A modular neural network architecture for inverse kinematics model learning. *Neurocomputing*, 2001, vol. 38–40, pp. 797–805. [https://doi.org/10.1016/S0925-2312\(01\)00416-7](https://doi.org/10.1016/S0925-2312(01)00416-7)
10. Bai Y., Luo M., Pang F. An algorithm for solving robot inverse kinematics based on FOA optimized BP neural network. *Applied Sciences*, 2021, vol. 11, no. 15, pp. 7129. <https://doi.org/10.3390/app11157129>
11. Kratzer P., Toussaint M., Mainprice J. Prediction of human full-body movements with motion optimization and recurrent neural networks. *Proc. of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 1792–1798. <https://doi.org/10.1109/icra40945.2020.9197290>
12. Bataineh M., Timothy M., Karim A.-M., Jasbir A. Neural network for dynamic human motion prediction. *Expert Systems with Applications*, 2016, vol. 48, pp. 26–34. <https://doi.org/10.1016/j.eswa.2015.11.020>
13. Kucherenko T., Beskow J., Kjellström H. A neural network approach to missing marker reconstruction in human motion capture. *arXiv*, 2018, arXiv:1803.02665. <https://doi.org/10.48550/arXiv.1803.02665>
14. Geigel J., Schweppe M. Motion capture for realtime control of virtual actors in live, distributed, theatrical performances. *Proc. of the 2011*

14. Geigel J., Schweppe M. Motion capture for realtime control of virtual actors in live, distributed, theatrical performances // Proc. of the 2011 IEEE International Conference on Automatic Face & Gesture Recognition (FG). 2011. P. 774–779. <https://doi.org/10.1109/FG.2011.5771347>
15. Degen R., Tauber A., Nüßgen A., Irmer M., Klein F., Schyr C., Leijon M., Ruschitzka M. Methodical approach to integrate human movement diversity in real-time into a virtual test field for highly automated vehicle systems // Journal of Transportation Technologies. 2022. V. 12. N 3. P. 296–309. <https://doi.org/10.4236/jtts.2022.123018>
16. Sers R., Forrester S., Moss E., Ward S., Ma J., Zecca M. Validity of the Perception Neuron inertial motion capture system for upper body motion analysis // Measurement. 2020. V. 149. P. 107024. <https://doi.org/10.1016/j.measurement.2019.107024>
17. Sers R., Forrester S., Moss E., Ward S., Ma J., Zecca M. Validity of the Perception Neuron inertial motion capture system for upper body motion analysis // *IEEE International Conference on Automatic Face & Gesture Recognition (FG)*, 2011, pp. 774–779. <https://doi.org/10.1109/FG.2011.5771347>
18. Degen R., Tauber A., Nüßgen A., Irmer M., Klein F., Schyr C., Leijon M., Ruschitzka M. Methodical approach to integrate human movement diversity in real-time into a virtual test field for highly automated vehicle systems. *Journal of Transportation Technologies*, 2022, vol. 12, no. 3, pp. 296–309. <https://doi.org/10.4236/jtts.2022.123018>
19. Sers R., Forrester S., Moss E., Ward S., Ma J., Zecca M. Validity of the Perception Neuron inertial motion capture system for upper body motion analysis. *Measurement*, 2020, vol. 149, pp. 107024. <https://doi.org/10.1016/j.measurement.2019.107024>

#### Авторы

**Обухов Артём Дмитриевич** — доктор технических наук, доцент, Тамбовский государственный технический университет, Тамбов, 392000, Российская Федерация, [sc 56104232400](https://orcid.org/0000-0002-3450-5213), <https://orcid.org/0000-0002-3450-5213>, [obuhov.art@gmail.com](mailto:obuhov.art@gmail.com)

**Теселкин Даниил Вячеславович** — студент, Тамбовский государственный технический университет, Тамбов, 392000, Российская Федерация, [sc 57362498400](https://orcid.org/0000-0002-1304-9490), <https://orcid.org/0000-0002-1304-9490>, [teselkin.dv@tstu.ru](mailto:teselkin.dv@tstu.ru)

Статья поступила в редакцию 07.02.2023  
Одобрена после рецензирования 05.06.2023  
Принята к печати 18.07.2023

#### Authors

**Artem D. Obukhov** — D.Sc., Associate Professor, Tambov State Technical University, Tambov, 392000, Russian Federation, [sc 56104232400](https://orcid.org/0000-0002-3450-5213), <https://orcid.org/0000-0002-3450-5213>, [obuhov.art@gmail.com](mailto:obuhov.art@gmail.com)

**Daniil V. Teselkin** — Student, Tambov State Technical University, Tambov, 392000, Russian Federation, [sc 57362498400](https://orcid.org/0000-0002-1304-9490), <https://orcid.org/0000-0002-1304-9490>, [teselkin.dv@tstu.ru](mailto:teselkin.dv@tstu.ru)

Received 07.02.2023  
Approved after reviewing 05.06.2023  
Accepted 18.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ  
MODELING AND SIMULATION

doi: 10.17586/2226-1494-2023-23-4-795-802

УДК 681.787

Погрешности алгоритма демодуляции с генерируемой фазой несущей,  
вносимые фильтром низкой частотыГеоргий Петрович Мирошниченко<sup>1</sup>, Алина Николаевна Аржаненкова<sup>2</sup>✉,  
Михаил Юрьевич Плотников<sup>3</sup><sup>1,2,3</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация<sup>1</sup> [gpmirosh@gmail.com](mailto:gpmirosh@gmail.com), <https://orcid.org/0000-0002-4265-8818><sup>2</sup> [11arzh11@gmail.com](mailto:11arzh11@gmail.com)✉, <https://orcid.org/0000-0003-4869-2838><sup>3</sup> [plotnikov-michael@yandex.ru](mailto:plotnikov-michael@yandex.ru), <https://orcid.org/0000-0003-2506-0379>**Аннотация**

**Введение.** Представлены результаты исследования погрешности метода гомодинной демодуляции на основе решений функции арктангенса (PGC-ATAN). Метод связан с использованием фильтра низких частот в данном алгоритме демодуляции фазы сигнала. **Метод.** Исследован метод демодуляции интерференционного сигнала методом PGC-ATAN. Обнаружены и рассмотрены погрешности на этапе фильтрации методом скользящего среднего. Аналитически рассчитаны поправки к сигналу с учетом погрешности, вносимой фильтром низких частот. **Основные результаты.** Выведены формулы для расчета поправок к сигналам  $S_1(t)$ ,  $S_2(t)$ ,  $S_3(t)$ ,  $S_4(t)$  полученным с помощью фильтрации исходного сигнала, умноженного на сигнал опорного генератора. Выполнено сравнение результатов расчетов с данными математического моделирования обработки интерференционного сигнала методом PGC-ATAN. Демодуляция сигнала с учетом поправок показала, что влияние на фазу сигнала оказывается небольшим при невысокой скорости нагрева. Замечено, что для высокоскоростных процессов погрешность может приводить к серьезным искажениям искомого фазы сигнала. **Обсуждение.** Рассчитанные поправки к обрабатываемому интерференционному сигналу позволят улучшить метод демодуляции на основе вычислений функции арктангенса и точнее рассчитывать искомую фазу сигнала.

**Ключевые слова**

PGC, интерферометрия, погрешности

**Благодарности**

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, госзадание № 2019-0923.

**Ссылка для цитирования:** Мирошниченко Г.П., Аржаненкова А.Н., Плотников М.Ю. Погрешности алгоритма демодуляции с генерируемой фазой несущей, вносимые фильтром низкой частоты // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 795–802. doi: 10.17586/2226-1494-2023-23-4-795-802

Errors in the demodulation algorithm with a generated carrier phase introduced  
by the low-pass filterGeorge P. Miroshnichenko<sup>1</sup>, Alina N. Arzhanenkova<sup>2</sup>✉, Michael Yu. Plotnikov<sup>3</sup><sup>1,2,3</sup> ITMO University, Saint Petersburg, 197101, Russian Federation<sup>1</sup> [gpmirosh@gmail.com](mailto:gpmirosh@gmail.com), <https://orcid.org/0000-0002-4265-8818><sup>2</sup> [11arzh11@gmail.com](mailto:11arzh11@gmail.com)✉, <https://orcid.org/0000-0003-4869-2838><sup>3</sup> [plotnikov-michael@yandex.ru](mailto:plotnikov-michael@yandex.ru), <https://orcid.org/0000-0003-2506-0379>**Abstract**

In this paper, we study the errors of the homodyne demodulation method based on arctangent function solutions (PGC-ATAN) which are associated with the use of a low-pass filter (LPF) in this signal phase demodulation algorithm. The

© Мирошниченко Г.П., Аржаненкова А.Н., Плотников М.Ю., 2023

method of demodulation of an interference signal by PGC-ATAN method is investigated in order to detect and consider in more detail the errors at the filtering stage (the article considers the moving average method), and corrections to the signal are analytically calculated, taking into account the error introduced by the low-pass filter. We obtained formulas for calculating corrections to the signals  $S_1(t)$ ,  $S_2(t)$ ,  $S_3(t)$ ,  $S_4(t)$  which received by filtering the original signal multiplied by the reference oscillator signal, the calculations were compared with the results of mathematical modeling of the interference signal processing by the PGC-ATAN method. The demodulation of the signal, taking into account the corrections, showed that, in general, the effect on the signal phase is small at a low heating rate, however, for high-speed processes, the error can lead to serious distortions of the desired signal phase. These calculated corrections for processed interference signal will improve the demodulation method based on the calculations of the arc tangent function and more accurately calculate the desired phase of the signal.

### Keywords

PGC, interferometry, measurement errors

### Acknowledgements

This work was supported by the Ministry of Science and Higher Education of Russian Federation, State Assignment no. 2019-0923.

**For citation:** Miroshnichenko G.P., Arzhanenkova A.N., Plotnikov M.Yu. Errors in the demodulation algorithm with a generated carrier phase introduced by the low-pass filter. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 795–802 (in Russian). doi: 10.17586/2226-1494-2023-23-4-795-802

## Введение

Волоконно-оптические интерферометрические датчики обладают рядом преимуществ благодаря высокой чувствительности, низкой стоимости и невосприимчивости к электромагнитным помехам. Данные датчики успешно применяются во многих областях, таких как вибромониторинг различных объектов, подводное акустическое зондирование, деформационный мониторинг объектов, измерение скорости прецессии гироскопов, некоторые химические или биомедицинские, а также температурные измерения. Для восстановления акустического сигнала с выхода интерференционного датчика используют алгоритм демодуляции с генерируемой фазой несущей (PGC). Этот алгоритм активно используется из-за его широкого динамического диапазона, хорошей линейности фазовой характеристики и высокой чувствительности.

Запишем сигнал с фотодиода  $S_{in}(t)$ , поступающий на схему гомодинного детектирования в виде:

$$S_{in}(t) = I_0(1 + n_I(t))(1 + V\cos(C\cos(\omega_c t) + \varphi(t) + n_p(t))) + n_c(t),$$

где  $n_p(t)$  — фазовый шум, возникающий из-за частотного шума лазерного источника;  $n_c(t)$  — шум электронной схемы (фотоэлектрические преобразования, усилители и схемы дискретизации), который можно смоделировать как аддитивный белый гауссовский шум;  $n_I(t)$  — шум интенсивности света, исходящий от лазерного источника, модулятора, оптических усилителей;  $I_0$  — средняя интенсивность интерференционного сигнала;  $V$  — параметр видности;  $C$  — глубина модуляции (амплитуда несущей);  $\omega_c = 2\pi\nu_c$  — частота модуляции;  $\nu_c$  — частота колебаний;  $\varphi(t)$  — фаза, зависящая от изучаемого сигнала.

Алгоритм PGC активно изучается, исследуются границы его применимости, рассчитываются погрешности и поправки к методу для того, чтобы качественнее и эффективнее выделять искомую фазу сигнала.

Цель работы — поиск погрешностей, вносимых фильтром низких частот (ФНЧ) в сигнал, представляющий собой умножение интерференционного сигнала и сигнала с опорного генератора.

## Обзор предыдущих исследований

Существует несколько вариантов алгоритма PGC. Все алгоритмы начинаются с процесса гомодинной демодуляции, основанной на умножении сигнала  $S_{in}(t)$  на сигнал опорного генератора, который может иметь гармоники частоты модуляции от первой до четвертой в зависимости от схемы, и последующей фильтрации результата с помощью фильтра низкой частоты. После фильтрации получают низкочастотные сигналы, обозначаемые  $S_k(t)$ , где  $k = 1, 2, 3, 4$  — номер умножаемой гармоники. Сигналы  $S_k(t)$  зависят от измеряемой фазы  $\varphi(t)$ , которую необходимо извлечь с помощью последующей обработки.

Отметим основные алгоритмы обработки сигналов. В работах [1–5] предложен метод дифференцированного перекрестного умножения (differentiate-cross-multiply, PGC-DCM), основанный на дифференцировании сигналов  $S_1(t)$  и  $S_2(t)$ , их перекрестном перемножении и интегрировании. В результате получен сигнал, пропорциональный изучаемой фазе  $\varphi(t)$ . Эта схема позволяет восстанавливать измеренные фазовые сигналы из сигнала помехи независимо от рабочей точки интерферометра, однако на нее влияет эффект возмущения интенсивности света (Light Intensity Disturbance, LID). В работах [6–8] предложен метод демодуляции, использующий метод PGC-ATAN, который устраняет влияние LID, однако результат демодуляции строго зависит от глубины модуляции  $C$ . Если глубина демодуляции  $C$  отклоняется от 2,63 рад, результаты демодуляции будут иметь серьезные гармонические искажения. В работах [7–14] сформулирован улучшенный алгоритм PGC, основанный на вычислении функции арктангенса и дифференциальном умножении (DSM). Для схемы демодуляции алгоритма PGC на основе подходов PGC-DCM [9] и PGC-ATAN разработаны гибридные алгоритмы [15–26], такие как PGC-RCM [19], улучшенный PGC [20], PGC-DCM-ATAN [21], PGC-ARCCOT-DSM [22], PGC-ATAN-DSM [23], PGC-DSMI [24], PGC-Elim-B [25], PGC-Elim-BC [26]. В работе [12] исследован алгоритм, основанный на вычислении четырех сигналов  $S_1(t)$ ,  $S_2(t)$ ,  $S_3(t)$ ,  $S_4(t)$ , получаемых с помощью четырех гармоник частоты

модуляции. Этот алгоритм устойчив к флуктуациям интенсивности лазера, глубины модуляции, положения рабочей точки и имеет большой динамический диапазон. В перечисленных работах исследованы шумовые характеристики демодулированной фазы  $\varphi(t)$  и отмечено, что ФНЧ вносит погрешность в измеряемую фазу. Детальное исследование погрешности от ФНЧ исследовано в работе [6]. Тем не менее, в указанных работах не исследованы погрешности, вносимые ФНЧ в демодулированные сигналы  $S_1(t), \dots, S_4(t)$ .

В настоящей работе использован метод PGC-ATAN. Пренебрежем шумовыми характеристиками в сигнале  $S_{in}(t)$ , который запишем в виде:

$$S_{in}(t) = A + B \cos(C \cos(\omega_c t) + \varphi(t)), \quad (1)$$

где  $A$  и  $B$  — константы, определяемые интенсивностью и видностью оптического излучения.

Исследуем поправки к полученным сигналам  $S_1(t), \dots, S_4(t)$  после их обработки ФНЧ.

На рис. 1 изображена схема гомодинной демодуляции на основе решений функции арктангенса (метод PGC-ATAN) [10].

### ФНЧ – скользящее среднее

Исчерпывающими характеристиками любого линейного фильтра являются в операторном пространстве — передаточная функция, во временном пространстве — импульсная характеристика. Используем в качестве ФНЧ нерекурсивный линейный дискретный фильтр. Запишем уравнение фильтрации, имеющее вид свертки сигнала с импульсной характеристикой:

$$y_k = \sum_{j=0}^{P-1} x_{k-j} b_j, \quad (2)$$

где  $y_k$  —  $k$ -ое значение выходного сигнала;  $x_j$  —  $j$ -ое значение входного сигнала;  $b_j$  — отсчеты импульсной характеристики фильтра;  $P$  — порядок фильтра. Исследуем простой фильтр скользящего среднего, для которого импульсная характеристика равна

$$b_j = \frac{1}{P}, j = 0, 1, \dots, P-1.$$

После преобразования запишем уравнение фильтрации для фильтра скользящего среднего

$$y_k = \frac{1}{P} \sum_{j=k-P+1}^k x_j$$

При  $k = n + P - 1$

$$y_n = \frac{1}{P} \sum_{j=n}^{n+P-1} x_j, \quad (3)$$

где  $P$  — интервал усреднения (порядок фильтра);  $\omega_d$  — частота дискретизации;  $\omega_c$  — частота модуляции:

$$P = \frac{\omega_d}{\omega_c}. \quad (4)$$

Найдем амплитудно-частотную характеристику  $K(\omega)$  фильтра (3)

$$K(\omega) = \left| \frac{1}{P} \sum_{k=n}^{n+P-1} \exp(-i\omega T_d k) \right| = \frac{\sin\left(\pi \frac{\omega}{\omega_c}\right)}{\sin\left(\pi \frac{\omega}{\omega_d}\right)}, \quad (5)$$

где  $T_d$  — период дискретизации.

Детектируемый сигнал алгоритма PGC (1) при постоянной фазе изменяется во времени на частотах, кратных частоте модуляции  $\omega_c$ . Для правильной работы

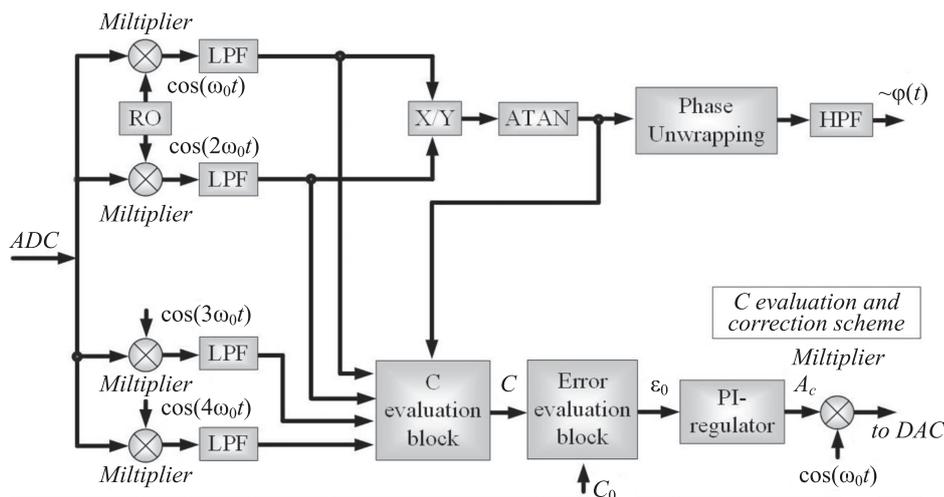


Рис. 1. Схема гомодинной демодуляции метода PGC-ATAN: ADC — аналого-цифровой преобразователь; RO — опорный генератор; LPF — фильтр низких частот; X/Y — операция деления; ATAN — операция вычисления значений функции арктангенса; Phase unwrapping — блок доворачивания фазы; HPF — фильтр высоких частот; C evaluation block — блок вычисления глубины модуляции; Error evaluation block — блок расчета ошибок; PI-regulator — схема обратной связи с интегратором; DAC — цифро-аналоговый преобразователь

Fig. 2. PGC-ATAN demodulation scheme: ADC — analog-to-digital converter; RO — reference oscillator; LPF — low pass filter; X/Y — division block; ATAN — arctangent function block; Phase unwrapping block; HPF — high pass filter; C — evaluation block and Error evaluation block; PI-regulator — integrator feedback scheme; DAC — digital-to-analog converter

алгоритма необходимо, чтобы эти частоты являлись нулями  $K(\omega)$ . Именно этим свойством обладает амплитудно-частотная характеристика ФНЧ-скользящее среднее (5). Недостатком этого фильтра, как и других дискретных фильтров, является ограничение диапазона фильтрации по частоте: частота сигнала не должна превосходить частоту Котельникова, равной половине  $\omega_d$ .

Иначе говоря,

$$\frac{\omega}{\omega_c} < \frac{P}{2}. \quad (6)$$

Область изменения частот в сигнале (1) зависит от параметра глубины модуляции  $C$ . Обычно выбирают оптимальную глубину модуляции:

$$C \approx 2,6.$$

Тогда в сигнале (1) можно ограничиться первыми пятью гармониками, и, как следствие, выбрать глубину фильтрации  $P \geq 20$ . Запишем формулу (3) в непрерывной форме:

$$\begin{aligned} T_d &= dt, T_{dj} = t \\ y(t) &= \frac{1}{T_c} \int_{t-T_c}^{t+T_c} x(t') dt', \end{aligned} \quad (7)$$

где  $T_c$  — период модуляции.

В этом случае амплитудно-частотная характеристика фильтра примет вид

$$K(\omega) = \left| \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \exp(-i\omega t') dt' \right| = \frac{\sin\left(\pi \frac{\omega}{\omega_c}\right)}{\pi \frac{\omega}{\omega_d}}. \quad (8)$$

Характеристики (5) и (8) совпадают, при условии (6). Для расчетов уравнение фильтра (7) в непрерывной форме удобнее, чем дискретное уравнение (3).

### Расчет поправок к сигналам

В блоке гомодинной демодуляции выполняется умножение сигнала  $S_{in}(t)$  (1) на гармоники частоты  $\omega_c$  задающего генератора с номерами  $r = 1, 2, 3, 4$ . Полученное произведение пропускается через ФНЧ, в качестве которого будем использовать фильтр скользящего среднего, функционирующий, согласно соотношению (7). В результате выделяются искомые низкочастотные сигналы  $S_r(t)$

$$S_r(t) = \frac{1}{T_c} \int_{t-T_c}^{t+T_c} S_{in}(t') \cos(r\omega_c t') dt'.$$

Подставим (1), получим

$$S_r(t) = S_r^{(c)}(t) + S_r^{(s)}(t),$$

где

$$S_r^{(c)}(t) = B \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \cos(C \cos(\omega_c t')) \cos(\varphi(t')) \cos(r\omega_c t') dt'$$

$$S_r^{(s)}(t) = -B \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \sin(C \cos(\omega_c t')) \sin(\varphi(t')) \cos(r\omega_c t') dt'$$

В формуле (1) параметры  $A$  и  $B$  полагаются не зависящими от времени, интеграл  $A \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \cos(r\omega_c t') dt'$  равен нулю (свойство (8) фильтра скользящего среднего).

Перепишем сигналы  $S_r^{(c)}(t)$  и  $S_r^{(s)}(t)$  с помощью формулы Якоби–Ангера [27].

$$\cos(C \cos(\omega_c t')) = J_0(C) + 2 \sum_{k=1}^{\infty} (-1)^k J_{2k}(C) \cos(2k\omega_c t'),$$

$$\sin(C \cos(\omega_c t')) = 2 \sum_{k=0}^{\infty} (-1)^k J_{2k+1}(C) \cos((2k+1)\omega_c t').$$

Тогда

$$S_r^{(c)}(t) = B(J_0(C)A_{0,r}(t) + 2 \sum_{k=1}^{\infty} (-1)^k J_{2k}(C)A_{2k,r}(t)),$$

$$S_r^{(s)}(t) = -2B \sum_{k=0}^{\infty} (-1)^k J_{2k+1}(C)B_{2k+1,r}(t),$$

где

$$\begin{aligned} A_{2k,r}(t) &= \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \cos(2k\omega_c t') \cos(\varphi(t')) \cos(r\omega_c t') dt'; \\ B_{2k+1,r}(t) &= \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \cos((2k+1)\omega_c t') \sin(\varphi(t')) \cos(r\omega_c t') dt'. \end{aligned} \quad (9)$$

Поясним, по какой причине для анализа погрешностей фильтрации был выбран ФНЧ скользящего среднего, обладающий свойством (8). Рассмотрим интеграл  $B_{2k+1,r}(t)$  при условии независимости искомой фазы  $\varphi(t)$  от времени. Имеем

$$B_{2k+1,r}(t) = \sin(\varphi) \frac{1}{T_c} \int_{t-T_c}^{t+T_c} \cos((2k+1)\omega_c t') \cos(r\omega_c t') dt'.$$

Из формулы (8) следует, что при условии  $r = 2k + 1$  получим  $B_{2k+1,r}(t) = \frac{1}{2} \sin(\varphi)$ , в противном случае этот интеграл равен нулю. Другими словами, ФНЧ скользящего среднего идеально отфильтровывает все гармоники задающего генератора и правильно выделяет низкочастотную часть сигнала (при условии, что фаза не зависит от времени). Если фаза меняется во времени, то фильтрация вносит погрешность в искомый низкочастотный сигнал. Поправки к  $S_1(t), \dots, S_4(t)$  найдем при условии

$$\xi = \frac{\dot{\varphi}(t')}{\omega_c} \ll 1, t \leq t' \leq t + T_c,$$

где  $\xi$  — параметр малости.

При этом условии можно разложить  $\varphi(t')$  в ряд Тейлора в интегралах (9):

$$\varphi(t') = \varphi(t) + \dot{\varphi}(t)(t' - t).$$

Ограничимся вычислением поправок первого порядка по параметру  $\xi$ :

$$S_r(t) = S_r^{(0)}(t) + \xi \Delta S_r(t). \quad (10)$$

Найдем интегралы  $A_{2k+1,r}(t), B_{2k+1,r}(t)$  с точностью до первого порядка по параметру  $\xi$ .

Для  $A_{2k,r}(t)$  получим

$$A_{2k,r}(t) = \delta_{2k,r}A_{2k,r}^{(0)}(t) + \xi A_{2k,r}^{(1)}(t),$$

$$A_{2k,2k}^{(0)}(t) = \frac{\cos(\varphi(t))}{2},$$

$$A_{2k,2k}^{(1)}(t) = -\frac{\sin(\varphi(t)) \sin(4k\omega_c t)}{2 \cdot 4k},$$

$$A_{2k,r}^{(1)}(t) = -\frac{\sin(\varphi(t))}{2} \left( \frac{\sin((2k+r)\omega_c t)}{2k+r} + \frac{\sin((2k-r)\omega_c t)}{2k-r} \right),$$

при условии, что  $r \neq 2k$ .

Для  $A_{2k+1,r}(t)$  получим

$$B_{2k+1,r}(t) = \delta_{2k+1,r}B_{2k+1,r}^{(0)}(t) + \xi B_{2k+1,r}^{(1)}(t),$$

$$B_{2k+1,2k+1}^{(0)}(t) = \frac{\sin(\varphi(t))}{2},$$

$$B_{2k+1,2k+1}^{(1)}(t) = \frac{\cos(\varphi(t)) \sin((4k+2)\omega_c t)}{2 \cdot 4k+2},$$

$$B_{2k+1,r}^{(1)}(t) = \frac{\cos(\varphi(t))}{2} \left( \frac{\sin((2k+1-r)\omega_c t)}{2k+1-r} + \frac{\sin((2k+1+r)\omega_c t)}{2k+1+r} \right),$$

при условии, что  $r \neq 2k+1$ , где  $\delta_{n,m}$  — символ Кронекера.

Рассчитаем нулевой порядок для  $S_r(t)$ :

$$S_1^{(0)}(t) = -2BJ_1(C)B_{1,1}^{(0)}(t) = -BJ_1(C)\sin(\varphi(t)),$$

$$S_2^{(0)}(t) = -2BJ_2(C)A_{2,2}^{(0)}(t) = -BJ_2(C)\cos(\varphi(t)),$$

$$S_3^{(0)}(t) = 2BJ_3(C)B_{3,3}^{(0)}(t) = BJ_3(C)\sin(\varphi(t)),$$

$$S_4^{(0)}(t) = 2BJ_4(C)A_{4,4}^{(0)}(t) = BJ_4(C)\cos(\varphi(t)),$$

где  $J_1, J_2, J_3, J_4$  — функции Бесселя первого рода 1, 2, 3 и 4 порядков соответственно.

Поправка первого порядка для  $S_r(t)$  (10) имеет вид

$$\Delta S_r(t) = B(J_0(C)A_{0,r}^{(1)}(t) + 2 \sum_{k=1}^{\infty} (-1)^k J_{2k}(C)A_{2k,r}^{(1)}(t) - 2 \sum_{k=0}^{\infty} (-1)^k J_{2k+1}(C)B_{2k+1,r}^{(1)}(t)).$$

где  $J_0$  — функция Бесселя нулевого порядка;  $J_{2k}$  и  $J_{2k+1}$  — функции Бесселя  $2k$  и  $2k+1$  порядков.

### Применение метода PGC-ATAN для вычисления фазы $\varphi(t)$

Кратко опишем протокол действия метода PGC-ATAN, используемый для численного моделирования:

1. Зададим сигнал с детектора (1), с зависящей от времени фазой  $\varphi(t)$  и частотой модуляции  $\nu_c = 2 \cdot 10^4$  Гц. Рассмотрим два случая: медленное изменение фазы

$$\varphi(t) = \pi \cos(0,05\nu_c t), \quad \xi_{\max} \approx 0,025 \quad (11)$$

и быстрое изменение фазы

$$\varphi(t) = 16\pi \cos(0,05\nu_c t), \quad \xi_{\max} \approx 0,4. \quad (12)$$

2. Умножим сигнал (1) на  $\cos(r\omega_c t)$ ,  $r = 1, 2, 3, 4$  и дискретизируем результат с частотой дискретизации  $\nu_d = 2 \cdot 10^6$  Гц.

3. Пропустим полученные векторы через ФНЧ-скользящее среднее (2) с порядком фильтрации  $P = 100$  (4). Получим четыре дискретизированных сигнала  $S_1(t), \dots, S_4(t)$ .

4. Применим процедуру арктангенса и вычислим искомую фазу:

$$\tilde{\varphi}(t) = \arctan \left( \sqrt{\frac{3 S_3(t)(S_3(t) - S_1(t))}{2 S_2(t)(S_2(t) - S_4(t))}} \right). \quad (13)$$

Равенство (13) справедливо благодаря известным свойствам функций Бесселя целого порядка.

Найденная таким образом фаза  $\tilde{\varphi}(t)$  будет отличаться от истинной фазы  $\varphi(t)$  из-за наличия поправок фильтрации. Найдем поправки первого порядка к фазе через поправку  $\Delta(t)$  к выражению

$$F(t) = \frac{3 S_3(t)(S_3(t) - S_1(t))}{2 S_2(t)(S_2(t) - S_4(t))} = \left( \frac{\sin(\varphi(t))}{\cos(\varphi(t))} \right)^2 (1 + \xi \Delta(t)). \quad (14)$$

Преобразуем выражение (14) с помощью формулы (10)

$$\Delta(t) = \frac{\Delta S_3(t)}{S_3^{(0)}(t)} + \frac{\Delta S_3(t) - \Delta S_1(t)}{S_3^{(0)}(t) - S_1^{(0)}(t)} - \frac{\Delta S_2(t)}{S_2^{(0)}(t)} - \frac{\Delta S_2(t) - \Delta S_4(t)}{S_2^{(0)}(t) - S_4^{(0)}(t)}.$$

Получим формулу для искаженной фазы с точностью первого порядка

$$\tilde{\varphi}(t) = \arctan(\sqrt{F(t)}) = \arctan \left( \frac{\sin(\varphi(t))}{\cos(\varphi(t))} \left( 1 + \frac{1}{2} \xi \Delta(t) \right) \right). \quad (15)$$

Вычисление арктангенса (15) даст искомую фазу с поправкой:

$$\tilde{\varphi}(t) = \varphi(t) + \xi \varphi^{(1)}(t). \quad (16)$$

В итоге запишем

$$\varphi^{(1)}(t) = \frac{1}{2} \cos(\varphi(t)) \left( \left( \frac{1}{J_3(C)} + \frac{C}{4J_2(C)} \right) \Delta S_3(t) - \frac{C}{4J_2(C)} \Delta S_1(t) \right) + \frac{1}{2} \sin(\varphi(t)) \left( \left( \frac{1}{J_2(C)} + \frac{C}{6J_3(C)} \right) \Delta S_2(t) - \frac{C}{6J_3(C)} \Delta S_4(t) \right). \quad (17)$$

### Результаты

Результаты численного моделирования по методу PGC-ATAN представлены на рис. 2 и рис. 3. Полученные графики погрешностей демодуляции фазы, рассчитаны с помощью математического моделирования (экспериментальная погрешность  $\delta_{\text{exp}}(t) = \tilde{\varphi}(t) - \varphi(t)$ ) и по формулам (16) и (17) (теоретическая погрешность  $\delta_{\text{th}}(t) = \xi(t)\varphi^{(1)}(t)$ ). На рис. 2 приведено сравнение погрешностей  $\delta_{\text{exp}}(t)$  и  $\delta_{\text{th}}(t)$  для медленного из-

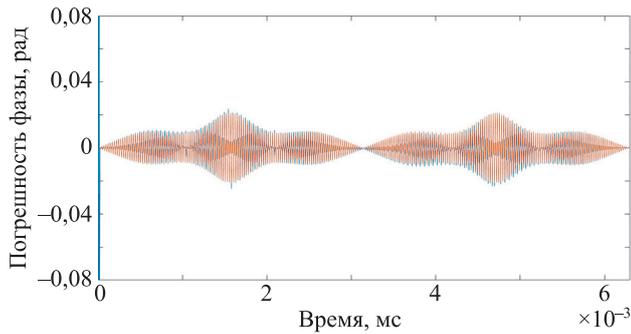


Рис. 2. Рассчитанная аналитическим методом погрешность  $\delta_{exp}(t)$  (синий цвет) и численный эксперимент  $\delta_{th}(t)$  (красный цвет) для небольшой скорости нагрева (11)

Fig. 2. Analytical calculated error  $\delta_{exp}(t)$  (in blue) and numerical experiment  $\delta_{th}(t)$  (in red) for a low heating rate (11)

менения фазы по формуле (11), на рис. 3 —  $\delta_{exp}(t)$  и  $\delta_{th}(t)$  для быстрого изменения фазы по формуле (12). Для периодического изменения фазы и ошибка изменяются периодически. На рисунках представлено несколько периодов. Ошибка велика в области максимальной скорости изменения фазы и мала в области минимальной скорости.

### Заключение

Исследованы погрешности в сигнале, создаваемые фильтром низких частот скользящего среднего. В рамках предыдущих исследований ошибки в сигналах  $S_1(t)$ ,  $S_2(t)$ ,  $S_3(t)$ ,  $S_4(t)$  из-за фильтра низких частот исследованы не были. Фильтр низких частот скользящего

### Литература

1. Dandridge A., Tveten A.B., Giallorenzi T.G. Homodyne demodulation scheme for fiber optic sensors using phase generated carrier // *IEEE Journal of Quantum Electronics*. 1982. V. 18. N 10. P. 1647–1653. <https://doi.org/10.1109/jqe.1982.1071416>
2. McGarrity C., Jackson D. Improvement on phase generated carrier technique for passive demodulation of miniature interferometric sensors // *Optics Communications*. 1994. V. 109. N 3-4. P. 246–248. [https://doi.org/10.1016/0030-4018\(94\)90687-4](https://doi.org/10.1016/0030-4018(94)90687-4)
3. Kersey A.D. A Review of recent developments in fiber optic sensor technology // *Optical Fiber Technology*. 1996. V. 2. N 3. P. 291–317. <https://doi.org/10.1006/ofte.1996.0036>
4. Wu K., Min Z., Liao Y. Signal dependence of the phase-generated carrier method // *Optical Engineering*. 2007. V. 46. N 10. P. 105602. <https://doi.org/10.1117/1.2799518>
5. Lin W.-W., Huang S.-C., Chen M.-H. Fiber optic microampere dc current sensor // *Optical Engineering*. 2003. V. 42. N 9. P. 2551–2557. <https://doi.org/10.1117/1.1597675>
6. Feng L., He J., Duan J.-Y., Li F., Liu Y.-L. Implementation of phase generated carrier technique for FBG laser sensor multiplexed system based on compact RIO // *Proc. 1st Asia-Pacific Optical Fiber Sensors Conference, APOS*. 2008. <https://doi.org/10.1109/APOS.2008.5226295>
7. Liu Y., Wang L., Tian C., Zhang M., Liao Y. Analysis and optimization of the PGC method in all digital demodulation systems // *Journal of Lightwave Technology*. 2008. V. 26. N 18. P. 3225–3233. <https://doi.org/10.1109/jlt.2008.928926>
8. He J., Wang L., Li F., Liu Y. An ameliorated phase generated carrier demodulation algorithm with low harmonic distortion and high stability // *Journal of Lightwave Technology*. 2010. V. 28. N 22. P. 3258–3265. <https://doi.org/10.1109/jlt.2010.2081347>
9. Azmi A.I., Leung I., Chen X., Zhou S., Zhu Q., Gao K., Childs P., Peng G. Fiber laser based hydrophone systems // *Photonic Sensors*.

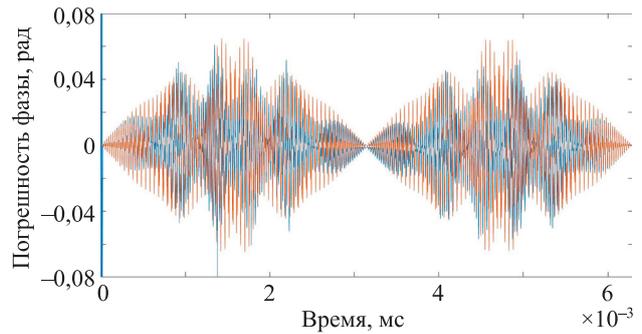


Рис. 3. Погрешности фазы интерференционного сигнала для быстрых процессов (12). Синим цветом показана теоретическая погрешность  $\delta_{th}(t)$ , красным — численный эксперимент  $\delta_{exp}(t)$

Fig. 3. Interference signal phase errors for fast processes (12). The theoretical error  $\delta_{th}(t)$  is shown in blue, numerical experiment  $\delta_{exp}(t)$  is shown in red

среднего был выбран неслучайно, он проще и идеально отфильтровывает все гармоники задающего генератора и правильно выделяет низкочастотную часть сигнала (при условии, что фаза не зависит от времени). Как следует из приведенного исследования, величина поправки к сигнальной фазе зависит от скорости ее изменения. Таким образом, динамический диапазон метода PGC-ATAN определяется не только величиной фазовых и амплитудных шумов, но и скоростью изменения сигнальной фазы. Полученные результаты позволяют улучшить качество метода PGC-ATAN и точнее определить его динамический диапазон.

### References

1. Dandridge A., Tveten A.B., Giallorenzi T.G. Homodyne demodulation scheme for fiber optic sensors using phase generated carrier. *IEEE Journal of Quantum Electronics*, 1982, vol. 18, no. 10, pp. 1647–1653. <https://doi.org/10.1109/jqe.1982.1071416>
2. McGarrity C., Jackson D. Improvement on phase generated carrier technique for passive demodulation of miniature interferometric sensors. *Optics Communications*, 1994, vol. 109, no. 3-4, pp. 246–248. [https://doi.org/10.1016/0030-4018\(94\)90687-4](https://doi.org/10.1016/0030-4018(94)90687-4)
3. Kersey A.D. A Review of recent developments in fiber optic sensor technology. *Optical Fiber Technology*, 1996, vol. 2, no. 3, pp. 291–317. <https://doi.org/10.1006/ofte.1996.0036>
4. Wu K., Min Z., Liao Y. Signal dependence of the phase-generated carrier method. *Optical Engineering*, 2007, vol. 46, no. 10, pp. 105602. <https://doi.org/10.1117/1.2799518>
5. Lin W.-W., Huang S.-C., Chen M.-H. Fiber optic microampere dc current sensor. *Optical Engineering*, 2003, vol. 42, no. 9, pp. 2551–2557. <https://doi.org/10.1117/1.1597675>
6. Feng L., He J., Duan J.-Y., Li F., Liu Y.-L. Implementation of phase generated carrier technique for FBG laser sensor multiplexed system based on compact RIO. *Proc. 1st Asia-Pacific Optical Fiber Sensors Conference, APOS*, 2008. <https://doi.org/10.1109/APOS.2008.5226295>
7. Liu Y., Wang L., Tian C., Zhang M., Liao Y. Analysis and optimization of the PGC method in all digital demodulation systems. *Journal of Lightwave Technology*, 2008, vol. 26, no. 18, pp. 3225–3233. <https://doi.org/10.1109/jlt.2008.928926>
8. He J., Wang L., Li F., Liu Y. An ameliorated phase generated carrier demodulation algorithm with low harmonic distortion and high stability. *Journal of Lightwave Technology*, 2010, vol. 28, no. 22, pp. 3258–3265. <https://doi.org/10.1109/jlt.2010.2081347>
9. Azmi A.I., Leung I., Chen X., Zhou S., Zhu Q., Gao K., Childs P., Peng G. Fiber laser based hydrophone systems. *Photonic Sensors*,

2011. V. 1. N 3. P. 210–221. <https://doi.org/10.1007/s13320-011-0018-3>
10. Plotnikov M.J., Kulikov A.V., Strigalev V.E., Meshkovsky I.K. Dynamic range analysis of the phase generated carrier demodulation technique // *Advances in Optical Technologies*. 2014. V. 2014. P. 815108. <https://doi.org/10.1155/2014/815108>
  11. Zhang A., Zhang S. High stability fiber-optics sensors with an improved PGC demodulation algorithm // *IEEE Sensors Journal*. 2016. V. 16. N 21. P. 7681–7684. <https://doi.org/10.1109/jсен.2016.2604348>
  12. Беликин М.Н., Плотников М.Ю., Стригалеv В.Е., Куликов А.В., Киреевков А.Ю. Экспериментальное сравнение алгоритмов гомодинной демодуляции сигналов для фазового волоконно-оптического датчика // *Научно-технический вестник информационных технологий, механики и оптики*. 2015. Т. 15. № 6. С. 1008–1014. <https://doi.org/10.17586/2226-1494-2015-15-6-1008-1014>
  13. Volkov A.V., Plotnikov M.Y., Mekhregin M.V., Miroshnichenko G.P., Aleynik A.S. Phase modulation depth evaluation and correction technique for the PGC demodulation scheme in fiber-optic interferometric sensors // *IEEE Sensors Journal*. 2017. V. 17. N 13. P. 4143–4150. <https://doi.org/10.1109/jсен.2017.2704287>
  14. Zhang Z., Zhang Z., Cheng J., Kong Y., Zhang L., Zhang D., Zhuang S., Peng W., Wu F., Zhu Y., Cui X. An improved phase generated carrier demodulation algorithm with high stability and low harmonic distortion // *Optics Communications*. 2022. V. 524. P. 128800. <https://doi.org/10.1016/j.optcom.2022.128800>
  15. Zhang S., Chen Y., Chen B., Yan L., Xie J., Lou Y. A PGC-DCDM demodulation scheme insensitive to phase modulation depth and carrier phase delay in an EOM-based SPM interferometer // *Optics Communications*. 2020. V. 474. P. 126183. <https://doi.org/10.1016/j.optcom.2020.126183>
  16. Li Q., Huang H., Lin F., Wu X. Optical micro-particle size detection by phase-generated carrier demodulation // *Optics Express*. 2016. V. 24. N 11. P. 11458–11465. <https://doi.org/10.1364/OE.24.011458>
  17. Yu Z., Dai H., Zhang M., Zhang J., Liu L., Jin X., Luo Y. High stability and low harmonic distortion PGC demodulation technique for interferometric optical fiber sensors // *Optics & Laser Technology*. 2019. V. 109. P. 8–13. <https://doi.org/10.1016/j.optlastec.2018.07.055>
  18. Christian T.R., Frank P.A., Houston B.H. Real-time analog and digital demodulator for interferometric fiber optic sensors // *Proceedings of SPIE*. 1994. V. 2191. P. 324–336. <https://doi.org/10.1117/12.173962>
  19. Zhang S., Zhang A., Pan H. Eliminating light intensity disturbance with reference compensation in interferometers // *IEEE Photonics Technology Letters*. 2015. V. 27. N 17. P. 1888–1891. <https://doi.org/10.1109/lpt.2015.2444421>
  20. Tong Y., Zeng H., Li L., Zhou Y. Improved phase generated carrier demodulation algorithm for eliminating light intensity disturbance and phase modulation amplitude variation // *Applied Optics*. 2012. V. 51. N 29. P. 6962–6967. <https://doi.org/10.1364/AO.51.006962>
  21. Yang X., Chen Z., Ng J.H., Pallayil V., Unnikrishnan C.K.C. A PGC demodulation based on differential-cross-multiplying (DCM) and arctangent (ATAN) algorithm with low harmonic distortion and high stability // *Proceedings of SPIE*. 2012. V. 8421. P. 84215J. <https://doi.org/10.1117/12.974939>
  22. Wang G.-Q., Xu T.-W., Li F. PGC demodulation technique with high stability and low harmonic distortion // *IEEE Photonics Technology Letters*. 2012. V. 24. N 23. P. 2093–2096. <https://doi.org/10.1109/lpt.2012.2220129>
  23. He J., Li F., Zhang W., Wang L., Xu T., Liu Y. High performance wavelength demodulator for DFB fiber laser sensor using novel PGC algorithm and reference compensation method // *Proceedings of SPIE*. 2011. V. 7753. P. 775333. <https://doi.org/10.1117/12.885823>
  24. Zhang W., Xia H., Pan C., Huang W., Li F. Differential self-multiplying-integrate phase generated carrier method for fiber optic sensors // *Proceedings of SPIE*. 2014. V. 9233. P. 92331U. <https://doi.org/10.1117/12.2069764>
  25. Peng F., Hou L., Yang J., Yuan Y., Li C., Yan D., Yuan L., Zheng H., Chang Z., Ma K., Yang J. An improved fixed phased demodulation method combined with phase generated carrier (PGC) and ellipse fitting algorithm // *Proceedings of SPIE*. 2015. V. 9620. P. 96200S. <https://doi.org/10.1117/12.2190842>
  26. Wu B., Yuan Y., Yang J., Liang S., Yuan L. Optimized phase generated carrier (PGC) demodulation algorithm insensitive to C value // *Proceedings of SPIE*. 2015. V. 9655. P. 96550C. <https://doi.org/10.1117/12.2184268>
  - 2011, vol. 1, no. 3, pp. 210–221. <https://doi.org/10.1007/s13320-011-0018-3>
  10. Plotnikov M.J., Kulikov A.V., Strigalev V.E., Meshkovsky I.K. Dynamic range analysis of the phase generated carrier demodulation technique. *Advances in Optical Technologies*, 2014, vol. 2014, pp. 815108. <https://doi.org/10.1155/2014/815108>
  11. Zhang A., Zhang S. High stability fiber-optics sensors with an improved PGC demodulation algorithm. *IEEE Sensors Journal*, 2016, vol. 16, no. 21, pp. 7681–7684. <https://doi.org/10.1109/jсен.2016.2604348>
  12. Belikin M.N., Plotnikov M.Yu., Strigalev V.E., Kulikov A.V., Kireenkov A.Yu. Experimental comparison of homodyne demodulation algorithms for phase fiber-optic sensor. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2015, vol. 15, no. 6, pp. 1008–1014. (in Russian). <https://doi.org/10.17586/2226-1494-2015-15-6-1008-1014>
  13. Volkov A.V., Plotnikov M.Y., Mekhregin M.V., Miroshnichenko G.P., Aleynik A.S. Phase modulation depth evaluation and correction technique for the PGC demodulation scheme in fiber-optic interferometric sensors. *IEEE Sensors Journal*, 2017, vol. 17, no. 13, pp. 4143–4150. <https://doi.org/10.1109/jсен.2017.2704287>
  14. Zhang Z., Zhang Z., Cheng J., Kong Y., Zhang L., Zhang D., Zhuang S., Peng W., Wu F., Zhu Y., Cui X. An improved phase generated carrier demodulation algorithm with high stability and low harmonic distortion. *Optics Communications*, 2022, vol. 524, pp. 128800. <https://doi.org/10.1016/j.optcom.2022.128800>
  15. Zhang S., Chen Y., Chen B., Yan L., Xie J., Lou Y. A PGC-DCDM demodulation scheme insensitive to phase modulation depth and carrier phase delay in an EOM-based SPM interferometer. *Optics Communications*, 2020, vol. 474, pp. 126183. <https://doi.org/10.1016/j.optcom.2020.126183>
  16. Li Q., Huang H., Lin F., Wu X. Optical micro-particle size detection by phase-generated carrier demodulation. *Optics Express*, 2016, vol. 24, no. 11, pp. 11458–11465. <https://doi.org/10.1364/OE.24.011458>
  17. Yu Z., Dai H., Zhang M., Zhang J., Liu L., Jin X., Luo Y. High stability and low harmonic distortion PGC demodulation technique for interferometric optical fiber sensors. *Optics & Laser Technology*, 2019, vol. 109, pp. 8–13. <https://doi.org/10.1016/j.optlastec.2018.07.055>
  18. Christian T.R., Frank P.A., Houston B.H. Real-time analog and digital demodulator for interferometric fiber optic sensors. *Proceedings of SPIE*, 1994, vol. 2191, pp. 324–336. <https://doi.org/10.1117/12.173962>
  19. Zhang S., Zhang A., Pan H. Eliminating light intensity disturbance with reference compensation in interferometers. *IEEE Photonics Technology Letters*, 2015, vol. 27, no. 17, pp. 1888–1891. <https://doi.org/10.1109/lpt.2015.2444421>
  20. Tong Y., Zeng H., Li L., Zhou Y. Improved phase generated carrier demodulation algorithm for eliminating light intensity disturbance and phase modulation amplitude variation. *Applied Optics*, 2012, vol. 51, no. 29, pp. 6962–6967. <https://doi.org/10.1364/AO.51.006962>
  21. Yang X., Chen Z., Ng J.H., Pallayil V., Unnikrishnan C.K.C. A PGC demodulation based on differential-cross-multiplying (DCM) and arctangent (ATAN) algorithm with low harmonic distortion and high stability. *Proceedings of SPIE*, 2012, vol. 8421, pp. 84215J. <https://doi.org/10.1117/12.974939>
  22. Wang G.-Q., Xu T.-W., Li F. PGC demodulation technique with high stability and low harmonic distortion. *IEEE Photonics Technology Letters*, 2012, vol. 24, no. 23, pp. 2093–2096. <https://doi.org/10.1109/lpt.2012.2220129>
  23. He J., Li F., Zhang W., Wang L., Xu T., Liu Y. High performance wavelength demodulator for DFB fiber laser sensor using novel PGC algorithm and reference compensation method. *Proceedings of SPIE*, 2011, vol. 7753, pp. 775333. <https://doi.org/10.1117/12.885823>
  24. Zhang W., Xia H., Pan C., Huang W., Li F. Differential self-multiplying-integrate phase generated carrier method for fiber optic sensors. *Proceedings of SPIE*, 2014, vol. 9233, pp. 92331U. <https://doi.org/10.1117/12.2069764>
  25. Peng F., Hou L., Yang J., Yuan Y., Li C., Yan D., Yuan L., Zheng H., Chang Z., Ma K., Yang J. An improved fixed phased demodulation method combined with phase generated carrier (PGC) and ellipse fitting algorithm. *Proceedings of SPIE*, 2015, vol. 9620, pp. 96200S. <https://doi.org/10.1117/12.2190842>
  26. Wu B., Yuan Y., Yang J., Liang S., Yuan L. Optimized phase generated carrier (PGC) demodulation algorithm insensitive to C value. *Proceedings of SPIE*, 2015, vol. 9655, pp. 96550C. <https://doi.org/10.1117/12.2184268>

27. Бейтмен Г., Эрдейи А. Высшие трансцендентные функции. Т. 2. Функции Бесселя, функции параболического цилиндра, ортогональные многочлены / 2-е изд., пер. с англ. Н. Я. Виленкина, М.: Наука, 1974. С. 15.

- Proceedings of SPIE*, 2015, vol. 9655, pp. 96550C. <https://doi.org/10.1117/12.2184268>  
27. Bateman H., Erdélyi A. Higher Transcendental Functions. Vol. 2. McGraw-Hill, 1954.

#### Авторы

**Мирошниченко Георгий Петрович** — доктор физико-математических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 55401299600](https://orcid.org/0000-0002-4265-8818), <https://orcid.org/0000-0002-4265-8818>, [gpmirosh@gmail.com](mailto:gpmirosh@gmail.com)  
**Аржаненкова Алина Николаевна** — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57674300400](https://orcid.org/0000-0003-4869-2838), <https://orcid.org/0000-0003-4869-2838>, [11arzh11@gmail.com](mailto:11arzh11@gmail.com)  
**Плотников Михаил Юрьевич** — кандидат технических наук, старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57193069973](https://orcid.org/0000-0003-2506-0379), <https://orcid.org/0000-0003-2506-0379>, [plotnikov-michael@yandex.ru](mailto:plotnikov-michael@yandex.ru)

*Статья поступила в редакцию 13.03.2023*  
*Одобрена после рецензирования 05.05.2023*  
*Принята к печати 18.07.2023*

#### Authors

**George P. Miroshnichenko** — D.Sc. (Physics & Mathematics), Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 55401299600](https://orcid.org/0000-0002-4265-8818), <https://orcid.org/0000-0002-4265-8818>, [gpmirosh@gmail.com](mailto:gpmirosh@gmail.com)  
**Alina N. Arzhanenkova** — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57674300400](https://orcid.org/0000-0003-4869-2838), <https://orcid.org/0000-0003-4869-2838>, [11arzh11@gmail.com](mailto:11arzh11@gmail.com)  
**Michael Yu. Plotnikov** — PhD, Senior Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57193069973](https://orcid.org/0000-0003-2506-0379), <https://orcid.org/0000-0003-2506-0379>, [plotnikov-michael@yandex.ru](mailto:plotnikov-michael@yandex.ru)

*Received 13.03.2023*  
*Approved after reviewing 05.05.2023*  
*Accepted 18.07.2023*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-803-811

УДК 629.7.036.621.373

## Моделирование процесса корректировки формы роторов шаровых гироскопов

Маргарита Алексеевна Тит<sup>1</sup>, Сергей Николаевич Беляев<sup>2</sup>,  
Александр Григорьевич Щербак<sup>3</sup>, Ольга Сергеевна Юльметова<sup>4</sup>✉

<sup>1,2,3,4</sup> АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург, 197046, Российская Федерация

<sup>4</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>1</sup> rita93.07.93.07@gmail.com, <https://orcid.org/0009-0003-0069-2940>

<sup>2</sup> serenible@mail.ru, <https://orcid.org/0000-0003-3765-9469>

<sup>3</sup> asherbak1946@gmail.com, <https://orcid.org/0000-0001-8507-7653>

<sup>4</sup> olga@yulmetova.ru✉, <https://orcid.org/0000-0001-9899-5939>

### Аннотация

**Введение.** Совершенствование технологии изготовления гироскопических приборов, являющихся автономными датчиками выработки параметров движения объектов, имеет стратегическое значение и является приоритетным направлением развития различных отраслей промышленности. Объектом исследования выступают роторы шаровых гироскопов, геометрические параметры которых определяют точностные характеристики прибора. В работе представлены результаты моделирования процесса корректировки формы ротора на стадии его изготовления в процессе напыления функционального покрытия. **Метод.** Предложенная математическая модель процесса напыления основана на размещении между ротором и источником подвижного экрана с отверстием. Ось экрана лежит на динамической оси ротора и обеспечивает нанесение покрытия на поверхность сферического сегмента. Целью для разработки математической модели послужила необходимость управления геометрией напыляемой поверхности. Управление можно выполнить с помощью экрана с отверстием при использовании точечного источника напыляемого материала. Перемещение экрана или ротора вдоль динамической оси ротора в процессе нанесения дополнительного слоя, а также вращение ротора вокруг динамической оси позволяет в широких пределах и с высокой точностью до сотых долей микрометра регулировать кривизну формируемого покрытия. **Основные результаты.** Разработана аналитическая модель технологического процесса корректировки формы сферических роторов шаровых гироскопов. Приведено математическое описание, управляющие факторы и значимые параметры процесса. Представлены результаты практической апробации разработанных моделей. **Обсуждение.** Представленная математическая модель позволяет осуществлять корректировку формы роторов при напылении функционального покрытия, расширяя технологические возможности процесса изготовления и обеспечивая повышение точности роторов.

### Ключевые слова

покрытие, корректировка формы, напыление, моделирование, сферический ротор

**Ссылка для цитирования:** Тит М.А., Беляев С.Н., Щербак А.Г., Юльметова О.С. Моделирование процесса корректировки формы роторов шаровых гироскопов // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 803–811. doi: 10.17586/2226-1494-2023-23-4-803-811

## Modeling of the process of spherical form correction for rotors of electrostatically suspended gyros

Margarita A. Tit<sup>1</sup>, Sergey N. Belyaev<sup>2</sup>, Alexandr G. Shcherbak<sup>3</sup>, Olga S. Yulmetova<sup>4</sup>✉

<sup>1,2,3,4</sup> Concern CSRI Elektropribor, JSC, Saint Petersburg, 197046, Russian Federation

<sup>4</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>1</sup> rita93.07.93.07@gmail.com, <https://orcid.org/0009-0003-0069-2940>

<sup>2</sup> serenible@mail.ru, <https://orcid.org/0000-0003-3765-9469>

<sup>3</sup> asherbak1946@gmail.com, <https://orcid.org/0000-0001-8507-7653>

<sup>4</sup> olga@yulmetova.ru✉, <https://orcid.org/0000-0001-9899-5939>

© Тит М.А., Беляев С.Н., Щербак А.Г., Юльметова О.С., 2023

**Abstract**

Improvement of the manufacturing technology for gyroscopic devices, which autonomously generate motion parameters of moving objects, has strategic importance and priority for various industries. The object of current research is a spherical rotor of an electrostatically suspended gyroscope which geometric parameters determine the accuracy characteristics of the device. The paper presents results of the process modeling of spherical form correction for rotors of electrostatically suspended gyroscopes at the stage of its manufacture during the coating deposition process. The proposed mathematical model of the deposition process is based on the placement of a movable screen with a hole between a rotor and a spray source. The axis of the hole lies on the dynamic axis of the rotor and it provides a formation of a spherical segment on the coating rotor surface. During deposition of an additional layer, the screen or rotor moves along the dynamic axis of the rotor changing the distance between the rotor and the screen, and there is additional rotation of the rotor around its dynamic axis. It allows adjusting the curvature of the formed coating on the rotor surface. An analytical model of the technological process for correcting the shape of spherical rotors of electrostatically suspended gyroscopes has been developed. A mathematical description, control factors and significant parameters of the process are given. The results of practical testing of the developed model are presented. The presented mathematical model makes it possible to correct the shape of the rotors during the deposition of a functional coating expanding the technological possibilities and increasing the accuracy of rotors.

**Keywords**

coating, shape correction, deposition, modeling, spherical rotor

**For citation:** Tit M.A., Belyaev S.N., Shcherbak A.G., Yulmetova O.S. Modeling of the process of spherical form correction for rotors of electrostatically suspended gyros. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 803–811 (in Russian). doi: 10.17586/2226-1494-2023-23-4-803-811

**Введение**

Точностные характеристики шаровых гироскопов [1, 2], таких как криогенный и бескарданный электростатический гироскоп (БЭСГ) [3], существенно зависят от технологического процесса изготовления сферического ротора — основного узла чувствительного элемента. Требования, предъявляемые к ротору по отклонению от сферичности и дисбалансу, составляют сотые доли микрометра, что создает предпосылки для разработки управляемой технологии [4], обеспечивающей возможность варьирования указанных параметров в заданных пределах.

Геометрические параметры заданы конечным диаметром ротора и некруглостью, контроль которой осуществляется на основе гармонического анализа круглограмм ротора в нескольких сечениях. Отклонение от круглости сферы ротора измеряется с помощью кругломера Talygond 73. Профиль регистрируется в виде полярного графика. Получаемая профилограмма конкретного сечения после компьютерной обработки дискретно раскладывается в ряд Фурье и представляется в виде совокупности гармоник

$$X_n = \sum_{k=0}^{N-1} R_k e^{-2\pi i n k / N},$$

где  $X_n$  — комплексная амплитуда  $n$ -ой гармоники;  $R_k$  — значение длины радиус-вектора на угловом отсчете  $2\pi k / N$ ;  $N$  — длина сигнала. Амплитудный спектр определяется как  $A_n = \frac{2X_n}{N}$ . Из полученных гармоник значимыми являются первые шесть, причем первая гармоника  $A_1$  характеризует эксцентриситет ротора относительно центра измерения, т. е. точность установки ротора. Следующие гармоники определяют в меридиональном сечении:  $A_2$  — эллиптичность,  $A_3$  — треугольную аномалию,  $A_4$  — квадратичность,  $A_5$  — звездобразную форму,  $A_6$  — грушевидность. К виду формы сферического ротора шарового гироскопа предъявляются требования к усредненным амплитудам гармоник

в четырех меридиональных сечениях и одном экваториальном.

Для сплошного ротора БЭСГ [3] паспортизируются амплитуды  $A$  пяти гармоник формы ( $A_2 \leq 0,05$  мкм;  $A_3 \leq 0,025$  мкм;  $A_4, A_5, A_6 \leq 0,015$  мкм). Ротор криогенного гироскопа должен иметь наружный диаметр  $10 \pm 0,005$  мм [3], средние значения амплитуд гармоник:  $A_2 \leq 0,02$  мкм,  $A_3, A_5 \leq 0,005$  мкм,  $A_4, A_6 \leq 0,01$  мкм.

Технологический цикл изготовления роторов предполагает формообразование заготовки методами механической обработки и последующее напыление тонкопленочных функциональных покрытий [5, 6] нитрида титана [7, 8] для БЭСГ и ниобия [9–13] для криогенного гироскопа. При необходимости, в зависимости от модификации прибора, на поверхности покрытия формируется светоконтрастный рисунок, обеспечивающий функционирование оптоэлектронной системы при получении информации [14–17].

При напылении покрытий, решающих задачи, в том числе и многоцелевые, формирования параметров ротора (дисбаланс и форма) оперируют величинами в сотые и тысячные доли микрометра, что недостижимо при других методах обработки. В связи с этим исследования, связанные с развитием и совершенствованием технологии нанесения функциональных тонкопленочных покрытий, являются важными и актуальными.

**Постановка задачи**

Формирование тонкопленочных покрытий осуществлено методом магнетронного распыления или конденсации ионной бомбардировкой с применением специальных средств технологического оснащения. Ориентация ротора в процессе напыления по отношению к потоку напыляемого материала определяет конечную форму поверхности. Схема фиксации ротора в приспособлении [7] предполагает крепление в четырех точках, соответствующих вершинам правильной трехгранной пирамиды. При этом с одной стороны ротора расположен один игольчатый упор, а с другой — три

упора, образующие трехконцевую вилку. Поскольку элементы фиксации размещены несимметрично относительно поверхности ротора, возможно возникновение локальных дефектов на уровне сотых долей микрометра, вызванных экранированием элементами фиксации.

Очевидно, что существенным фактором является величина отклонения от номинальной толщины покрытия  $\Delta h$  в зоне экранирования, определяемой сферическим сегментом в области размещения трехконцевой вилки и характеризуемая углом  $\theta_{\max}$  (рис. 1).

Фактически на поверхности ротора образуется незаполненный материалом покрытия фрагмент, который, рассматривая зону экранирования как поверхность сферического сегмента, имеющего исходный радиус  $R_c$  (рис. 1), ограничен двумя сферическими поверхностями —  $R_c$  и требуемого радиуса  $R$ . При этом можно ввести понятие расстояния  $\Delta$  между центрами сфер радиусами  $R_c$  и  $R$ , а сам фрагмент определить как выпукло-вогнутый сегмент или выпукло-вогнутый мениск (линзу).

Цель работы — разработка математической модели и создание средств оснащения для корректировки формы роторов.

### Моделирование процесса корректировки формы и дисбаланса сферического ротора

Моделирование [18–21] процесса корректировки формы ротора обусловлено формированием целевой функции, в которой оптимизируемыми аргументами являются сочетание управляющих факторов.

У реального ротора значения  $\Delta h$ ,  $R_c$  и угол  $\theta_{\max}$  определяются из профилограммы меридионального сечения, полученной при измерении формы ротора после напыления. При этом корректировка геометрии

ротора возможна за счет дополнительного напыления, обеспечивающего заполнение указанного фрагмента материалом покрытия на втором этапе напыления. Формирование покрытия должно осуществляться только в зоне сферического сегмента, определяемого углом  $\theta_{\max}$ , что на практике может быть обеспечено за счет использования экрана с отверстием диаметра  $d$ , размещенного между источником напыляемого материала и ротором [22].

Комплексно рассмотрена технология формирования поверхностного покрытия и сформулированы принципы, определяющие оптимизацию процесса корректировки функциональных параметров ротора в процессе дополнительного этапа напыления:

- размещение между ротором и источником экрана с отверстием диаметра  $d$ , ось которого лежит на динамической оси ротора, обеспечивает нанесение покрытия на поверхность сферического сегмента, определяемого углом  $\theta_{\max}$ ;
- осуществление в процессе нанесения дополнительного слоя перемещения экрана или ротора вдоль динамической оси ротора, изменяя расстояние между ротором и экраном, и вращение ротора вокруг динамической оси, что позволяет регулировать кривизну формируемого покрытия;
- использование точечного источника напыляемого материала, что позволяет при применении экрана с отверстием управлять геометрией напыляемой поверхности.

Рассмотрим предложенную схему корректировки формы методом дополнительного напыления и определим зависимость желаемого профиля покрытия от значимых параметров  $\Delta h$  и  $\theta_{\max}$ , характеризующих отклонение от сферичности поверхности ротора, полученное на этапе напыления.

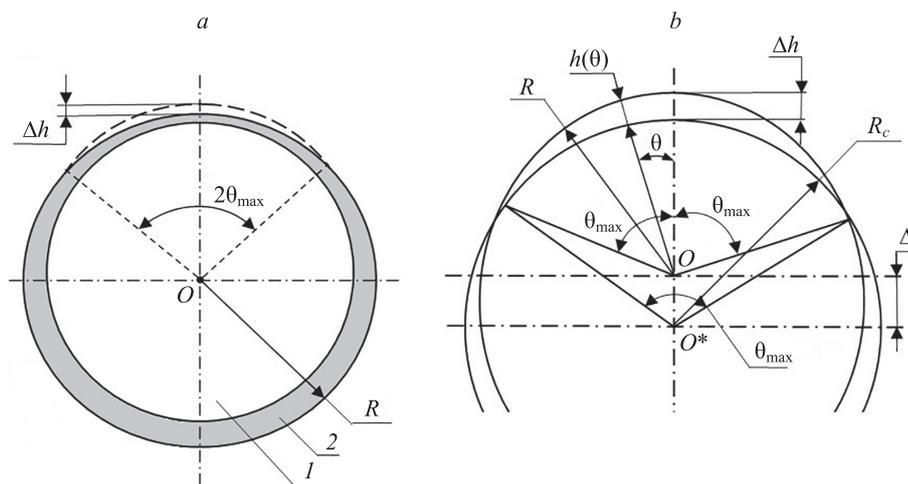


Рис. 1. Схема напыления при неподвижном экране: искажение формы ротора (а) и формируемый фрагмент (б).

$\theta_{\max}$  — угол, определяющий зону экранирования; 1 — ротор; 2 — покрытие;  $R$  — радиус ротора;  $\Delta h$  — отклонение по толщине покрытия;  $R_c$  — радиус сферического сегмента в зоне экранирования;  $h(\theta)$  — толщина покрытия в произвольной точке;  $\Delta$  — расстояние между центрами сфер радиусами  $R_c$  и  $R$

Fig. 1. Deposition scheme with a fixed screen. Rotor shape distortion.

1 — rotor; 2 — coating;  $R$  is the radius of the rotor;  $\Delta h$  is the deviation in the coating thickness;  $\theta_{\max}$  is the angle defining the shielding zone (a); Fragment being formed:  $R_c$  is the radius of the spherical segment in the shielding zone;  $h(\theta)$  is the coating thickness at an arbitrary point;  $\Delta$  is the distance between the centers of spheres with radii  $R_c$  and  $R$  (b)

Будем считать, что  $h$  — толщина покрытия на оси симметрии фрагмента (максимальная толщина покрытия) и используем понятие расстояния  $\Delta$  между центрами  $O$  и  $O^*$  сфер радиусами  $R_c$  и  $R$  (рис. 1). Последовательными преобразованиями на основе геометрических построений получим выражение:

$$h = \frac{\Delta h - \Delta + \Delta \cos \theta^*}{1 - \frac{\Delta}{R_c} \cos \theta^*}. \quad (1)$$

Подставим в выражение (1) соотношение  $\Delta h = \frac{\Delta h}{1 - \cos \theta_{\max}}$  и получим желаемый профиль в зависимости от угла экранирования  $\theta_{\max}$ :

$$h(\theta) = \Delta h \frac{\cos \theta - \cos \theta_{\max}}{1 - \cos \theta_{\max}}. \quad (2)$$

С точки зрения практической реализации реальный профиль толщины получаемого покрытия можно получить при условии использования неподвижного экрана, размещенного между источником напыляемого материала и ротором. Реальный профиль определим зависимостью скорости осаждения покрытия  $V_k$  от угла  $\alpha$  между направлением потока и нормалью к поверхности ротора в данной точке, а также расходимостью потока, определяемой соотношением  $1/l^2$ , где  $l$  — расстояние от источника до данной точки поверхности ротора.

Согласно [7], определим скорость осаждения:

$$V_k \sim A \frac{\cos \alpha}{l^2}, \quad (3)$$

где  $A$  — размерная константа, зависящая от конкретных характеристик источника.

По теореме косинусов из рис. 1, считая, что  $\theta_{\max} \approx \theta^*_{\max}$ , определим выражения для  $l^2$  и для угла  $\alpha$  в виде:

$$l = (R_c)^2 + (L + R_c)^2 - 2(R_c + L)R_c \cos \theta^*, \quad (4)$$

$$\cos \alpha = \frac{(L + R_c) \cos \theta^* - R_c}{((R_c)^2 + (L + R_c)^2 - 2(R_c + L)R_c \cos \theta^*)^{1/2}} \quad (5)$$

где  $L$  — минимальное расстояние от источника до поверхности ротора. Учитывая выражения (4) и (5), из зависимости (3) можно получить скорость осаждения:

$$V_k = A \frac{(L + R_c) \cos \theta^* - R_c}{((R_c)^2 + (L + R_c)^2 - 2(R_c + L)R_c \cos \theta^*)^{3/2}} \quad (6)$$

В выражении (6) можно пренебречь малым отличием углов  $\theta$  и  $\theta^*$  и радиусов  $R$  и  $R_c$ .

При этом профиль толщины покрытия  $h(\theta)$  получим из соотношения:

$$h(\theta) = V_k(\theta) \tau = \tau A \frac{(L + R_c) \cos \theta^* - R_c}{((R_c)^2 + (L + R_c)^2 - 2(R_c + L)R_c \cos \theta^*)^{3/2}} \quad (7)$$

где  $\tau$  — время напыления.

Таким образом, выражение (2) определяет желаемый профиль напыления, а выражение (7) — реальный. Видно, что функциональные зависимости от угла  $\theta$  в выражениях (2) и (7) различны, поэтому ни при каких значениях параметров  $(L, R, \theta_{\max})$  они не будут тождественны. Можно лишь подобрать значения, чтобы разница была минимальна.

На рис. 2 показаны графические зависимости, определяющие желаемый профиль, и рассчитанные в соответствии с выражением (7) профили толщины для разных величин  $L$  и углов  $\theta_{\max}$ . Профили были нормированы по  $\Delta h$ . Радиус ротора  $R_c = 5$  мм.

Из графиков рис. 2 видно, что можно добиться компенсации толщины покрытия на оси сферического сегмента. При этом профиль толщины покрытия в зоне всего сегмента будет отличаться от желаемого.

В ряде случаев, когда параметры искажения формы после первого этапа напыления покрытия имеют сравнительно небольшие величины, может быть использован рассматриваемый метод корректировки формы, поскольку несовпадение профилей будет составлять тысячные доли микрометра. Однако в широком применении предложенная схема корректировки может оказаться недостаточной и необходимо обеспечить формирование желаемого профиля толщины покрытия без отклонений.

Чтобы добиться максимально точного совпадения, требуемого и фактически напыляемого профилей толщины, предложено использовать кинематическую схему, в которой осуществлено поступательное движение экрана или ротора, синхронизированное с процессом напыления покрытия [22]. Таким образом, целесообразно

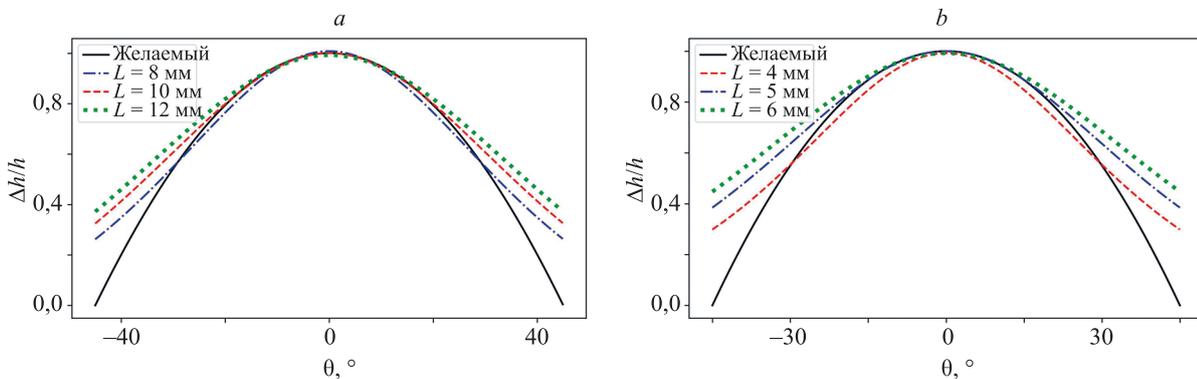


Рис. 2. Нормированные зависимости профилей толщины для разных величин  $L$  при  $2\theta_{\max}$ :  $90^\circ$  (a) и  $60^\circ$  (b)

Fig. 2. Normalized dependencies of thickness profiles for different values of  $L$  and for  $2\theta_{\max}$ :  $90^\circ$  (a) and  $60^\circ$  (b)

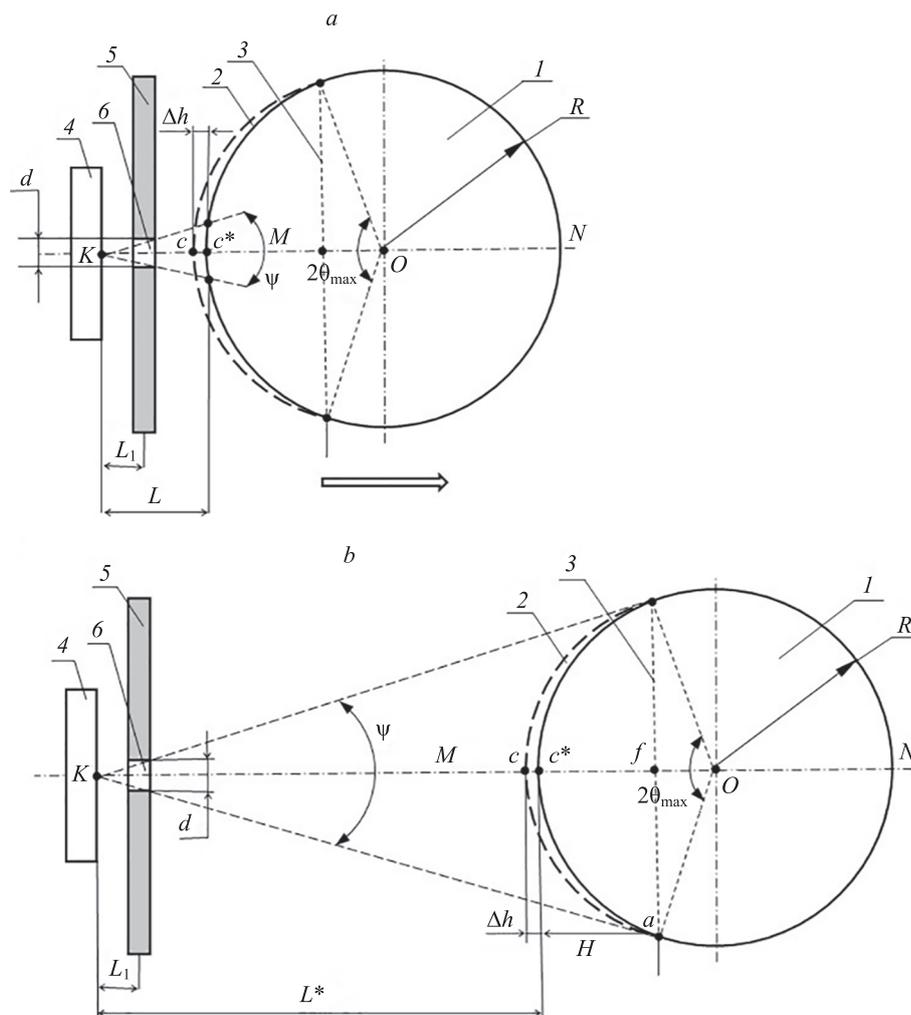


Рис. 3. Схема процесса напыления в начале (а) и конце (б) цикла напыления при подвижном экране.

1 — ротор; 2 — шаровой сегмент, соответствующий зоне экранирования; 3 — основание шарового сегмента 2; 4 — источник напыляемого материала с позицией К точечного испарения; 5 — экран; 6 — отверстие в экране диаметром  $d$ ;  $L_1$  — расстояние от источника 4 до экрана 5;  $L$  — минимальное и  $L^*$  максимальное расстояние от источника 4 до поверхности ротора 1;  $c^*$  — вершина шарового сегмента 2;  $MN$  — динамическая ось ротора;  $\psi$  — угол, определяющий поток напыляемого материала;  $H$  — высота шарового сегмента 2;  $af$  и  $bf$  — радиусы основания 3 шарового сегмента 2

Fig. 3. Scheme of the deposition process at the beginning (a) and end (b) of the deposition cycle with a moving screen.

1 — rotor; 2 — spherical segment corresponding to the shielding zone; 3 — base of spherical segment 2; 4 — source of sprayed material with position K of point evaporation; 5 — screen; 6 — hole in the screen with a diameter  $d$ ;  $L_1$  is the distance from source 4 to screen 5;  $L$  is a minimum and  $L^*$  is a maximum distance from source 4 to the surface of rotor 1;  $c^*$  is the top of the spherical segment 2;  $MN$  is the dynamic axis of the rotor;  $\psi$  is the angle that determines the flow of the sprayed material;  $H$  is the height of the spherical segment 2;  $af$  and  $bf$  are the radii of the base 3 of the spherical segment 2

но рассмотреть подход, предполагающий поступательное перемещение ротора, движущегося от минимального расстояния  $L$  от источника до поверхности ротора до максимального расстояния  $L^*$  (рис. 3).

Обозначив постоянное расстояние от источника до экрана  $L_1$ , а расстояние от экрана до ротора  $x$ , и рассматривая ротор, равномерно движущийся от экрана со скоростью  $u$ , в начальный момент времени расположенный вплотную к экрану, зависимость  $x(\tau)$  представим в виде  $x(\tau) = u\tau$ .

Для точек поверхности ротора, задаваемых углом  $\psi$ , которые в силу конечной ширины отверстия в экране «открыты» для напыления на протяжении всего времени движения ротора, получим выражение для определения толщины покрытия:

$$h(\theta) = \int_0^T \frac{A \cos \theta}{(L_1 + ut)^2} dt, \quad (8)$$

где  $T$  — полное время движения. Вычислив интеграл в (8), получим:

$$h(\theta) = \frac{A \cos \theta L_1}{L_1^2 u} \left( 1 - \frac{1}{1 + uT/L_1} \right). \quad (9)$$

Так как задано, что при нулевом угле  $\theta$  необходимо компенсировать толщину  $\Delta h$ , то должно выполняться соотношение:

$$\Delta h = h(\theta = 0) = \frac{A L_1}{L_1^2 u} \left( 1 - \frac{1}{1 + uT/L_1} \right). \quad (10)$$

Подставим выражение (10) в (9), тогда итоговое выражение имеет вид:

$$h(\theta) = \Delta h \cos \theta. \quad (11)$$

Для точек поверхности ротора, которые «закрываются» для напыления в начальные моменты времени, выражение для толщины напыления будет иметь вид, аналогичный выражению (8).

Можно показать, что для углов  $\theta$ , не удовлетворяющих этому условию, выражение для толщины напыления имеет вид:

$$h(\theta) = \int_{\frac{L_1}{u} \left( \frac{2R}{d} \sin \theta - 1 \right)}^T \frac{A \cos \theta}{(L_1 + ut)^2} dt. \quad (12)$$

Вычислив интеграл в (12), получим:

$$h(\theta) = \frac{A \cos \theta L_1}{L_1^2} \left( \frac{2R}{u d \sin \theta} - \frac{1}{1 + uT/L_1} \right).$$

Введем обозначение  $uT = x_{\max} = L_1 \left( \frac{2R}{d} \sin \theta_{\max} - 1 \right)$ , и используя соотношение (10), получим:

$$h(\theta) = \Delta h \cos \theta \frac{\frac{1}{\sin \theta} - \frac{1}{\sin \theta_{\max}}}{\frac{2R}{d} - \frac{1}{\sin \theta_{\max}}}. \quad (13)$$

Таким образом, для углов  $\theta$ , определяющих точки, которые в силу конечной ширины отверстия в экране «открыты» для напыления на протяжении всего времени движения ротора, толщина напыления определяется выражением (8), а остальных углов — (13).

Сравним получающиеся в результате моделирования профили толщины с желаемыми, которые определены в (2).

На рис. 4 представлены желаемые и рассчитанные в результате моделирования в соответствии с выражениями (11) и (13) профили толщины напыляемого фрагмента при подвижном экране для различных значений  $\theta_{\max}$ .

Видно, что в отличие от неподвижного экрана, где характерной особенностью было несоответствие про-

филя поверхности фрагмента и ротора по краям фрагмента и фактическое наличие ступеньки, при подвижном экране зона наибольшего несоответствия толщины располагается в средней части фрагмента. Также имеет место закономерность — меньшему углу соответствует меньшее отклонение профиля от желаемого.

Для корректной оптимизации сочетаний управляющих факторов, обеспечивающих формирование требуемого фрагмента покрытия и получения взаимосвязи значимых параметров процесса напыления, рассмотрим подобные (по признаку равных углов) прямоугольные треугольники в системах «источник 4 — экран 5» и «источник 4 — ротор 1».

Для системы «источник 4 — экран 5» из соотношения  $\text{tg} \frac{\psi}{2} = \frac{d}{2L_1}$  следует

$$\frac{\psi}{2} = \text{arctg} \frac{d}{2L_1}$$

или

$$\psi = 2 \text{arctg} \frac{d}{2L_1}. \quad (14)$$

Для системы «источник 4 — ротор 1» запишем  $\text{tg} \frac{\psi}{2} = \frac{af}{L^* + H}$  или  $\psi = 2 \text{arctg} \frac{af}{L^* + H}$ , а поскольку  $af = R \cos(\theta_{\max}/2)$ , то имеем

$$\psi = 2 \text{arctg} \frac{R \cos \frac{\theta_{\max}}{2}}{L^* + H}. \quad (15)$$

Приравняем выражения (14) и (15), получим

$$2 \text{arctg} \frac{d}{2L_1} = 2 \text{arctg} \frac{R \cos \frac{\theta_{\max}}{2}}{L^* + H},$$

или

$$\frac{d}{2L_1} = \frac{R \cos \frac{\theta_{\max}}{2}}{L^* + H}.$$

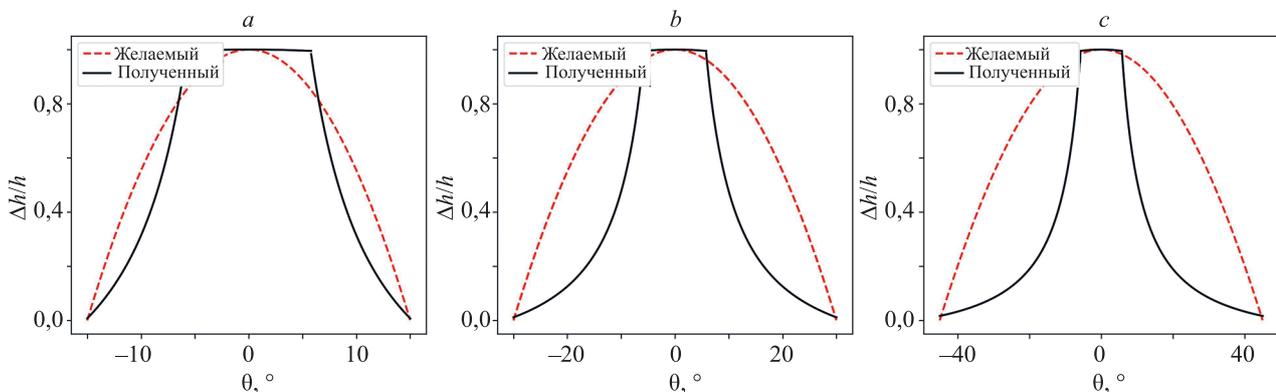


Рис. 4. Желаемые и рассчитанные профили толщины для схемы напыления при подвижном экране для значений  $2\theta_{\max}$ : 30° (a), 60° (b), 90° (c)

Fig. 4. Desired and actual profiles of thickness for deposition scheme with moving screen for  $2\theta_{\max}$ : 30° (a), 60° (b), 90° (c)

Время напыления  $\tau$  и, соответственно, время перемещения ротора  $l$  определим, исходя из полученного эмпирическим путем значения скорости  $V_k$  осаждения покрытия с учетом геометрических параметров ротора  $l$  и шарового сегмента 2. Последовательным преобразованием выражений, определяющих связь времени напыления с толщиной покрытия и скоростью осаждения, это можно представить в виде соответствующей зависимости.

Достаточно очевидной является необходимость вращения ротора при напылении дополнительного слоя вокруг динамической оси, что позволяет компенсировать влияние возможной неоднородности потока в зоне напыления.

Таким образом, определив реальные для каждого ротора значения  $\Delta h$  и  $\theta_{\max}$  с помощью профилограммы, возможно подобрать сочетание управляющих факторов, в частности соотношение параметров  $d$ ,  $L^*$  и  $L_1$ , обеспечивающих необходимые условия для корректировки формы ротора. Важным аспектом в данном случае является то, что время напыления  $\tau$  соответствует времени перемещения экрана  $T$ . Это обеспечивает условие формирования требуемой толщины  $\Delta h$  дополнительного слоя покрытия на вершине сегмента и монотонное убывание толщины по мере приближения к основанию этого сегмента, где толщина дополнительного слоя равна нулю.

#### Апробация результатов моделирования

Апробация разработанной модели для объективной оценки ее адекватности выполнена следующим образом.

Нанесено корректирующее дополнительное покрытие при неподвижном экране, когда искажение формы после первого этапа напыления покрытия имело сравнительно небольшую величину — отклонение гармоники  $A_3$  от предельно допустимого значения не превышало 0,01 мкм. Сформировано дополнительное покрытие с перемещением экрана в условиях, когда амплитуды гармоники  $A_3$  превышали допустимые на 0,01–0,025 мкм.

При неподвижном экране нанесено покрытие при расстоянии  $L = 12$  мм (рис. 2). В результате обеспече-

но соответствие амплитуд гармоник  $A_3$  техническим требованиям.

Проведена корректировка формы ротора с подвижным экраном, если амплитуды гармоники  $A_3$  находятся в интервале 0,003–0,005 мкм. После корректировки значения амплитуд получены в пределах 0,018–0,023 мкм. Расстояние  $L^*$  составило 60 мм,  $L_1$  — 16 мм, а диаметр отверстия  $d$  в экране — 1 мм при радиусе  $R$  ротора 5 мм.

Геометрия роторов до и после процесса нанесения покрытия по результатам измерения остальных гармоник осталась в пределах допуска, значения амплитуд изменились на тысячные доли микрометра.

Положительные результаты практического использования подтвердили адекватность предложенных моделей.

#### Заключение

При исследовании влияния схемы фиксации ротора при напылении покрытия показано, что асимметричное расположение элементов фиксации приводит к искажению формы ротора, что вызывает изменение дисбаланса.

Предложена математическая модель корректировки формы ротора за счет дополнительного напыления выпукло-вогнутого сегмента в зоне экранирования. Обосновано, что изготовление ротора с требуемыми параметрами предполагает две последовательные и взаимосвязанные операции нанесения покрытия. Рассмотрены два альтернативных варианта формирования корректирующего покрытия, основанные на использовании размещаемого между ротором и источником напыляемого материала экрана с отверстием расчетного диаметра:

- с фиксированным расстоянием между ротором, экраном и источником;
- ротором, перемещающимся от минимального расстояния от экрана до поверхности ротора до заданного максимального расстояния.

Экспериментальная проверка подтвердила адекватность разработанных моделей. Предложенный подход позволяет расширить технологические возможности процесса изготовления сферических роторов.

#### Литература

1. Peshekhonov V.G. The outlook for gyroscopy // *Gyroscopy and Navigation*. 2020. V. 11. N 3. P. 193–197. <https://doi.org/10.1134/S2075108720030062>
2. Gusinsky V.Z., Litmanovich Y.A., Lesyuchevsky V.M. Spacecraft attitude determination using electrostatically suspended gyros // *IFAC Proceedings Volumes*. 2001. V. 34. N 15. P. 26–31. [https://doi.org/10.1016/S1474-6670\(17\)40700-2](https://doi.org/10.1016/S1474-6670(17)40700-2)
3. Ландау Б.Е., Белаш А.А., Гуревич С.С., Левин С.Л., Романенко С.Г., Цветков В.Н. Электростатический гироскоп в системах ориентации космических аппаратов // *Гироскопия и навигация*. 2021. № 3. С. 69–79. <https://doi.org/10.17285/0869-7035.0071>
4. Юльметова О.С. Ионно-плазменные и лазерные технологии в гироскопическом приборостроении: диссертация на соискание ученой степени доктора технических наук. СПб.: ГНЦ РФ АО «Концерн «ЦНИИ «Электроприбор», 2018. 244 с.

#### References

1. Peshekhonov V.G. The outlook for gyroscopy. *Gyroscopy and Navigation*, 2020, vol. 11, no. 3, pp. 193–197. <https://doi.org/10.1134/S2075108720030062>
2. Gusinsky V.Z., Litmanovich Y.A., Lesyuchevsky V.M. Spacecraft attitude determination using electrostatically suspended gyros. *IFAC Proceedings Volumes*, 2001, vol. 34, no. 15, pp. 26–31. [https://doi.org/10.1016/S1474-6670\(17\)40700-2](https://doi.org/10.1016/S1474-6670(17)40700-2)
3. Landau B.E., Belash A.A., Gurevich S.S., Levin S.L., Romanenko S.G., Tsvetkov V.N. Electrostatic gyroscope in spacecraft attitude reference systems. *Gyroscopy and Navigation*, 2021, vol. 12, no. 3, pp. 247–253. <https://doi.org/10.1134/s2075108721030056>
4. Iulmetova O.S. *Ion-plasma and laser technologies in gyroscopic instrumentation*. Thesis for the degree of Doctor of Technical Sciences. St. Petersburg, Concern CSRI Elektropribor, 2018, 244 p. (in Russian)

5. Grigoriev S., Vereschaka A., Zelenkov V., Sitnikov N., Bublikov J., Milovich F., Andreev N., Sotova C. Investigation of the influence of the features of the deposition process on the structural features of microparticles in PVD coatings // *Vacuum*. 2022. V. 202. P. 111144. <https://doi.org/10.1016/j.vacuum.2022.111144>
6. Maksakova O., Simoëš S., Pogrebnyak A., Bondar O., Kravchenko Y., Beresnev V., Erdybaeva N. The influence of deposition conditions and bilayer thickness on physical-mechanical properties of CA-PVD multilayer ZrN/CrN coatings // *Materials Characterization*. 2018. V. 140. P. 189–196. <https://doi.org/10.1016/j.matchar.2018.03.048>
7. Фомичев А.М., Щербак А.Г., Юльметова О.С., Беляев С.Н. Устройство для напыления тонкопленочных покрытий на сферические роторы электростатического гироскопа. Патент RU2555699C1. Бюл. 2015. № 19.
8. Ou Y.X., Wang H.Q., Ouyang X., Zhao Y.Y., Zhou Q., Luo C.W., Hua Q.S., Ouyang X.P., Zhang S. Recent advances and strategies for high-performance coatings // *Progress in Materials Science*. 2023. V. 136. P. 101125. <https://doi.org/10.1016/j.pmatsci.2023.101125>
9. Махаев Е.А., Рябова Л.П., Чесноков П.А., Щербак А.Г., Кузнецов С.А., Мартюшов Г.Г. Разработка конструкции и технологии изготовления ротора криогироскопа // *Материалы XXX конференции памяти Н.Н. Острякова*. Санкт-Петербург, 2016. С. 116–123.
10. Dubrovskiy A.R., Okunev M., Makarova O., Kuznetsov S. Corrosion resistance of the substrates for the cryogenic gyroscope and electrodeposition of the superconductive niobium coatings // *Journal of Physics: Conference Series*. 2017. V. 857. N 1. P. 012008. <https://doi.org/10.1088/1742-6596/857/1/012008>
11. Tit M.A., Yulmetova O.S., Sisyukov A.N. Analysis of niobium thin film modification based on thermodynamic approach // *Proc. of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. 2020. P. 1049–1052. <https://doi.org/10.1109/eiconrus49466.2020.9038932>
12. Тит М.А., Беляев С.Н. Влияние стехиометрического состава на функциональные характеристики тонкопленочных поверхностных структур на узлах гироскопических приборов // *Вопросы материаловедения*. 2019. № 4(100). С. 78–90. <https://doi.org/10.22349/1994-6716-2019-100-4-78-90>
13. Tit M.A., Yulmetova O.S., Andreeva V.D., Sisyukov A.N., Yulmetova R.F. Application of niobium nitride thin films for improvement of performance characteristics of cryogenic gyroscope // *Proc. of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. 2021. P. 1232–1235. <https://doi.org/10.1109/EIConRus51938.2021.9396714>
14. Scherbak A.G., Yulmetova O.S., Contrast image formation based on thermodynamic approach and surface laser oxidation process for optoelectronic read-out system // *Optics & Laser Technology*. 2018. V. 101. P. 242–247. <https://doi.org/10.1016/j.optlastec.2017.11.030>
15. Yulmetova O.S., Scherbak A.G. Composition analysis of thin films formed on beryllium surfaces under pulsed laser action by the method of chemical thermodynamics // *International Journal of Advanced Manufacturing Technology*. 2018. V. 97. N 9-12. P. 3231–3236. <https://doi.org/10.1007/s00170-018-2216-2>
16. Yulmetova O.S., Tumanova M.A. Laser marking of contrast images for optical read-out systems // *Journal of Physics: Conference Series*. 2017. V. 917. N 5. P. 052007. <https://doi.org/10.1088/1742-6596/917/5/052007>
17. Yulmetova O.S., Tit M.A., Sisyukov A.N., Yulmetova R.F. Comparative analysis of electrochemical and laser marking for gyroscopic application // *Proc. of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. 2020. P. 1053–1055. <https://doi.org/10.1109/EIConRus49466.2020.9039076>
18. Chen L., Zhao Y., Song B., Yu T., Liu Z. Modeling and simulation of 3D geometry prediction and dynamic solidification behavior of Fe-based coatings by laser cladding // *Optics & Laser Technology*. 2021. V. 139. P. 107009. <https://doi.org/10.1016/j.optlastec.2021.107009>
19. Katranidis V., Gu S., Allcock B., Kamnis S. Experimental study of high velocity oxy-fuel sprayed WC-17Co coatings applied on complex geometries. Part A: Influence of kinematic spray parameters on thickness, porosity, residual stresses and microhardness // *Surface and Coatings Technology*. 2017. V. 311. P. 206–215. <https://doi.org/10.1016/j.surfcoat.2017.01.015>
20. Sahoo D.K., Neupane N., Chaudhary S. Predicting the coating geometries of friction surfaced aluminium cladding over carbon steel by artificial neural network modeling. *Materials Today: Proceedings*, 2023, in press. <https://doi.org/10.1016/j.matpr.2023.01.056>
5. Grigoriev S., Vereschaka A., Zelenkov V., Sitnikov N., Bublikov J., Milovich F., Andreev N., Sotova C. Investigation of the influence of the features of the deposition process on the structural features of microparticles in PVD coatings. *Vacuum*, 2022, vol. 202, pp. 111144. <https://doi.org/10.1016/j.vacuum.2022.111144>
6. Maksakova O., Simoëš S., Pogrebnyak A., Bondar O., Kravchenko Y., Beresnev V., Erdybaeva N. The influence of deposition conditions and bilayer thickness on physical-mechanical properties of CA-PVD multilayer ZrN/CrN coatings. *Materials Characterization*, 2018, vol. 140, pp. 189–196. <https://doi.org/10.1016/j.matchar.2018.03.048>
7. Fomichev A.M., Shcherbak A.G., Jul'metova O.S., Beljaev S.N. Device for sputtering of thin film coatings on spherical rotor of electrostatic gyroscope. *Patent RU2555699C1*, 2015. (in Russian)
8. Ou Y.X., Wang H.Q., Ouyang X., Zhao Y.Y., Zhou Q., Luo C.W., Hua Q.S., Ouyang X.P., Zhang S. Recent advances and strategies for high-performance coatings. *Progress in Materials Science*, 2023, vol. 136, pp. 101125. <https://doi.org/10.1016/j.pmatsci.2023.101125>
9. Makhaev E.A., Ryabova L.P., Chesnokov P.A., Shcherbak A.G., Kuznetsov S.A., Martiushov G.G. Development of a construction and manufacturing technology for the cryogyroscope rotor. *Proc. of the 30th Conference of Memory N.N. Ostryakov*. St. Petersburg, 2016, pp. 116–123. (in Russian)
10. Dubrovskiy A.R., Okunev M., Makarova O., Kuznetsov S. Corrosion resistance of the substrates for the cryogenic gyroscope and electrodeposition of the superconductive niobium coatings. *Journal of Physics: Conference Series*, 2017, vol. 857, no. 1, pp. 012008. <https://doi.org/10.1088/1742-6596/857/1/012008>
11. Tit M.A., Yulmetova O.S., Sisyukov A.N. Analysis of niobium thin film modification based on thermodynamic approach. *Proc. of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2020, pp. 1049–1052. <https://doi.org/10.1109/eiconrus49466.2020.9038932>
12. Tit M.A., Belyaev S.N. Effect of Stoichiometric Composition on the Functional Characteristics of Thin-Film Surface Structures on Gyroscope Components. *Inorganic Materials: Applied Research*, 2020, vol. 11, no. 6, pp. 1370–1377. <https://doi.org/10.1134/s2075113320060283>
13. Tit M.A., Yulmetova O.S., Andreeva V.D., Sisyukov A.N., Yulmetova R.F. Application of niobium nitride thin films for improvement of performance characteristics of cryogenic gyroscope. *Proc. of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2021, pp. 1232–1235. <https://doi.org/10.1109/EIConRus51938.2021.9396714>
14. Scherbak A.G., Yulmetova O.S., Contrast image formation based on thermodynamic approach and surface laser oxidation process for optoelectronic read-out system. *Optics & Laser Technology*, 2018, vol. 101, pp. 242–247. <https://doi.org/10.1016/j.optlastec.2017.11.030>
15. Yulmetova O.S., Scherbak A.G. Composition analysis of thin films formed on beryllium surfaces under pulsed laser action by the method of chemical thermodynamics. *International Journal of Advanced Manufacturing Technology*, 2018, vol. 97, no. 9-12, pp. 3231–3236. <https://doi.org/10.1007/s00170-018-2216-2>
16. Yulmetova O.S., Tumanova M.A. Laser marking of contrast images for optical read-out systems. *Journal of Physics: Conference Series*, 2017, vol. 917, no. 5, pp. 052007. <https://doi.org/10.1088/1742-6596/917/5/052007>
17. Yulmetova O.S., Tit M.A., Sisyukov A.N., Yulmetova R.F. Comparative analysis of electrochemical and laser marking for gyroscopic application. *Proc. of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, 2020, pp. 1053–1055. <https://doi.org/10.1109/EIConRus49466.2020.9039076>
18. Chen L., Zhao Y., Song B., Yu T., Liu Z. Modeling and simulation of 3D geometry prediction and dynamic solidification behavior of Fe-based coatings by laser cladding. *Optics & Laser Technology*, 2021, vol. 139, pp. 107009. <https://doi.org/10.1016/j.optlastec.2021.107009>
19. Katranidis V., Gu S., Allcock B., Kamnis S. Experimental study of high velocity oxy-fuel sprayed WC-17Co coatings applied on complex geometries. Part A: Influence of kinematic spray parameters on thickness, porosity, residual stresses and microhardness. *Surface and Coatings Technology*, 2017, vol. 311, pp. 206–215. <https://doi.org/10.1016/j.surfcoat.2017.01.015>
20. Sahoo D.K., Neupane N., Chaudhary S. Predicting the coating geometries of friction surfaced aluminium cladding over carbon steel by artificial neural network modeling. *Materials Today: Proceedings*, 2023, in press. <https://doi.org/10.1016/j.matpr.2023.01.056>

- by artificial neural network modeling // *Materials Today: Proceedings*. 2023. in press. <https://doi.org/10.1016/j.matpr.2023.01.056>
21. Paturi U.M.R., Cheruku S., Geereddy S.R. Process modeling and parameter optimization of surface coatings using artificial neural networks (ANNs): State-of-the-art review // *Materials Today: Proceedings*. 2021. V. 38. Part 5. P. 2764–2774. <https://doi.org/10.1016/j.matpr.2020.08.695>
  22. Тит М.А., Ландау Б.Е., Филиппов А.Ю., Щербак А.Г. Способ изготовления ротора шарового гироскопа. Патент RU2743492C1. Бюл. 2020. № 5.
  21. Paturi U.M.R., Cheruku S., Geereddy S.R. Process modeling and parameter optimization of surface coatings using artificial neural networks (ANNs): State-of-the-art review. *Materials Today: Proceedings*, 2021, vol. 38, part 5, pp. 2764–2774. <https://doi.org/10.1016/j.matpr.2020.08.695>
  22. Tit M.A., Landau B.E., Filippov A.Y., Shcherbak A.G. Ball gyroscope rotor manufacturing method. *Patent RU2743492C1*, 2020. (in Russian)

#### Авторы

**Тит Маргарита Алексеевна** — научный сотрудник, АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург, 197046, Российская Федерация, [sc 57216270946](https://orcid.org/0009-0003-0069-2940), <https://orcid.org/0009-0003-0069-2940>, [rita93.07.93.07@gmail.com](mailto:rita93.07.93.07@gmail.com)

**Беляев Сергей Николаевич** — кандидат технических наук, старший научный сотрудник, АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург, 197046, Российская Федерация, [sc 57216270946](https://orcid.org/0000-0003-3765-9469), <https://orcid.org/0000-0003-3765-9469>, [serenible@mail.ru](mailto:serenible@mail.ru)

**Щербак Александр Григорьевич** — доктор технических наук, начальник сектора, АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург, 197046, Российская Федерация, [sc 57197817003](https://orcid.org/0000-0001-8507-7653), <https://orcid.org/0000-0001-8507-7653>, [asherbak1946@gmail.com](mailto:asherbak1946@gmail.com)

**Юльметова Ольга Сергеевна** — доктор технических наук, начальник сектора, АО «Концерн «ЦНИИ «Электроприбор», Санкт-Петербург, 197046, Российская Федерация; доцент практики, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 37068364500](https://orcid.org/0000-0001-9899-5939), <https://orcid.org/0000-0001-9899-5939>, [olga@yulmetova.ru](mailto:olga@yulmetova.ru)

#### Authors

**Margarita A. Tit** — Scientific Researcher, Concern CSRI Elektropribor, JSC, Saint Petersburg, 197046, Russian Federation, [sc 57216270946](https://orcid.org/0009-0003-0069-2940), <https://orcid.org/0009-0003-0069-2940>, [rita93.07.93.07@gmail.com](mailto:rita93.07.93.07@gmail.com)

**Sergey N. Belyaev** — PhD, Senior Researcher, Concern CSRI Elektropribor, JSC, Saint Petersburg, 197046, Russian Federation, [sc 57216270946](https://orcid.org/0000-0003-3765-9469), <https://orcid.org/0000-0003-3765-9469>, [serenible@mail.ru](mailto:serenible@mail.ru)

**Alexandr G. Shcherbak** — D.Sc., Head of Sector, Concern CSRI Elektropribor, JSC, Saint Petersburg, 197046, Russian Federation, [sc 57197817003](https://orcid.org/0000-0001-8507-7653), <https://orcid.org/0000-0001-8507-7653>, [asherbak1946@gmail.com](mailto:asherbak1946@gmail.com)

**Olga S. Yulmetova** — D.Sc., Head of Sector, Concern CSRI Elektropribor, JSC, Saint Petersburg, 197046, Russian Federation; Associate Professor of Practice, ITMO University, Saint Petersburg, Russian Federation, [sc 37068364500](https://orcid.org/0000-0001-9899-5939), <https://orcid.org/0000-0001-9899-5939>, [olga@yulmetova.ru](mailto:olga@yulmetova.ru)

*Статья поступила в редакцию 24.04.2023*  
*Одобрена после рецензирования 19.05.2023*  
*Принята к печати 16.07.2023*

*Received 24.04.2023*  
*Approved after reviewing 19.05.2023*  
*Accepted 16.07.2023*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-812-819

УДК 519.724.2; 621.391

## Метод пространственного мультиплексирования в многоантенных системах связи

Алексей Юрьевич Гришенцев✉

Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

[AGrishentsev@yandex.ru](mailto:AGrishentsev@yandex.ru)✉, <https://orcid.org/0000-0003-1373-0670>

### Аннотация

**Введение.** Исследованы вопросы надежности системы связи с пространственным мультиплексированием. Повышение пропускной способности каналов радиосвязи за счет пространственного мультиплексирования является одним из наиболее востребованных и актуальных направлений современных исследований в области радиосвязи. Решение задачи пространственного мультиплексирования во временной области при многолучевом распространении сопряжено со значительным ростом размерности задачи и избыточными вычислениями. Детектирование во временной области затрудняет учет частотной зависимости амплитуды и фазы принимаемых сигналов, что в свою очередь снижает вероятность верного распознавания сигналов. **Метод.** При многолучевом распространении предложено решение задачи пространственного мультиплексирования в частотной области путем применения теоремы о свертке. Произведена оценка вероятности ошибки при использовании предлагаемого метода детектирования. Выполнена оценка устойчивости решения в зависимости от обусловленности матрицы амплитудно-фазовых параметров. Выведено математическое выражение для оценки верхней границы вероятности ошибки в субканале в зависимости от числа обусловленности матрицы амплитудно-фазовых параметров и спектральной плотности шума в физических каналах связи. Разработан алгоритм адаптивного формирования матриц амплитудно-фазовых параметров, позволяющий выбирать среди избыточного числа приемных антенн такие антенны, которые дают возможность повысить устойчивость детектирования за счет снижения числа обусловленности матрицы коэффициентов системы линейных уравнений. **Основные результаты.** Разработана теоретическая основа метода пространственного мультиплексирования в многоантенных системах связи. Предлагаемый метод позволяет повысить эффективность вычислений за счет снижения размерности задачи детектирования по сравнению с решением во временной области. Предложено решать задачи детектирования только на частотах, на которых ожидается прием полезного сигнала, что особенно полезно при узкополосных частотных и фазовых, ортогональных и биортогональных видах модуляции, часто используемых в многоантенных цифровых системах связи. Выведены соотношения, позволяющие дать оценку вероятности ошибки в субканале. Разработан алгоритм адаптивного формирования матриц амплитудно-фазовых параметров, позволяющий повысить устойчивость решения задачи детектирования. **Обсуждение.** Результаты исследований могут найти применение при разработке многоантенных систем связи с пространственным мультиплексированием.

### Ключевые слова

радиосвязь, пространственное мультиплексирование, многоантенные системы, метод пространственного кодирования сигнала, MIMO

### Благодарности

Работа выполнена при поддержке Министерства науки и высшего образования Российской Федерации, проект тематики научных исследований № 2019-0898.

**Ссылка для цитирования:** Гришенцев А.Ю. Метод пространственного мультиплексирования в многоантенных системах связи // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 812–819. doi: 10.17586/2226-1494-2023-23-4-812-819

## Method of spatial multiplexing in multi-antenna communication systems

Alexey Yu. Grishentsev✉

ITMO University, Saint Petersburg, 197101, Russian Federation

AGrishentsev@yandex.ru✉, <https://orcid.org/0000-0003-1373-0670>

### Abstract

Reliability of the communication system with spatial multiplexing has been studied. Increasing the bandwidth of radio communication channels due to spatial multiplexing is one of the most popular and relevant areas of modern research in the field of radio communications. Solving the problem of spatial multiplexing in the time domain with multipath propagation is associated with a significant increase in the dimension of the problem and redundant calculations. Detection in the time domain makes it difficult to take into account the frequency dependence of the amplitude and phase of the received signals, which in turn reduces the probability of correct signal recognition. In multipath propagation, a solution to the problem of spatial multiplexing in the frequency domain is proposed by applying the convolution theorem. The probability of error is estimated when using the proposed detection method. The stability of the solution is estimated depending on the conditionality of the matrix of amplitude-phase parameters. The expression of the estimation of the upper bound of the error probability in the subchannel is derived depending on the number of conditionality of the matrix of amplitude-phase parameters and the spectral density of noise in physical communication channels. An algorithm for adaptive formation of matrices of amplitude-phase parameters has been developed which selects such antennas among an excessive number of receiving antennas allowing to increase the stability of detection by reducing the number of conditionality of the matrix of coefficients of a system of linear equations. The theoretical basis of the spatial multiplexing method in multi-antenna communication systems has been developed. The proposed method makes it possible to increase the efficiency of calculations by reducing the dimension of the detection problem in comparison with the solution in the time domain. It is proposed to solve the detection problem only at frequencies at which a useful signal is expected to be received, which is especially useful for narrow-band frequency and phase, orthogonal and biorthogonal types of modulation often used in multi-antenna digital communication systems. Expressions for estimating the probability of error in the subchannel are derived. An algorithm for adaptive formation of matrices of amplitude-phase parameters has been developed, which makes it possible to increase the stability of the solution of the detection problem. The research results are applicable in the development of multi-antenna communication systems with spatial multiplexing.

### Keywords

radio communication, spatial multiplexing, multi-antenna systems, Multiple Input Multiple Output, MIMO

### Acknowledgements

This work was supported by the Ministry of Science and Higher Education of the Russian Federation, passport of the State Assignment no. 2019-0898.

**For citation:** Grishentsev A.Yu. Method of spatial multiplexing in multi-antenna communication systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 812–819 (in Russian). doi: 10.17586/2226-1494-2023-23-4-812-819

### Введение

Увеличение пропускной способности каналов связи является актуальной задачей [1]. Значительное внимание уделяется возможности увеличения полосы пропускания за счет пространственного мультиплексирования, называемого в международном сообществе MIMO (Multiple Input Multiple Output — множественный вход, множественный выход) по отношению к каналу связи [2]. Практика исследования и использования MIMO не оправдала в полной мере надежды, образовавшиеся в начале XXI века на принципиальное увеличение пропускной способности. Потому появившиеся в начале развития технологии MIMO ожидания и прогнозы можно назвать сверхоптимистичными. При этом технология MIMO прочно заняла свое место в повседневных задачах связи и продолжает развиваться. Существенной проблемой повышения пропускной способности MIMO является необходимость высокоточного определения характеристик канала связи при многолучевом распространении и необходимость большого числа вычислений в реальном режиме времени [3]. Остается нерешенной проблема динамической оптимизации конфигурации передающих и принимающих антенн [4–7]. Наиболее надежными [8] и потому

востребованными на сегодняшний день являются системы, использующие две передающие и две излучающие антенны с пространственным и поляризационным разделением, а также технология изменения диаграммы направленности за счет поворота фазы в антенне-решетке — так называемая схема Алмоути [9]. Отметим, что в условиях прямой видимости [10, 11] канал MIMO успешно используется в оптическом диапазоне волн, крайне высокочастотном и высокочастотном радиодиапазонах. Есть все основания полагать, что в условиях прямой видимости применение систем MIMO будет увеличиваться. А с ростом доступности скоростных вычислительных мощностей неизбежно увеличение пространственных субканалов в системах связи с отсутствием прямой видимости. В данном исследовании произведена разработка и анализ метода детектирования в частотной области многолучевых сигналов с пространственным мультиплексированием. Выведено выражение для расчета отношения сигнала к шуму в субканале, предложен метод оптимизационного выбора подмножества антенн для повышения устойчивости детектирования сигналов. Исследован канал связи с пространственным мультиплексированием, не исключающий применения временного, частотного и/или кодового разделения сигналов.

**Базовая модель пространственного мультиплексирования в многоантенных системах связи типа «точка-точка»**

В качестве базовой модели [12] использована модель однолучевого распространения дискретного времени узкополосного канала ММО для систем связи типа «точка-точка»:

$$\mathbf{Y} = \mathbf{H}\mathbf{\Psi} + \mathbf{\Omega}, \quad (1)$$

где  $\mathbf{Y}$  — вектор (размера  $M_r$ ) принимаемого сигнала на каждой антенне;  $\mathbf{H}$  — матрица коэффициентов затухания сигнала размера  $M_r \times M_t$ ;  $\mathbf{\Psi}$  — вектор (размера  $M_t$ ) передаваемых сигналов с каждой антенны;  $\mathbf{\Omega}$  — вектор аддитивного шума размера  $M_r$ ;  $M_t$  и  $M_r$  — число передающих и принимающих антенн.

Для эффективного распознавания переданных сигналов на основе решения системы (1), а значит, пространственного мультиплексирования, необходимо иметь дополнительную информацию о состояниях каналов в приемнике (Channel State Information at the Receiver, CSIR) и в передатчике (Channel State Information at the Transmitter, CSIT) [13].

Параметры CSIR используем для решения задачи детектирования сообщений в приемнике. Наличие информации CSIT о канале связи на передающей стороне позволяет корректировать следующие характеристики излучения: диаграмму направленности антенн; мощность излучения; используемые сигнальные алфавиты; метод кодирования и временной зазор, препятствующий межсимвольной интерференции. На принимающей стороне информация CSIR дает возможность оптимизировать диаграмму направленности и частотную избирательность приемных устройств. Если процесс выбора оптимальных характеристик на передающей и/или принимающей стороне происходит в динамике с учетом текущего состояния канала связи, то канал связи — *адаптивный*.

В системах ММО для оптимального декодирования обычно применяют метод максимального правдоподобия [12, 14]. Существенный недостаток базовой модели (1) заключается в том, что не учитывается многолучевое распространение. Данное допущение снижает вычислительные затраты, но значительно ограничивает возможности, а значит, и область применения модели.

**Модель пространственного мультиплексирования в многоантенных системах связи при многолучевом распространении**

При многолучевом распространении сложность модели канала связи с пространственным мультиплексированием значительно возрастает.

Каждый сигнал, принятый антенной с индексом  $m_r = 1 \dots M_r$ , выражается в виде уравнения:

$$y_{m_r}(t) = \sum_{m_t=1}^{M_t} \sum_{k=1}^{K_{m_r,m_t}} \alpha_{m_r,m_t,k} x_{m_t}(t + \Delta t_{m_r,m_t,k}) + n_{m_r}(t), \quad (2)$$

где  $\alpha_{m_r,m_t,k}$  и  $\Delta t_{m_r,m_t,k}$  — коэффициент затухания и время задержки при распространении  $k$ -го луча, от  $m_t$ -ой ан-

тенны к  $m_r$ -ой антенне;  $y_{m_r}$  и  $n_{m_r}(t)$  — сигнал и аддитивный шум принимаемый  $m_r$ -ой антенной;  $x_{m_t}$  — сигнал излученный  $m_t$ -ой антенной;  $K_{m_r,m_t}$  — число учитываемых лучей при распространении от  $m_t$ -ой антенны к  $m_r$ -ой антенне. В случае учета достаточного числа лучей  $k$ , компонент  $n_{m_r}(t)$  — является аддитивным белым гауссовым шумом (Additive White Gaussian Noise, AWGN) с нулевым средним и дисперсией  $\sigma_{m_r}^2 = \frac{1}{2} N_0^{m_r}$ , принимаемый  $m_r$ -ой антенной;  $N_0^{m_r}$  — спектральная плотность мощности AWGN.

Учитывая, что при пространственном мультиплексировании используется множество  $m_t = 1 \dots M_t$  передающих антенн и множество  $m_r = 1 \dots M_r$  принимающих антенн совокупность выражений (2) образует систему уравнений:

$$\begin{cases} y_1(t) = \sum_{m_t=1}^{M_t} \sum_{k=1}^{K_{1,m_t}} \alpha_{1,m_t,k} x_{m_t}(t + \Delta t_{1,m_t,k}) + n_1(t) \\ y_{m_r}(t) = \sum_{m_t=1}^{M_t} \sum_{k=1}^{K_{m_r,m_t}} \alpha_{m_r,m_t,k} x_{m_t}(t + \Delta t_{m_r,m_t,k}) + n_{m_r}(t). \\ y_{M_r}(t) = \sum_{m_t=1}^{M_t} \sum_{k=1}^{K_{M_r,m_t}} \alpha_{M_r,m_t,k} x_{m_t}(t + \Delta t_{M_r,m_t,k}) + n_{M_r}(t) \end{cases} \quad (3)$$

Для решения системы уравнений (3) и определения  $x_{m_t}(t + \Delta t_{m_r,m_t,k})$  необходимо иметь набор параметров CSIR, характеризующих канал связи с пространственным разделением:  $\alpha_{m_r,m_t,k}$ ;  $\Delta t_{m_r,m_t,k}$ ; число лучей  $K_{m_r,m_t}$  при распространении сигнала  $x_{m_t}$  в субканале от  $m_t$ -ой антенны к  $m_r$ -ой антенне. Для решения (3) принимающая многоантенная система должна быть системой с памятью, чтобы была возможность детектирования переданных сигналов с учетом задержки распространения в различных лучах.

В общем случае  $\Delta t_{m_r,m_t,k}$  — время задержки распространения  $k$ -го луча сигнала  $x_{m_t}(t)$  в субканале ( $m_r, m_t$ ) является функцией от частоты и разности фаз  $\theta(S, M_t)$  сигналов, излучаемых различными антеннами из общего массива  $M_t$ , т. е.  $\Delta t_{m_r,m_t,k}(j\omega, \theta(S, M_t))$ ; и  $\alpha_{m_r,m_t,k}$  — коэффициенты затухания  $k$ -го луча сигнала  $x_{m_t}(t)$  в субканале ( $m_r, m_t$ ) является функцией от частоты и фазы, т. е.  $\alpha_{m_r,m_t,k}(j\omega, \theta(S, M_t))$ , где  $S$  — мощность применяемого сигнального алфавита. От разности фаз  $\theta(S, M_t)$  сигналов, излучаемых различными антеннами, зависит диаграмма направленности в случае использования массива антенн-решеток, что используется для блочного кодирования по схеме Алмоути [9]. Частотная зависимость параметров  $\Delta t_{m_r,m_t,k}$  и  $\alpha_{m_r,m_t,k}$  особенно актуальна для широкополосных и тем более сверхширокополосных систем связи. Здесь и далее по тексту:  $j^2 = -1$  — квадрат базисного нормированного вектора мнимой единицы;  $\omega = 2\pi f$  — круговая частота;  $f$  — частота сигнала;  $j\omega$  — комплексный параметр преобразования Фурье.

Сложность решения (3) во временной области при многолучевом распространении заключается в том, что при наличии  $\Delta t_{m_r,m_t,k} \neq 0$  для решения (3) в заданный момент времени  $t$  приходится совместно решать уравнения (3) в моменты времени  $t$  и  $t - \Delta t_{m_r,m_t,k}$ ,  $\Delta t_{m_r,m_t,k} \neq 0$ ,

что, в свою очередь, приводит к разрастанию размерности системы уравнений и необходимости значительных вычислительных затрат. При решении в частотной области системы (3) имеется возможность производить решение для заданных частот независимо от других частот, что может быть особенно полезно при частотном разделении сигналов, так как позволяет производить решение и детектирование сигналов только в той области спектра, в которой ожидается прием сигнала, при этом размерность системы линейных уравнений останется равной числу принимающих антенн  $M_r$ .

В многоканальных системах связи с многолучевым распространением для описания статистики флуктуаций традиционно используют распределение Релея или Накагами [12, 15], и такой подход действительно актуален для модели (1). В модели (3), при учете достаточного числа лучей, шумовая компонента соответствует каналу связи с аддитивным белым гауссовым шумом. Распределение Накагами также будет актуально для модели (3), в случае если остаются неучтенными лучи, средняя суммарная мощность которых будет оказывать значимое влияние на вид распределения [16].

### Модель пространственного мультиплексирования в частотной области при многолучевом распространении

Преобразование Фурье [17, 18] выражения (2) имеет следующий вид:

$$Y_{m_r}(j\omega) = \sum_{m_t=1}^{M_t} \sum_{k=1}^{K_{m_r, m_t}} \alpha_{m_r, m_t, k}(j\omega) \frac{1}{\sqrt{T}} \int_{t_0}^{t_0+T} x_{m_t} \times \\ \times (t + \Delta t_{m_r, m_t, k}(j\omega)) e^{-j\omega t} dt + \frac{1}{\sqrt{T}} \int_{t_0}^{t_0+T} n_{m_r}(t) e^{-j\omega t} dt, \quad (4)$$

где  $T = \tau + \Delta t_{\max}$  — период интегрирования;  $\tau$  — длительность сигнального символа;  $\Delta t_{\max} = \max(\Delta t_{m_r, m_t, k}(j\omega))$  — максимальное время задержки.

Применив теорему о задержке [18, 19] к выражению (4), получим:

$$Y_{m_r}(j\omega) = \sum_{m_t=1}^{M_t} \sum_{k=1}^{K_{m_r, m_t}} \alpha_{m_r, m_t, k}(j\omega) e^{-j\omega \Delta t_{m_r, m_t, k}(j\omega)} X_{m_t}(j\omega) + N_{m_r}(j\omega). \quad (5)$$

Введем обозначение для суммы амплитудных (коэффициентов затухания) и фазовых (время задержки) параметров многолучевого субканала  $m_t$  распространения сигнала  $x_{m_t}(t)$  от  $m_t$ -ой антенны к  $m_r$ -ой антенне:

$$\gamma_{m_r, m_t}(j\omega) = \sum_{k=1}^{K_{m_r, m_t}} \alpha_{m_r, m_t, k}(j\omega) e^{-j\omega \Delta t_{m_r, m_t, k}(j\omega)},$$

тогда выражение (5) примет вид:

$$Y_{m_r}(j\omega) = \sum_{m_t=1}^{M_t} \gamma_{m_r, m_t}(j\omega) X_{m_t}(j\omega) + N_{m_r}(j\omega). \quad (6)$$

Применив аналогичные (4)–(6) преобразования к каждому уравнению системы (3), получим:

$$\begin{cases} Y_1(j\omega) = \sum_{m_t=1}^{M_t} \gamma_{1, m_t}(j\omega) X_{m_t}(j\omega) + N_1(j\omega) \\ \dots \\ Y_{m_r}(j\omega) = \sum_{m_t=1}^{M_t} \gamma_{m_r, m_t}(j\omega) X_{m_t}(j\omega) + N_{m_r}(j\omega). \\ \dots \\ Y_{M_r}(j\omega) = \sum_{m_t=1}^{M_t} \gamma_{M_r, m_t}(j\omega) X_{m_t}(j\omega) + N_{M_r}(j\omega) \end{cases} \quad (7)$$

Полученная система (7) по форме аналогична системе уравнений (1), с тем отличием, что в случае системы (7) вместо матрицы коэффициентов затухания фигурирует  $\Gamma(j\omega) = (\gamma_{m_r, m_t}(j\omega))$  — матрица, для частоты  $\omega$  составленная из членов ряда преобразования Фурье комплексных амплитудно-фазовых коэффициентов размера  $M_r \times M_t$ . Обозначим матрицу  $\Gamma(j\omega)$  без указания зависимости от частоты, т. е.  $\Gamma$ , кроме случаев, где необходимо обратить дополнительное внимание на частотную зависимость матрицы  $\Gamma$ . Запишем (7) в матричном виде:

$$\Gamma \mathbf{X} = \mathbf{Y} - \mathbf{N}. \quad (8)$$

При условии неравенства нулю определителя  $\det(\Gamma) \neq 0$ , решение системы (8) будет иметь вид:

$$\mathbf{X} = \Gamma^{-1}(\mathbf{Y} - \mathbf{N}) = \Gamma^{-1}\mathbf{Y} - \Gamma^{-1}\mathbf{N}. \quad (9)$$

Решение (7) позволяет учесть энергию некоторого множества лучей при многолучевом распространении. По своей сути повышение отношения сигнал/шум за счет аддитивного учета энергии отдельных лучей при многолучевом распространении относится к технологии RAKE. В контексте решаемой задачи технология RAKE применяется для каждого субканала ( $m_r, m_t$ ).

### Оценка вероятности ошибки и устойчивости решения в частотной области

Из выражения (9) видно, что точность и возможность решения системы (8) определяется матрицей  $\Gamma$  и вектором  $\mathbf{N}$ . Выполним анализ влияния матрицы  $\Gamma$  и вектора  $\mathbf{N}$  на решение и вероятность ошибки. Канал связи, заданный системой уравнений (8), позволяет реализовать максимальное число субканалов мультиплексирования в том случае, если число принимающих антенн  $M_r$  больше или равно числу передающих антенн  $M_t$ , т. е.  $M_r \geq M_t$ . При условии  $M_r \geq M_t$  система (8) будет разрешима. Положим, что  $M_r = M_t$ , тогда число субканалов равно  $M_r = M_t$ .

Запишем матрицу  $\Gamma^{-1}(j\omega)$  комплексных амплитудно-фазовых коэффициентов в виде, удобном для дальнейших рассуждений:

$$\Gamma^{-1}(j\omega) = \begin{pmatrix} \beta_{1,1}(j\omega) & \dots & \beta_{1, M_r}(j\omega) \\ \vdots & \ddots & \vdots \\ \beta_{M_t, 1}(j\omega) & \dots & \beta_{M_t, M_r}(j\omega) \end{pmatrix}, \quad (10)$$

где  $\beta_{m_r, m_t}(j\omega)$  — элементы матрицы  $\Gamma^{-1}$ .

В соответствии с (9) решение матрицы (10) определим выражением:

$$X_{m_i}(j\omega) = \sum_{m_r=1}^{M_r} \beta_{m_i, m_r}(j\omega)(Y_{m_r}(j\omega) + N_{m_r}(j\omega)). \quad (11)$$

Учитывая (10) и (11), произведем анализ влияния малого возмущения правой части системы уравнений (8) и матрицы  $\Gamma$  на устойчивость решения. Запишем выражение (9) в виде:

$$\bar{X} + \Delta X = \Gamma^{-1}(Y + N), \quad (12)$$

где  $\bar{X}$  — вектор полезных сигналов, полученных в ходе решения (8);  $N = (N_{m_r}(j\omega))$  — вектор возмущения правой части;  $\Delta X = (N_{m_i}(j\omega))$  — вектор возмущения решения (8),  $Y = (Y_{m_r}(j\omega))$  — вектор полезных сигналов принятых антеннами. Отметим, что возмущение правой части есть AWGN, в данном случае, при нулевом математическом ожидании, знак аддитивной компоненты не имеет значения (замечка для сравнения с (9)). Используя известное выражение [20], запишем отношение для оценки относительной погрешности:

$$\frac{\|\Delta X\|}{\|\bar{X}\|} \leq \chi(\Gamma) \frac{\|N\|}{\|Y\|}, \quad (13)$$

где  $\chi(\Gamma) = \|\Gamma\| \|\Gamma^{-1}\| \geq 1$  — число обусловленности матрицы  $\Gamma$ . В качестве нормы матриц и векторов в (13) и далее будем использовать евклидовы нормы, также называемые сферическими нормами. Применим для матрицы  $\Gamma$  норму по сумме квадратов элементов столбца, для матрицы  $\Gamma^{-1}$  норму по сумме квадратов элементов строки. Как известно, решение будет устойчивым, если число обусловленности  $\chi(\Gamma)$  близко по значению к единице [21].

Перепишем выражение (12) в виде, связывающем отношение энергии шума к энергии сигнала для частоты  $\omega$  на выходе антенно-фидерных устройств и в субканале:

$$\frac{1}{\chi^2(\Gamma(j\omega))} \frac{\sum_{m_r=1}^{M_r} |Y_{m_r}(j\omega)|^2}{\sum_{m_r=1}^{M_r} |N_{m_r}(j\omega)|^2} \leq \frac{\sum_{m_i=1}^{M_i} |\bar{X}_{m_i}(j\omega)|^2}{\sum_{m_i=1}^{M_i} |N_{m_i}(j\omega)|^2}. \quad (14)$$

Учитывая, что  $M_r = M_i$ , и произведя статистическое усреднение, можно сделать вывод, что отношение сигнала к шуму  $SNR_r(j\omega)$  на антеннах связано с отношением сигнала к шуму в субканалах  $SNR_t(j\omega)$  следующим образом:

$$\frac{SNR_r(j\omega)}{\chi^2(\Gamma(j\omega))} \leq SNR_t(j\omega). \quad (15)$$

Выражение (15) позволяет производить расчет вероятности ошибки в субканале при различных видах сигнального кодирования известными методами. Рассмотрим пример оценки вероятности ошибки в субканале при ортогональном кодировании.

### Оценка вероятности ошибки в субканале при ортогональном кодировании

В качестве примера решения задачи оценки вероятности ошибочного детектирования рассмотрим случай

использования для передачи сообщений элементов множества ортогональных сигнальных символов (сигнального алфавита мощностью  $S = 2^v$ ,  $v = 1, 2, 3, \dots$ ) априори известных на принимающей стороне  $\{\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_s(t), \dots, \hat{x}_S(t)\}$  и/или их частотных образов  $\{\hat{X}_1(j\omega), \hat{X}_2(j\omega), \dots, \hat{X}_s(j\omega), \dots, \hat{X}_S(j\omega)\}$ . Критерий правдоподобия [14, 22] для детектирования сигналов имеет вид:

$$\max_{m_i} \{p(X_{m_i}(j\omega) | \hat{X}_s(j\omega))\}, \quad (16)$$

где  $p(X_{m_i}(j\omega) | \hat{X}_s(j\omega))$  — функция условной плотности распределения вероятности, часто называемая функцией правдоподобия гипотезы о том, что передавался символ  $\hat{X}_s(j\omega)$  при данной фиксированной реализации сигнала  $X_{m_i}(j\omega)$ , определенного при решении (8). Отметим, что для использования множества  $\{\hat{x}_1(t), \hat{x}_2(t), \dots, \hat{x}_s(t), \dots, \hat{x}_S(t)\}$  при решении (16) необходимо выполнить выравнивание  $\hat{x}_s(t)$ ,  $s = 1, \dots, S$  по времени с  $\tau$  до  $T = \tau + \Delta t_{\max}$  дополнив  $\hat{x}_s(t)$  нулями длительностью  $\Delta t_{\max}$ , при этом произойдет выравнивание частотного разрешения для  $\hat{X}_s(j\omega)$  и  $X_{m_i}(j\omega)$ .

Не приводя подробности всех преобразований (подробный вывод имеется, например в [16, 23]), запишем выражение для оценки верхней границы вероятности ошибки ( $v$ -битового) символа в субканале:

$$P_{m_i} \leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left( 1 - \left( \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-0.5a^2} da \right)^{S-1} \right) \times \exp\left(-\frac{1}{2} \left( b - \sqrt{\frac{2E_{m_i}}{N_0^{m_i}}} \right)^2\right) db, \quad (17)$$

где  $a, b$  — добавленные аргументы интегральных выражений. Для вычисления (17) необходимое отношение энергии сигнала к односторонней спектральной плотности шума в субканале  $\frac{E_{m_i}}{N_0^{m_i}}$  можно выразить через аналогичное отношение на приемных антеннах в соответствии с выражением (14):

$$\frac{1}{\chi^2(\Gamma)} \frac{E_{m_r}}{N_0^{m_r}} \leq \frac{E_{m_i}}{N_0^{m_i}}. \quad (18)$$

В случае если ортогональные сигналы равновероятны, все вероятности ошибок на символ равновероятны, и возникают с вероятностью:  $\frac{P_{m_i}}{S-1}$ . Имеется  $C_g^v = \binom{v}{g}$  возможных реализаций, при которых из  $v$  переданных битов  $g$  приняты с ошибкой. Следовательно, среднее число ошибочных битов на  $v$ -битовый символ равно:

$$\sum_{g=1}^v g \binom{v}{g} \frac{P_{m_i}}{2^v - 1} = v \frac{2^{v-1}}{2^v - 1} P_{m_i}.$$

Средняя вероятность ошибки на бит в субканале  $m_i$ :

$$P_{m_i}^b = \frac{2^{v-1}}{2^v - 1} P_{m_i}.$$

Отметим, что в цифровых системах связи обычно используют критерий отношения энергии сигнала на бит  $E_{m_i}^b = \frac{1}{v} E_{m_i}$  к спектральной плотности мощности

шума  $N_0^{m_i}$ , т. е.  $\frac{E_{m_i}^b}{N_0^{m_i}}$ , в контексте рассматриваемого случая — в субканале.

### Алгоритм адаптивного формирования матриц амплитудно-фазовых параметров

Линейную независимость векторов амплитудно-фазовых параметров, образующих матрицу  $\Gamma$  невозможно обеспечить за счет задержания и/или усиления сигнала. Линейная независимость векторов амплитудно-фазовых параметров может быть достигнута только за счет уникальной для каждой антенны конфигурации (амплитудной и фазовой) принимаемых лучей при многолучевом распространении. Кроме пространственной избирательности, за счет диаграммы направленности излучающих и приемных антенн, дополнительно для повышения линейной независимости векторов амплитудно-фазовых параметров можно использовать поляризационное разделение, которое является традиционным способом для многоантенных систем типа ММО [2]. Линейная независимость векторов амплитудно-фазовых параметров, образующих матрицу  $\Gamma(j\omega)$  на одной частоте, например  $\omega_1$ , не гарантирует линейную независимость тех же векторов на частоте  $\omega_2$  ( $\omega_1 \neq \omega_2$ ).

Для построения модели адаптивного формирования матриц амплитудно-фазовых параметров для частоты  $\omega$  зададим множество приемных антенн  $\mathbf{A} = \{a_1, a_2, \dots, a_q, \dots, a_{Q-1}, a_Q\}$ , где  $q = 1, \dots, Q$  — номера доступных антенн из множества  $\mathbf{A}$ , так что  $\text{card}(\mathbf{A}) > M_r$ , т. е. мощность множества доступных для использования приемных антенн больше числа  $M_r$ . Для каждой антенны из множества  $\mathbf{A}$  известен набор амплитудно-фазовых параметров  $\gamma_{q,m_r}(j\omega)$ . Тогда для каждой частоты сформируем матрицу  $\Gamma(j\omega)$ , обладающую минимальным значением числа обусловленности  $\chi(\Gamma(j\omega))$ . Алгоритм формирования основан на принципе выбора, для каждой необходимой при детектировании частоты  $\omega$ , матрицы  $\Gamma(j\omega)$  с наибольшим определителем, составленной из нормированных значений амплитудно-фазовых параметров:

$$\max \left\{ \frac{\det(\Gamma_v(j\omega))}{\|\Gamma_v(j\omega)\|^{M_r}} \right\} \rightarrow \Gamma(j\omega), v = 1, \dots, C_{\text{card}(\mathbf{A})}^{M_r}.$$

Предположим, что число доступных антенн равно  $\text{card}(\mathbf{A})$ . Сформируем массивы используемых антенн  $M_r$ . Число возможных матриц  $\Gamma_v(j\omega)$  вычислим как число сочетаний из  $\text{card}(\mathbf{A})$  по  $M_r$ , т. е.  $C_{\text{card}(\mathbf{A})}^{M_r} = \frac{\text{card}(\mathbf{A})!}{M_r!(\text{card}(\mathbf{A}) - M_r)!}$ . Учитывая, что в многоантенных системах с пространственным мультиплексированием преимущественно используется узкополосная модуляция, то практическая необходимость вычисления определителя ограничена заданным набором частот  $\omega$ . Оценим вычислительную сложность расчета определителя как  $O(M_r^3)$  а общее число вычислений —  $O(M_r^3)C_{\text{card}(\mathbf{A})}^{M_r}$ . При этом фактическая вычислительная сложность будет значительно ниже, так как формирование каждой новой матрицы можно получить заменой лишь одной строки, что позволяет использовать ранее

вычисленные результаты. Заметим, что сама задача вычисления определителя хорошо распараллеливается, что позволяет существенно ускорить расчеты. Для стационарных систем связи, с квазистационарным паттерном многолучевых трасс, поиск оптимальной антенной конфигурации необходимо производить относительно редко, в отличие от подвижных систем связи или систем с быстрым изменением паттерна многолучевых трасс, для которых поиск оптимальной антенной конфигурации может быть необходим относительно часто.

Предлагаемый алгоритм и метод использования дополнительных антенн ( $\text{card}(\mathbf{A}) - M_r$ ) не является обязательной частью детектирования в частотной области, а позиционируется как дополнительный инструмент, ориентированный на повышение устойчивости решения задачи детектирования. С другой стороны, используя аналогичный подход, можно увеличить число передающих антенн до числа большего чем  $M_r$ , адаптивно используя только те, которые позволяют минимизировать значение числа обусловленности при решении системы (7) на принимающей стороне.

В системах связи с пространственным мультиплексированием существует несколько устоявшихся подходов к определению информации о состоянии канала в приемнике (CSIR), на основе которого формируются векторы амплитудно-фазовых параметров для каждой антенны:

- на основе математического моделирования канала связи;
- на основе измерений характеристик канала связи производимых самой приемно-передающей системой;
- смешанный подход, основанный на моделировании и измерениях.

Наилучшие результаты удается получить с помощью измерений характеристик канала связи, производимых самой приемно-передающей системой, и на основе смешанного подхода, основанного на измерениях и моделировании [24].

### Численное моделирование канала связи с пространственным разделением и детектированием сигналов в частотной области

В качестве численного моделирования выполним оценку вероятности ошибки в канале с пространственным мультиплексированием по четырем субканалам при использовании общего для всех субканалов ортогонального четырех символьного алфавита.

В таблице приведены характеристики матриц  $\Gamma$  и  $\Gamma^{-1}$ , имеющие значения для расчета вероятности ошибки. На основании численного моделирования построены графики (рисунок) средней вероятности ошибки по всем субканалам в зависимости от отношения сигнала к шуму на бит, для четырех различных значений числа обусловленности матрицы  $\Gamma$ . Моделирование произведено в среде MatLab<sup>1</sup>.

<sup>1</sup> MathWorks — Makers of MatLab and Simulink — MatLab & Simulink. MathWorks. [Электронный ресурс]. URL: <https://www.mathworks.com/> (дата обращения: 18.02.2023).

Таблица. Параметры численного моделирования  
Table. Numerical simulation parameters

Характеристика матриц	Рис. 1, <i>a</i>	Рис. 1, <i>b</i>	Рис. 1, <i>c</i>	Рис. 1, <i>d</i>
$ \det(\Gamma) $	1,000	0,21357	0,03415	0,00233
$\ \Gamma\ $	1,000	1,00800	0,99502	1,00505
$\ \Gamma^{-1}\ $	1,000	2,71739	3,12500	8,34937
$\chi(\Gamma)$ , дБ	1,000	2,73913	3,10943	8,39151

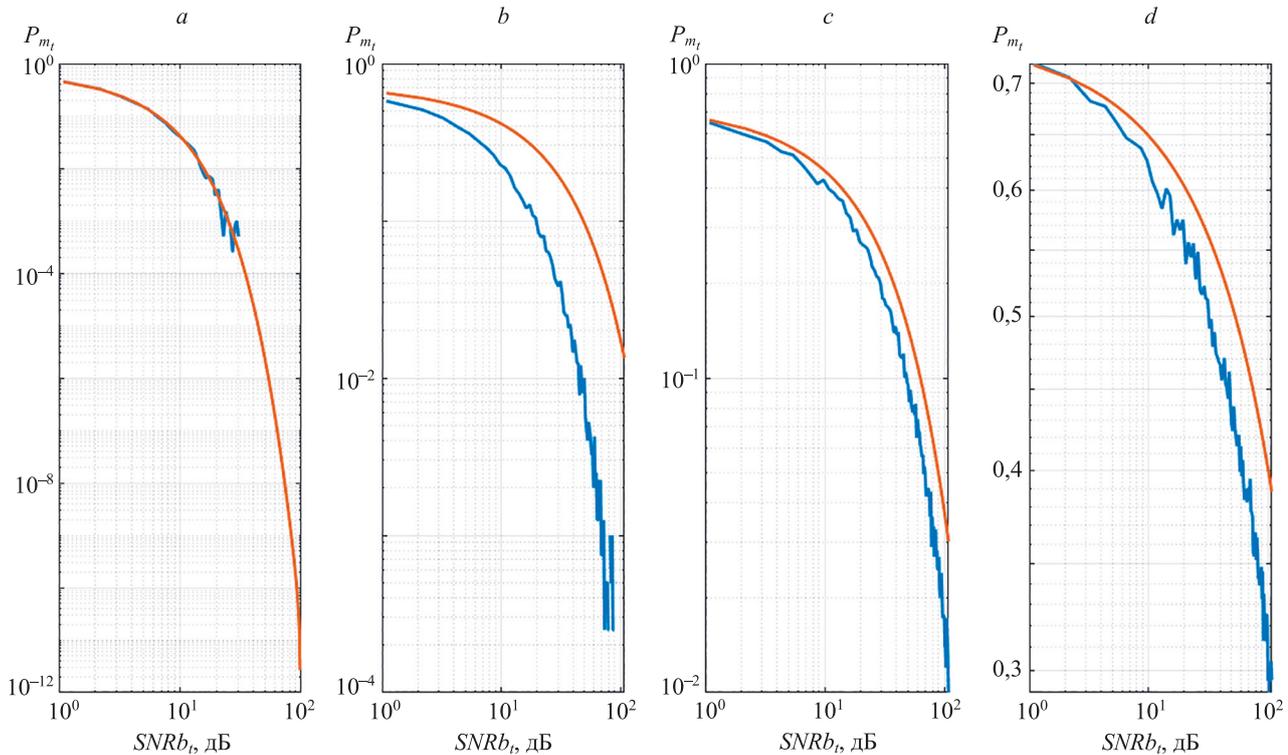


Рисунок. Моделирование вероятности средней статистической ошибки ( $P_{m_i}$ ) в субканалах в зависимости от отношения сигнала к шуму на бит ( $SNRb_i$ ) для четырех различных значений числа обусловленности матрицы  $\chi(\Gamma)$ : 1,000 (a); 2,73913 (b); 3,10943 (c); 8,39151 (d)

Теоретическая верхняя граница (красная линия) и результаты численного моделирования канала MIMO (синяя линия)

Figure. Modeling the probability of an average statistical error in subchannels vs. on the signal — to-noise ratio per bit for four different values of the matrix conditionality number:  $\chi(\Gamma)$ : 1.000 (a); 2.73913 (b); 3.10943 (c); 8.39151 (d).

Theoretical upper bound (red line) and results of numerical simulation of the MIMO channel (blue line)

Численное моделирование вероятности ошибки осуществлено с помощью разработанной модели MIMO канала с когерентным детектированием. Для каждой точки выполнена отправка 1 тыс. сообщений по четырем субканалам, поэтому усредненная статистика для каждой точки рассчитана по 4 тыс. значений. Полученные графики показали достаточно точную оценку верхней границы предложенным методом (17), (18).

### Заключение

Разработан метод решения задачи пространственного мультиплексирования в многоантенных системах

связи. Для многолучевого распространения предложено решение, реализуемое в частотной области, что позволяет значительно оптимизировать задачу детектирования за счет понижения ее размерности по отношению к решению во временной области. Выполнено исследование и анализ разработанного метода, выведено выражение для оценки вероятности ошибки в субканале многоантенной системы связи с пространственным разделением. Разработан алгоритм адаптивного формирования матриц амплитудно-фазовых параметров, что позволяет повысить устойчивость решения за счет оптимального использования увеличенного массива антенн.

## Литература

- Ипатов В. Широкополосные системы и кодовое разделение сигналов. принципы и приложения. М.: Техносфера, 2007. 487 с.
- Слюсар В.И. Системы ММО: принципы построения и обработка сигналов // Электроника: наука, технология, бизнес. 2005. № 8. С. 52–58.
- Hu C., Dai L., Mir T., Gao Z., Fang J. Super-resolution channel estimation for mmWave massive MIMO with hybrid precoding // IEEE Transactions on Vehicular Technology. 2018. V. 67. N 9. P. 8954–8958. <https://doi.org/10.1109/tvt.2018.2842724>
- Xue C.-D., Zhang X.Y., Cao Y.F., Hou Z., Ding C.F. MIMO antenna using hybrid electric and magnetic coupling for isolation enhancement // IEEE Transactions on Antennas and Propagation. 2017. V. 65. N 10. P. 5162–5170. <https://doi.org/10.1109/TAP.2017.2738033>
- Ратынский М.В. Адаптация и сверхразрешение в антенных решетках / 2-е изд. М.: Ленанд, 2022. 240 с.
- Bekkerman I., Tabrikian J. Target detection and localization using MIMO radars and sonars // IEEE Transactions on Signal Processing. 2006. V. 54. N 10. P. 3873–3883. <https://doi.org/10.1109/tsp.2006.879267>
- Хансен Р.С. Фазированные антенные решетки / 2-е изд. М.: Техносфера, 2012. 560 с.
- Богатырев В.А., Богатырев С.В., Богатырев А.В. Функциональная надежность вычислительных систем с перераспределением запросов // Известия высших учебных заведений. Приборостроение. 2012. Т. 55. № 10. С. 53–56.
- Alamouti S.I. A simple transmitter diversity technique for wireless communications // IEEE Journal on Selected Areas in Communications. 1998. V. 16. N 8. P. 1451–1458. <https://doi.org/10.1109/49.730453>
- Гришенцев А.Ю., Коробейников А.Г., Бондаренко И.Б. К вопросу о сверхразрешении чувствительных матриц // Журнал радиоэлектроники (электронный журнал). 2016. № 10. С. 12.
- Арслан Х., Чен Чж.Н., Ди Бендетто М. Сверхширокополосная беспроводная связь. М.: Техносфера, 2012. 640 с.
- Голдсмит А. Беспроводные коммуникации. М.: Техносфера, 2011. 904 с.
- Santamaria I., Elvira V., Via J., Ramirez D., Perez J., Ibanez J., Eickoff R., Ellinger F. Optimal MIMO transmission schemes with adaptive antenna combining in the RF path // Proc. of the 16<sup>th</sup> European Signal Processing Conference (EUSIPCO 2008). 2008. P. 102–107.
- Тихонов В.И. Статистическая радиотехника / 2-е изд., перераб. и доп. М.: Радио и связь, 1982. 624 с.
- Nakagami M. The m-Distribution—A general formula of intensity distribution of rapid fading // Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium Held at the University of California, Los Angeles, June 18–20, 1958. 1960. P. 3–36. <https://doi.org/10.1016/b978-0-08-009306-2.50005-4>
- Прокис Дж. Цифровая связь. М.: Радио и связь, 2000. 800 с.
- Письменный Д.Т. Конспект лекций по высшей математике: полный курс / 9-е изд. М.: Айрис-пресс, 2010. 608 с.
- Дёч Г. Руководство к практическому применению преобразования Лапласа и Z-преобразования. М.: Наука, 1971. 228 с.
- Оппенгейм А., Шафер Р. Цифровая обработка сигналов / 3-е изд, испр. М.: Техносфера, 2019. 1048 с.
- Самарский А.А., Гулин А.В. Численные методы. М.: Наука, 1989. 432 с.
- Воеводин В.В., Кузнецов Ю.А. Матрицы и вычисления. М.: Наука, 1984. 320 с.
- Кловский Д.Д. Теория передачи сигналов. М.: Связь, 1973. 376 с.
- Скляр Б. Цифровая связь. Теоретические основы и практическое применение / 2-е изд. М.: Вильямс, 2017. 1100 с.
- Deekshith P.K., Vinod S. Finite blocklength rates over a fading channel with CSIT and CSIR // Proc. of the 2018 IEEE International Conference on Communications (ICC). 2018. P. 1–7. <https://doi.org/10.1109/ICC.2018.8422087>

## Автор

**Гришенцев Алексей Юрьевич** — доктор технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 56321138400](https://orcid.org/0000-0003-1373-0670), <https://orcid.org/0000-0003-1373-0670>, [AGrishentsev@yandex.ru](mailto:AGrishentsev@yandex.ru)

## References

- Ipatov V.P. *Broadband Systems and Code Division of Signals. Principles and Applications*. Moscow, Tekhnosfera Publ., 2007, 487 p. (in Russian)
- Slusar V. MIMO systems. construction and signal processing principles. *Electronics: Science, Technology, Business*, 2005, no. 8, pp. 52–58. (in Russian)
- Hu C., Dai L., Mir T., Gao Z., Fang J. Super-resolution channel estimation for mmWave massive MIMO with hybrid precoding. *IEEE Transactions on Vehicular Technology*, 2018, vol. 67, no. 9, pp. 8954–8958. <https://doi.org/10.1109/tvt.2018.2842724>
- Xue C.-D., Zhang X.Y., Cao Y.F., Hou Z., Ding C.F. MIMO antenna using hybrid electric and magnetic coupling for isolation enhancement. *IEEE Transactions on Antennas and Propagation*, 2017, vol. 65, no. 10, pp. 5162–5170. <https://doi.org/10.1109/TAP.2017.2738033>
- Ratynskii M.V. *Adaptation and Superresolution in Antenna Arrays*. Moscow, Lenand Publ., 2022, 240 p. (in Russian)
- Bekkerman I., Tabrikian J. Target detection and localization using MIMO radars and sonars. *IEEE Transactions on Signal Processing*, 2006, vol. 54, no. 10, pp. 3873–3883. <https://doi.org/10.1109/tsp.2006.879267>
- Hansen R.C. *Phased Array Antennas*. Wiley, 2009, 580 p.
- Bogatyrev V.A., Bogatyrev S.V., Bogatyrev A.V. Functional reliability of computing systems with redistribution of inquiries. *Journal of Instrument Engineering*, 2012, vol. 55, no. 10, pp. 53–56. (in Russian)
- Alamouti S.I. A simple transmitter diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 1998, vol. 16, no. 8, pp. 1451–1458. <https://doi.org/10.1109/49.730453>
- Grishentsev A.Y., Korobeinikov A.G., Bondarenko I.B. Studies on superresolution image sensor. *Journal of Radio Electronics*, 2016, no. 10, pp. 12. (in Russian)
- Arslan H., Di Benedetto M.-G., Chen Z.N. *Ultra Wideband Wireless Communication*. Wiley, 2006, 500 p.
- Goldsmith A. *Wireless Communications*. Cambridge University Press, 2005, 644 p.
- Santamaria I., Elvira V., Via J., Ramirez D., Perez J., Ibanez J., Eickoff R., Ellinger F. Optimal MIMO transmission schemes with adaptive antenna combining in the RF path. *Proc. of the 16<sup>th</sup> European Signal Processing Conference (EUSIPCO 2008)*, 2008, pp. 102–107.
- Tikhonov V.I. *Statistical Radio Engineering*. Moscow, Radio i svjaz' Publ., 1982, 624 p. (in Russian)
- Nakagami M. The m-Distribution—A general formula of intensity distribution of rapid fading. *Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium Held at the University of California, Los Angeles, June 18–20, 1958*, 1960, pp. 3–36. <https://doi.org/10.1016/b978-0-08-009306-2.50005-4>
- Proakis J.G. *Digital Communications*. New York, McGraw-Hill Book Co., 1989.
- Pismennyi D.T. *Higher Mathematics: Lecture Notes*. Moscow, Ajris-Press Publ., 2010, 608 p. (in Russian)
- Doetsch G. *Anleitung zum praktischen Gebrauch der Laplace-transformation und der Z-transformation*. Oldenbourg, 1967.
- Oppenheim A.V., Schaffer R.W. *Discrete-time Signal Processing*. Pearson, 2010, 1108 p.
- Samarskii A.A., Gulina A.V. *Numerical Computations*. Moscow, Nauka Publ., 1989, 432 p. (in Russian)
- Voevodin V.V., Kuznetsov Iu.A. *Matrices and Calculations*. Moscow, Nauka Publ., 1984, 320 p. (in Russian)
- Klovskii D.D. *Signal Transmission Theory*. Moscow, Svjaz' Publ., 1973, 376 p. (in Russian)
- Sklar B. *Digital Communications*. Pearson Education, 2009, 1164 p.
- Deekshith P.K., Vinod S. Finite blocklength rates over a fading channel with CSIT and CSIR. *Proc. of the 2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7. <https://doi.org/10.1109/ICC.2018.8422087>

## Author

**Alexey Yu. Grishentsev** — D.Sc., Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 56321138400](https://orcid.org/0000-0003-1373-0670), <https://orcid.org/0000-0003-1373-0670>, [AGrishentsev@yandex.ru](mailto:AGrishentsev@yandex.ru)

doi: 10.17586/2226-1494-2023-23-4-820-827

УДК 536.24

## Математическое моделирование теплообменного аппарата с учетом сильной зависимости вязкости нефти от температуры

Динара Есентаевна Курманова<sup>1</sup>, Нурболат Жумабекович Джайчибеков<sup>2</sup>✉<sup>1,2</sup> Евразийский национальный университет им. Л.Н. Гумилева, Астана, 010008, Республика Казахстан<sup>1</sup> [dikonya89\\_29@mail.ru](mailto:dikonya89_29@mail.ru), <https://orcid.org/0009-0009-9787-7426><sup>2</sup> [jaich@mail.ru](mailto:jaich@mail.ru)✉, <https://orcid.org/0000-0002-3053-8288>

### Аннотация

**Введение.** Подогрев нефти и нефтепродуктов широко применяется для уменьшения энергопотерь при транспортировке. Течение в межтрубном пространстве теплообменника имеет сложный характер и зависит от многих факторов. Использование тонких трубок в теплообменных аппаратах геликоидного типа приводит к необходимости учета перехода режима течения от ламинарного к турбулентному. Традиционно используемые в численных расчетах полуэмпирические модели турбулентности не учитывают ламинарно-турбулентный переход. Разработан подход к определению эффективной длины теплообменного аппарата и температуры холодного теплоносителя на его выходе в случае сильной зависимости вязкости нефти от температуры с учетом возможности ламинарно-турбулентного перехода. В качестве нагреваемого теплоносителя рассмотрена нефть, а нагревающего компонента — вода. **Метод.** Новизна разработанного подхода заключается в применении модели турбулентности, учитывающей ламинарно-турбулентный переход, к расчету теплообменных аппаратов геликоидного типа. Для численного моделирования применены осредненные по Рейнольдсу уравнения Навье–Стокса, замкнутые при помощи модели турбулентности  $\gamma-Re_{0\tau}$ , учитывающей ламинарно-турбулентный переход. **Основные результаты.** Выполнено сравнение результатов численных расчетов с данными, полученными на основе метода среднелогарифмической разницы температур при постоянной и переменной вязкостях. В случае переменной вязкости нефти обнаружен переход ламинарного режима течения в турбулентный, который оказывает существенное влияние на эффективную длину теплообменного аппарата. **Обсуждение.** Результаты численных расчетов могут быть полезны при проектировании теплообменных аппаратов геликоидного типа.

### Ключевые слова

энергетика, теплообмен, вязкость, гидродинамика, нефтепродукты, численное моделирование, ламинарно-турбулентный переход

**Ссылка для цитирования:** Курманова Д.Е., Джайчибеков Н.Ж. Математическое моделирование теплообменного аппарата с учетом сильной зависимости вязкости нефти от температуры // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 820–827. doi: 10.17586/2226-1494-2023-23-4-820-827

## Modeling and simulation of heat exchanger with strong dependence of oil viscosity on temperature

Dinara E. Kurmanova<sup>1</sup>, Nurbolat Zh. Jaichibekov<sup>2</sup>✉<sup>1,2</sup> L.N. Gumilyov Eurasian National University, Astana, 010008, Kazakhstan<sup>1</sup> [dikonya89\\_29@mail.ru](mailto:dikonya89_29@mail.ru), <https://orcid.org/0009-0009-9787-7426><sup>2</sup> [jaich@mail.ru](mailto:jaich@mail.ru)✉, <https://orcid.org/0000-0002-3053-8288>

### Abstract

Heating of oil and oil products is widely used to reduce energy losses during transportation. The flow in the annular space of the heat exchanger is complex and depends on many factors. The use of thin tubes in helicoid-type heat exchangers makes it necessary to take into account the transition of the flow regime from laminar to turbulent. The semi-empirical turbulence models traditionally used in numerical calculations do not take into account the laminar-turbulent transition.

© Курманова Д.Е., Джайчибеков Н.Ж., 2023

An approach is developed to determine the effective length of the heat exchanger and the temperature of the cold coolant at its outlet in the case of a strong dependence of oil viscosity on temperature, taking into account the possibility of a laminar-turbulent transition. Oil is considered as a heated coolant, and water is considered as a heating component. The novelty of the developed approach lies in the application of the turbulence model, which takes into account the laminar-turbulent transition, to the calculation of helicoid-type heat exchangers. For numerical simulation, the Reynolds-averaged Navier–Stokes equations are used which are closed using  $\gamma$ - $Re_{0t}$  turbulence model that takes into account the laminar-turbulent transition. The results of numerical calculations are compared with the data obtained on the basis of the log-mean temperature difference method at constant and variable viscosity. In the case of variable oil viscosity, a transition from the laminar flow regime to the turbulent one is manifested which has a significant effect on the effective length of the heat exchanger. The results of numerical calculations can be useful in designing helicoid-type heat exchangers.

#### Keywords

energy, heat transfer, viscosity, hydrodynamics, oil products, simulation, laminar and turbulent transition

**For citation:** Kurmanova D.E., Jaichibekov N.Zh. Modeling and simulation of heat exchanger with strong dependence of oil viscosity on temperature. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 820–827 (in Russian). doi: 10.17586/2226-1494-2023-23-4-820-827

### Введение

Изменение качественного состояния сырьевой базы приводит к освоению и вовлечению в эксплуатацию месторождений нефти с высоким содержанием парафинов, смол, асфальтенов. Разработка таких месторождений требует применения нетрадиционных методов добычи нефти и ее подготовки к транспортировке. Светлые нефтепродукты (бензин, керосин) легко транспортируются по трубопроводам в любое время года и операции с ними не вызывают особых затруднений. Операции с темными нефтепродуктами (мазут, смазочные масла) и сырой нефтью вызывают значительные трудности, связанные тем, что темные нефтепродукты при понижении температуры воздуха становятся более вязкими, и их транспортировка без подогрева невозможна. Для трубопроводной транспортировки нефти и нефтепродуктов используется подход, основанный на регулировании реологических свойств нефти, например, при помощи нагрева нефти с ее последующей транспортировкой по трубопроводу с повышенной теплоизоляцией (горячая перекачка нефти). В некоторых случаях увеличение вязкости нефти при понижении температуры приводит к недопустимым напряжениям на стенках трубы и остановке транспортировки.

Процессы теплообмена осуществляются в теплообменных аппаратах (heat exchanger) различных типов и конструкций [1–5]. Для подогрева применяются различные теплоносители, например, горячая вода или водяной пар. Потребление энергии является одним из важных факторов, который оказывает существенное влияние на конструкцию теплообменника [6]. В нефтегазовой отрасли применение находят кожухотрубные теплообменники, которые обеспечивают хорошие характеристики производительности в широком диапазоне рабочих условий, высокую надежность и низкую стоимость.

Состав нефти (в частности, содержание асфальтенов, смол, парафинов) оказывает существенное влияние на зависимость вязкости от температуры [7, 8]. Эмпирические формулы, описывающие изменение кинематической вязкости в зависимости от температуры, имеют вид различных функций (экспоненциальные, полиномиальные, степенные и т. д.), которые характеризуются наличием коэффициентов, зависящих от

свойств жидкости. Постоянные коэффициенты определяются на основе значений измеренных кинематических вязкостей в экспериментальных точках. Для расчетов термодинамических параметров нефти, газовых конденсатов и их фракций используется обобщенное уравнение состояния Ли–Кесслера [9]. В работе [10] проведены исследования зависимости кинематической вязкости нефти и смесей нефтей от температуры, а также проанализированы существующие формулы для расчета кинематической вязкости нефти в магистральных трубопроводах.

Течение жидкости в межтрубном пространстве теплообменника имеет сложный характер и зависит от многих факторов. Численное моделирование теплообмена в теплообменных устройствах различных конструкций проведено в работах [11, 12]. Результаты численных расчетов применяются для поиска оптимальных способов интенсификации процессов теплообмена [13–15]. Полученные результаты указывают на уменьшение влияния вязкости перекачиваемой нефти на гидравлическую характеристику трубопровода при перекачке на развитых турбулентных режимах.

Сравнение точности различных моделей турбулентности, используемых для замыкания уравнений Рейнольдса, выполнено в [16–19]. В основном в расчетах используются  $k$ - $\epsilon$  и  $k$ - $\omega$  модели турбулентности, а также Shear Stress Transport (SST) модель  $k$ - $\omega$ . Полученные результаты сравнения моделей позволяют добиться удовлетворительного совпадения результатов расчетов с данными промышленных экспериментов. В результате возможно решить задачи управления процессами, повысить эффективность производства и определить оптимальные режимные параметры технологических процессов. Вместе с тем имеющиеся расчеты с использованием двухпараметрических моделей турбулентности не учитывают ламинарно-турбулентный переход, что сказывается на определении эффективной длины теплообменного аппарата [20].

В классических теплообменных аппаратах пучок труб для одного теплоносителя помещается внутри кожуха, по которому движется другой теплоноситель. В конструкции геликоидных теплообменников применяются профилированные трубки и ребра винтового профиля, с помощью которых улучшаются условия теплообмена. Трубки в таких аппаратах имеют малый

диаметр и тонкие стенки (около 0,3 мм). В случае, когда вязкость зависит от температуры, режим течения в таких тонких трубках может изменяться от ламинарного до турбулентного. Обычно точка перехода ламинарного режима течения в турбулентный определяется по косвенным признакам, например, по резкому изменению коэффициента трения.

В настоящей работе разработана математическая модель теплообменного аппарата, учитывающая ламинарно-турбулентный переход. Выбрана схема теплообменника «труба в трубе» с тонкой и гладкой внутренней трубкой. Приведен метод расчета теплообменника прямооточного типа, в котором во внутреннем трубопроводе рабочей жидкостью является нефть (холодный теплоноситель), а во внешней трубе — вода (горячий теплоноситель). Расчеты проведены для модельной конструкции теплообменного аппарата на базе двух подходов, основанных на методе среднелогарифмической разницы температур (Log-Mean Temperature Difference, LMTD) при постоянной и переменной вязкости и на основе средств вычислительной гидродинамики (Computational Fluid Dynamics, CFD). Осредненные по Рейнольдсу уравнения Навье–Стокса замкнуты при помощи модели турбулентности, учитывающей ламинарно-турбулентный переход. Выполнено сравнение данных, полученных в рамках двух подходов.

### Математическая модель

В расчетах использованы методы: среднелогарифмической разницы температур при постоянной и переменной вязкости; вычислительной гидродинамики.

**Метод среднелогарифмической разницы температур.** Для оценки тепловых потоков от горячего теплоносителя к холодному применяется модель теплоносителя с постоянной или переменной по длине вязкостью, основанная на использовании среднелогарифмической разницы температур. В случае сильной зависимости вязкости от температуры теплообменник разбивается на элементарные участки по длине. На каждом участке делается предположение о малом изменении вязкости.

Число Нуссельта зависит от режима течения (ламинарный, переходный, турбулентный) и режима теплообмена (нагрев или охлаждение). Локальные числа Нуссельта в ламинарном и турбулентном режимах течения рассчитываются при помощи соотношений, приведенных в работах [21–23]. Для кольцевого канала в турбулентном режиме течения используется соотношение, как и для течения в трубе, но со своим эквивалентным гидравлическим диаметром. Переход между различными критериальными соотношениями, соответствующими ламинарному, переходному и турбулентному режимам течения, реализуется в зависимости от числа Рейнольдса, которое находится по среднemasовой скорости (при  $Re \sim 2 \cdot 10^3$  и  $Re \sim 10^4$ ).

**Метод вычислительной гидродинамики.** Результаты расчетов при помощи метода среднелогарифмической разницы температур сравниваются с данными, полученными методами вычислительной гидродинамики. Течение считается стационарным и осесимметричным,

нефть — ньютоновской жидкостью с постоянной плотностью и переменными теплофизическими свойствами. Расчеты проводятся при помощи численного решения осредненных по Рейнольдсу уравнений Навье–Стокса (Reynolds-Averaged Navier–Stokes, RANS) для вязкой несжимаемой жидкости, замкнутой благодаря модели турбулентности, учитывающей ламинарно-турбулентный переход [24].

Модель турбулентности SST  $k-\omega$  разработана для эффективного сочетания надежной и точной модели  $k-\omega$  в пристеночной области и модели  $k-\epsilon$  в свободном потоке [25, 26]. Для переключения между моделями применяется специальная функция, которая принимает единичное значение в пристеночной области (стандартная модель  $k-\omega$ ) и нулевое значение вдали от стенки (модель  $k-\epsilon$ ).

Модель, учитывающая ламинарно-турбулентный переход (Local-Correlation Transition Model,  $\gamma-Re_{\theta t}$  transition model), основана на сочетании уравнений SST  $k-\omega$  модели турбулентности с двумя дополнительными уравнениями переноса для параметра перемежаемости  $\gamma$  и критического числа Рейнольдса  $Re_{\theta t}$ , построенного по толщине потери импульса [27, 28]. Для упрощения модели уравнение для  $Re_{\theta t}$  не рассматривается, а в уравнении для параметра перемежаемости делается предположение о малости конвективных слагаемых [29]. Такой подход приводит к алгебраическим соотношениям для нахождения параметра перемежаемости.

Модель турбулентности  $\gamma-Re_{\theta t}$ , используемая в расчетах, была разработана для решения задач внешней аэродинамики. В дальнейшем модель успешно применялась для моделирования широкого круга течений. В частности, в работе [30] проведено сравнение различных моделей турбулентности ( $k-\epsilon$ ,  $k-\omega$ ,  $k-\omega$  SST,  $v^2-f$ ,  $\gamma-Re_{\theta t}$ ) применительно к течениям в каналах и показаны преимущества модели  $\gamma-Re_{\theta t}$  перед другими моделями турбулентности.

В научных работах применяются различные зависимости вязкости от температуры. В нефтяной отрасли при расчете кинематической вязкости  $\nu$ , зависящей от температуры  $T$ , используется формула Вальтера [10]

$$\lg[\lg(\nu + 0,8)] = a + b \lg T, \quad (1)$$

где  $a$  и  $b$  — эмпирические коэффициенты, определяемые для исследуемой жидкости экспериментальным путем. Коэффициенты  $a$  и  $b$  в формуле (1) находятся из соотношений

$$a = \lg[\lg(\nu_1 + 0,8)] - b \lg T,$$

$$b = \frac{\lg[\lg(\nu_1 + 0,8)] - \lg[\lg(\nu_2 + 0,8)]}{\lg T_1 - \lg T_2},$$

где  $\nu_1$  и  $\nu_2$  — значения кинематической вязкости жидкости при температурах  $T_1$  и  $T_2$ .

### Численное моделирование

Реализация метода среднелогарифмической разницы температур заключается в интегрировании системы обыкновенных дифференциальных уравнений.

Уравнения решены методом конечных разностей на интервале  $x \in [0, L]$ , где  $L$  — длина теплообменника, что позволяет применить разработанный подход для расчета прямоточных и противоточных теплообменников. Для стабилизации итерационного процесса при линейаризации системы применен метод нижней релаксации. Длина промежутка интегрирования  $L$  заранее неизвестна, которая определена методом Ньютона.

Выполнена дискретизация осредненных по Рейнольдсу уравнений Навье–Стокса и уравнений модели турбулентности с помощью методов конечных объемов на неструктурированных сетках и коррекции давления Semi-Implicit Method for Pressure Linked Equations (SIMPLE) [31]. Дискретизация невязких потоков осуществлена с использованием монотонной противоточной схемы для законов сохранения (Monotonic Upstream Schemes for Conservation Laws, MUSCL), а вязких потоков — центрированная схема второго порядка точности. Схема MUSCL позволила повысить порядок аппроксимации по пространственным переменным без потери монотонности решения, соответствует условию Total Variation Diminishing (TVD) и представляет собой комбинацию центрированных конечных разностей второго порядка и диссипативного члена, для переключения между которыми служит ограничитель потока, построенный на основе характеристических переменных. Для решения системы разностных уравнений использован геометрический многосеточный метод [31].

Для расчетов ядра потока выбрана сетка, состоящая из 19 461 ячеек, из которых  $500 \times 24$  ячеек размещены в области, заполненной нефтью,  $500 \times 5$  ячеек — в области из стали, а  $500 \times 13$  — в области, заполненной водой. Сгущение ячеек сетки произведено около стенок трубы таким образом, чтобы  $y^+ < 2$ , где  $y^+$  — безразмерная пристеночная координата. Для этого в пристеночном слое размещено 15 узлов сетки, расстояние между которыми изменяется по закону геометрической прогрессии.

### Результаты и обсуждение

В качестве примера рассмотрим нефть Узеньского нефтегазового месторождения, расположенного в Мангистауской области Казахстана. Залежи нефти находятся на глубине 0,9–2,4 км. Нефть имеет следующие параметры: плотность составляет 844–874 кг/м<sup>3</sup>; вязкость — 3,4–8,15 мПа·с; содержание серы — 0,16–2 %, парафинов 16–22 % и смол — 8–20 %.

Выполним сравнение результатов расчетов по формуле Вальтера с экспериментальными значениями динамической вязкости  $\mu = \rho\nu$ , где  $\rho$  — плотность (рис. 1) для нефти Узеньского месторождения. Температура изменяется от 283 до 383 К.

Для расчета эффективной длины теплообменника и температуры горячего теплоносителя на выходе заданы следующие параметры:  $T_{ci} = 303$  К и  $T_{co} = 328$  К (входная и выходная температуры холодного теплоносителя);  $T_{hi} = 423$  К (входная температура горячего теплоносителя);  $v_c = 4$  мс и  $v_h = 4$  мс (скорости холодного и горячего теплоносителя);  $G_c = 0,3814$  кг/с и  $G_h = 0,6386$  кг/с

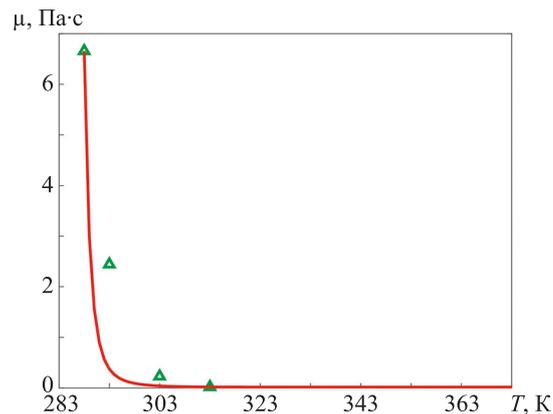


Рис. 1. Зависимость динамической вязкости нефти Узеньского месторождения от температуры.

Экспериментальные данные [20] (треугольные значки) и расчеты по формуле Вальтера (1) (сплошная линия)

Fig. 1. The dependence of the dynamic viscosity of oil from the Uzen field vs. temperature. Triangular symbols correspond to experimental data [20], solid line corresponds to calculations by the Walther formula (1)

(расходы горячего и холодного теплоносителя); геометрические характеристики (толщина стенок трубок  $\delta_w = 1$  мм, внутренний и внешний диаметры трубок  $d = 12$  мм и  $d_c = 14$  мм, внутренний диаметр кожуха  $D = 20$  мм); физические свойства материала трубок.

В результате решения уравнения теплового баланса методом конечных разностей получено, что длина теплообменника составляет  $L = 5,28$  м (для постоянной вязкости) и  $L = 4,26$  м (для переменной вязкости). При этом температура горячего теплоносителя на выходе достигла 416 К.

Выполним сравнение полученных результатов с данными численного моделирования. На рис. 2, а показаны распределения среднемассовой температуры нефти (холодного теплоносителя) ( $T_m$ ) вдоль длины теплообменника, полученные при помощи метода конечных разностей и на основе численного моделирования. Среднемассовая температура нефти возрастает вдоль длины за счет нагревания от источника тепла (горячего теплоносителя). Результаты аналитических и численных расчетов достаточно хорошо согласуются между собой.

Распределения среднемассовой температуры воды (горячего теплоносителя) вдоль длины, полученные на основе аналитических и численных расчетов, показаны на рис. 2, б. Среднемассовая температура воды убывает вдоль длины трубы за счет передачи тепла от нее к холодному теплоносителю. Следует отметить хорошее согласование результатов расчетов, полученных на основе различных подходов.

На рис. 2 видны характерные изменения кривизны линий на расстоянии порядка 2,5 м от входного сечения, где заметны резкие изменения градиентов температур. Изменения происходят на расстоянии, на котором ламинарный режим течения переходит в турбулентный. На рис. 3 приведены графики изменения числа Рейнольдса вдоль длины теплоносителей, полученные при аналитическом расчете с моделью переменной

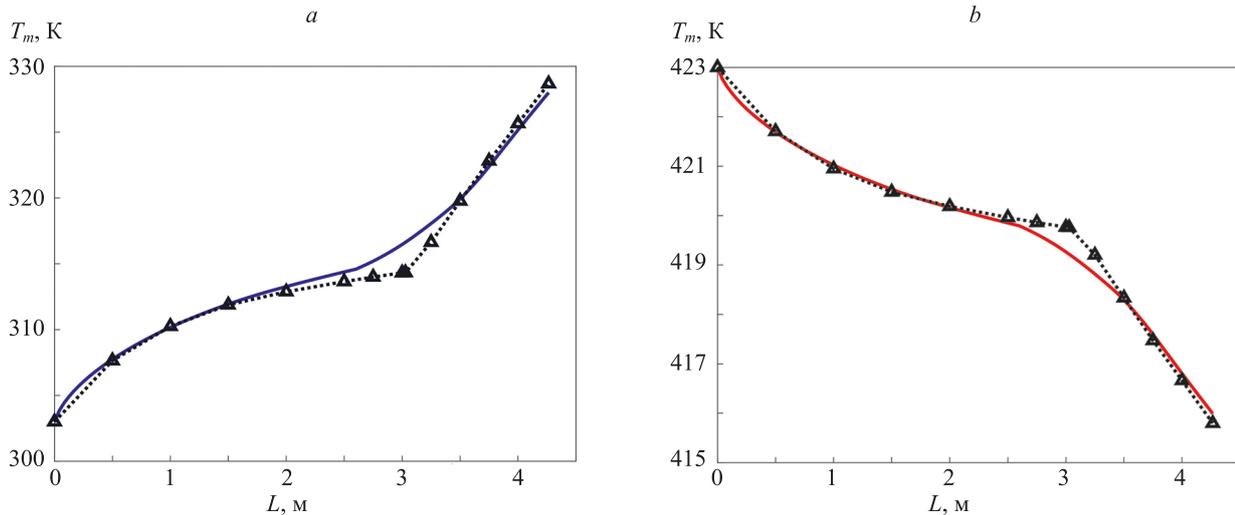


Рис. 2. Распределения среднемассовых температур нефти (а) и воды (б) вдоль длины теплообменника. Результаты, полученные на основе теоретического подхода (сплошная линия) и численных расчетов (пунктирная линия с треугольными значками)

Fig. 2. Distribution of average mass temperatures of oil (a) and water (b) along the length of the heat exchanger. The results obtained on the basis of a theoretical approach (solid line) and numerical calculations (dotted line with triangular icons)

вязкости. Для воды число Рейнольдса практически не меняется при постоянном значении вязкости и незначительном изменении температуры. Обратная картина наблюдается для холодного теплоносителя (нефти). Число Рейнольдса существенно возрастает, что связано с резким уменьшением вязкости нефти при ее практически неизменной плотности. Отметим, что наблюдается переходный участок на расстоянии порядка от 1,85 до 3,8 м, где происходит переход ламинарного течения в турбулентный.

На рис. 4 приведен график изменения числа Рейнольдса нефти (холодного теплоносителя) вдоль

длины теплообменного аппарата при переменных значениях вязкости нефти (линия 1). Видно, что значения числа Рейнольдса резко возрастают вдоль продольной координаты, что объясняется уменьшением вязкости нефти за счет ее нагревания. На рисунке приведен также график поведения числа Рейнольдса нефти при средней температуре теплоносителя (линия 2). При постоянной вязкости число Рейнольдса холодного теплоносителя постоянно и не может правильно охарактеризовать теплообмен в данном теплообменном аппарате.

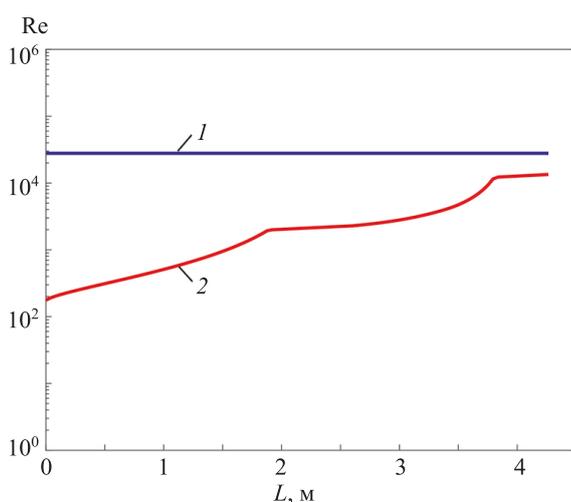


Рис. 3. Распределения чисел Рейнольдса по длине, полученные при аналитическом расчете для воды (линия 1) и нефти (линия 2)

Fig. 3. Length distributions of Reynolds numbers obtained in an analytical calculation. Line 1 corresponds to water and line 2 to oil

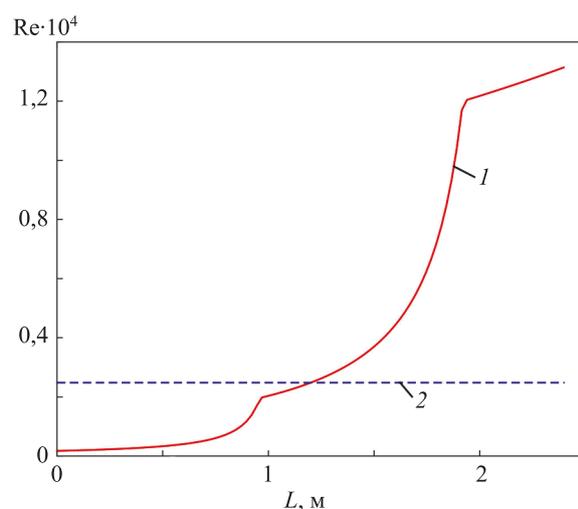


Рис. 4. Распределения чисел Рейнольдса нефти вдоль длины теплоносителя при численном расчете при переменной вязкости (линия 1) и теоретическом расчете (линия 2)

Fig. 4. Distributions of oil Reynolds numbers along length. Line 1 corresponds to the numerical calculation at variable viscosity, and line 2 corresponds to the theoretical calculation

### Заклучение

Снижение вязкости нефти при нагреве является одним из способов повышения энергоэффективности процесса перекачки высоковязкой нефти при добыче и транспортировке. Численное моделирование позволяет сократить сроки и трудоемкость испытаний новых конструкций теплообменных аппаратов в нефтегазовой отрасли.

Проведено численное моделирование кожухотрубчатого теплообменника средствами вычислительной гидродинамики. Получены распределения скорости и

температуры в теплообменном аппарате. Выполнено сравнение результатов численных расчетов с данными, полученными на основе метода среднелогарифмической разницы температур в случае постоянной и переменной вязкости. При переменной вязкости нефти осуществлен переход от ламинарного режима в турбулентный. Результаты эксперимента показали, что модель с постоянной вязкостью приводит к заниженной примерно на 20 % длине теплообменного аппарата по сравнению с расчетами, учитывающими зависимость вязкости нефти от температуры.

### Литература

1. Nakaso K., Mitani H., Fukai J. Convection heat transfer in a shell-and-tube heat exchanger using sheet fins for effective utilization of energy // *International Journal of Heat and Mass Transfer*. 2015. V. 82. P. 581–587. <https://doi.org/10.1016/j.ijheatmasstransfer.2014.11.033>
2. Wang Y., Gu X., Jin Z., Wang K. Characteristics of heat transfer for tube banks in crossflow and its relation with that in shell-and-tube heat exchangers // *International Journal of Heat and Mass Transfer*. 2016. V. 93. P. 584–594. <https://doi.org/10.1016/j.ijheatmasstransfer.2015.10.018>
3. Araavind S., Athreya A.S. CFD analysis of shell and tube heat exchanger for pre-heating of biodiesel // *International Journal of Advance Research in Science and Engineering*. 2017. V. 6. N 7. P. 687–693.
4. Abda A.A., Kareema M.Q., Najib S.Z. Performance analysis of shell and tube heat exchanger: parametric study // *Case Studies in Thermal Engineering*. 2018. V. 12. P. 563–568. <https://doi.org/10.1016/j.csite.2018.07.009>
5. Лычаков В.Д., Егоров М.Ю., Щеглов А.А., Сиволов А.С., Матяш А.С., Балунов Б.Ф. Анализ теплоотдачи сборок оребренных теплообменных элементов // *Теплоэнергетика*. 2022. № 3. С. 63–71.
6. Darbandi M., Abdollahpour M.-S., Hasanpour-Matkolaei M. A new developed semi-full-scale approach to facilitate the CFD simulation of shell and tube heat exchangers // *Chemical Engineering Science*. 2021. V. 245. P. 116836. <https://doi.org/10.1016/j.ces.2021.116836>
7. Bizhan K.G., Mohammad R.D., Hossein P. Prediction of kinematic viscosity of petroleum fractions using artificial neural networks // *Iranian Journal of Oil and Gas Science and Technology*. 2014. V. 3. N 2. P. 51–65.
8. Boda M.A., Bhasagi P.N., Sawade A.S., Andodgi R.A. Analysis of kinematic viscosity for liquids by varying temperature // *International Journal of Innovative Research in Science, Engineering and Technology*. 2015. V. 4. N 4. P. 1951–1954. <https://doi.org/10.15680/IJRSET.2015.0404020>
9. Lee B.I., Kesler M.G. A generalized thermodynamic correlation based on three-parameter corresponding states // *AIChE Journal*. 1975. V. 21. N 3. P. 510–527. <https://doi.org/10.1002/aic.690210313>
10. Аралов О.В., Буянов И.В., Саванин А.С., Иорданский Е.И. Исследование методов расчета кинематической вязкости нефти в магистральном нефтепроводе // *Наука и технологии трубопроводного транспорта нефти и нефтепродуктов*. 2017. Т. 7. № 5. С. 97–105.
11. Yogesh S.S., Selvaraj A.S., Ravi D.K., Rajagopal T.K.R. Heat transfer and pressure drop characteristics of inclined elliptical fin tube heat exchanger of varying ellipticity ratio using CFD code // *International Journal of Heat and Mass Transfer*. 2018. V. 119. P. 26–39. <https://doi.org/10.1016/j.ijheatmasstransfer.2017.11.094>
12. Chen K., Mohammed H.I., Mahdi J.M., Rahbari A., Cairns A., Talebizadehsardari P. Effects of non-uniform fin arrangement and size on the thermal response of a vertical latent heat triple-tube heat exchanger // *Journal of Energy Storage*. 2022. V. 45. P. 103723. <https://doi.org/10.1016/j.est.2021.103723>
13. Osley W.G., Droegemueller P., Ellerby P. CFD investigation of heat transfer and flow patterns in tube side laminar flow and the potential for enhancement // *Chemical Engineering Transactions*. 2013. V. 35. P. 997–1002. <https://doi.org/10.3303/CET1335166>

### References

1. Nakaso K., Mitani H., Fukai J. Convection heat transfer in a shell-and-tube heat exchanger using sheet fins for effective utilization of energy. *International Journal of Heat and Mass Transfer*, 2015, vol. 82, pp. 581–587. <https://doi.org/10.1016/j.ijheatmasstransfer.2014.11.033>
2. Wang Y., Gu X., Jin Z., Wang K. Characteristics of heat transfer for tube banks in crossflow and its relation with that in shell-and-tube heat exchangers. *International Journal of Heat and Mass Transfer*, 2016, vol. 93, pp. 584–594. <https://doi.org/10.1016/j.ijheatmasstransfer.2015.10.018>
3. Araavind S., Athreya A.S. CFD analysis of shell and tube heat exchanger for pre-heating of biodiesel. *International Journal of Advance Research in Science and Engineering*, 2017, vol. 6, no. 7, pp. 687–693.
4. Abda A.A., Kareema M.Q., Najib S.Z. Performance analysis of shell and tube heat exchanger: parametric study. *Case Studies in Thermal Engineering*, 2018, vol. 12, pp. 563–568. <https://doi.org/10.1016/j.csite.2018.07.009>
5. Lychakov V.D., Scheglov A.A., Sivovolov A.S., Matyash A.S., Balunov B.F., Egorov M.Y. An analysis of heat transfer in assemblies of finned heat-exchange elements. *Thermal Engineering*, 2022, vol. 69, no. 3, pp. 202–209. <https://doi.org/10.1134/S0040601522030089>
6. Darbandi M., Abdollahpour M.-S., Hasanpour-Matkolaei M. A new developed semi-full-scale approach to facilitate the CFD simulation of shell and tube heat exchangers. *Chemical Engineering Science*, 2021, vol. 245, pp. 116836. <https://doi.org/10.1016/j.ces.2021.116836>
7. Bizhan K.G., Mohammad R.D., Hossein P. Prediction of kinematic viscosity of petroleum fractions using artificial neural networks. *Iranian Journal of Oil and Gas Science and Technology*, 2014, vol. 3, no. 2, pp. 51–65.
8. Boda M.A., Bhasagi P.N., Sawade A.S., Andodgi R.A. Analysis of kinematic viscosity for liquids by varying temperature. *International Journal of Innovative Research in Science, Engineering and Technology*, 2015, vol. 4, no. 4, pp. 1951–1954. <https://doi.org/10.15680/IJRSET.2015.0404020>
9. Lee B.I., Kesler M.G. A generalized thermodynamic correlation based on three-parameter corresponding states. *AIChE Journal*, 1975, vol. 21, no. 3, pp. 510–527. <https://doi.org/10.1002/aic.690210313>
10. Aralov O.V., Buyanov I.V., Savanin A.S., Iordansky E.I. Research of methods for oil kinematic viscosity calculation in the oil-trunk pipeline. *Science & Technologies: Oil and Oil Products Pipeline Transportation*, 2017, vol. 7, no. 5, pp. 97–105. (in Russian)
11. Yogesh S.S., Selvaraj A.S., Ravi D.K., Rajagopal T.K.R. Heat transfer and pressure drop characteristics of inclined elliptical fin tube heat exchanger of varying ellipticity ratio using CFD code. *International Journal of Heat and Mass Transfer*, 2018, vol. 119, pp. 26–39. <https://doi.org/10.1016/j.ijheatmasstransfer.2017.11.094>
12. Chen K., Mohammed H.I., Mahdi J.M., Rahbari A., Cairns A., Talebizadehsardari P. Effects of non-uniform fin arrangement and size on the thermal response of a vertical latent heat triple-tube heat exchanger. *Journal of Energy Storage*, 2022, vol. 45, pp. 103723. <https://doi.org/10.1016/j.est.2021.103723>
13. Osley W.G., Droegemueller P., Ellerby P. CFD investigation of heat transfer and flow patterns in tube side laminar flow and the potential

14. Karar O., Emami S., Gounder S.M., Myo Thant M.M., Mukhtar H., Shariffpur M., Sadeghzadeh M. Experimental and numerical investigation on convective heat transfer in actively heated bundle-pipe // *Engineering Applications of Computational Fluid Mechanics*. 2021. V. 15. N 1. P. 848–864. <https://doi.org/10.1080/19942060.2021.1920466>
15. Rana S., Zunaid M., Kumar R. CFD approach for the enhancement of thermal energy storage in phase change material charged heat exchanger // *Case Studies in Thermal Engineering*. 2022. V. 33. P. 101921. <https://doi.org/10.1016/j.csite.2022.101921>
16. Allouche Y., Varga S., Bouden C., Oliveira A.C. Validation of a CFD model for the simulation of heat transfer in a tubes-in-tank PCM storage unit // *Renewable Energy*. 2016. V. 89. P. 371–379. <https://doi.org/10.1016/j.renene.2015.12.038>
17. Balaji D., Prakash L.S.S. CFD analysis of a pressure drop in a staggered tube bundle for a turbulent cross flow // *International Advanced Research Journal in Science, Engineering and Technology*. 2016. V. 3. N 2. P. 35–40. <https://doi.org/10.17148/IARJSET.2016.3209>
18. Czarnota T., Wagner C. Turbulent convection and thermal radiation in a cuboidal Rayleigh–Bénard cell with conductive plates // *International Journal of Heat and Fluid Flow*. 2016. V. 57. P. 150–172. <https://doi.org/10.1016/j.ijheatfluidflow.2015.10.006>
19. Mohanan A.K., Prasad B.V., Vengadesan S. Flow and heat transfer characteristics of a cross-flow heat exchanger with elliptical tubes // *Heat Transfer Engineering*. 2021. V. 42. N 21. P. 1846–1860. <https://doi.org/10.1080/01457632.2020.1826742>
20. Тугунов П.И., Новоселов В.Ф., Коршак А.А., Шаммазов А.М. Типовые расчеты при проектировании и эксплуатации нефтебаз и нефтепроводов. М.: Дизайн Полиграф Сервис, 2002. 234 с.
21. Кириллов П.Л., Юрьев Ю.С., Бобков В.П. Справочник по теплогидравлическим расчетам (ядерные реакторы, теплообменники, парогенераторы). М.: Энергоатомиздат, 1990. 360 с.
22. Петухов Б.С., Генин Л.Г., Ковалев С.А. Теплообмен в ядерных энергетических установках. М.: Атомиздат, 1974. 408 с.
23. Кутателадзе С.С. Теплопередача и гидродинамическое сопротивление. М.: Энергоатомиздат, 1990. 367 с.
24. 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>
25. Menter F., Esch T., Kubacki S. Transition modelling based on local variables // *Engineering Turbulence Modelling and Experiments*. Elsevier, 2002. P. 555–564. <https://doi.org/10.1016/b978-008044114-6/50053-3>
26. Langtry R.B., Menter F.R. Correlation-based transition modeling for unstructured parallelized computational fluid dynamics codes // *AIAA Journal*. 2009. V. 47. N 12. P. 2894–2906. <https://doi.org/10.2514/1.42362>
27. Coder J.G., Maughmer M.D. Computational fluid dynamics compatible transition modeling using an amplification factor transport equation // *AIAA Journal*. 2014. V. 52. N 11. P. 2506–2512. <https://doi.org/10.2514/1.j052905>
28. Menter F.R., Smirnov P.E., Liu T., Avancha R. A one-equation local correlation-based transition model // *Flow, Turbulence and Combustion*. 2015. V. 95. N 4. P. 583–619. <https://doi.org/10.1007/s10494-015-9622-4>
29. Gorji S., Seddighi M., Ariyaratne C., Vardy A.E., O'Donoghue T., Pokrajac D., He S. A comparative study of turbulence models in a transient channel flow // *Computers and Fluids*. 2014. V. 83. P. 111–123. <https://doi.org/10.1016/j.compfluid.2013.10.037>
30. Volkov K. Numerical analysis of Navier–Stokes equations on unstructured meshes // *Handbook on Navier–Stokes Equations: Theory and Analysis*. Nova Science, 2016. P. 365–442.
31. Volkov K. Multigrid and preconditioning techniques in CFD applications // *CFD Techniques and Thermo-Mechanics Applications*. Springer, 2018. P. 83–149. [https://doi.org/10.1007/978-3-319-70945-1\\_6](https://doi.org/10.1007/978-3-319-70945-1_6)
- for enhancement. *Chemical Engineering Transactions*, 2013, vol. 35, pp. 997–1002. <https://doi.org/10.3303/CET1335166>
14. Karar O., Emami S., Gounder S.M., Myo Thant M.M., Mukhtar H., Shariffpur M., Sadeghzadeh M. Experimental and numerical investigation on convective heat transfer in actively heated bundle-pipe. *Engineering Applications of Computational Fluid Mechanics*, 2021, vol. 15, no. 1, pp. 848–864. <https://doi.org/10.1080/19942060.2021.1920466>
15. Rana S., Zunaid M., Kumar R. CFD approach for the enhancement of thermal energy storage in phase change material charged heat exchanger. *Case Studies in Thermal Engineering*, 2022, vol. 33, pp. 101921. <https://doi.org/10.1016/j.csite.2022.101921>
16. Allouche Y., Varga S., Bouden C., Oliveira A.C. Validation of a CFD model for the simulation of heat transfer in a tubes-in-tank PCM storage unit. *Renewable Energy*, 2016, vol. 89, pp. 371–379. <https://doi.org/10.1016/j.renene.2015.12.038>
17. Balaji D., Prakash L.S.S. CFD analysis of a pressure drop in a staggered tube bundle for a turbulent cross flow. *International Advanced Research Journal in Science, Engineering and Technology*, 2016, vol. 3, no. 2, pp. 35–40. <https://doi.org/10.17148/IARJSET.2016.3209>
18. Czarnota T., Wagner C. Turbulent convection and thermal radiation in a cuboidal Rayleigh–Bénard cell with conductive plates. *International Journal of Heat and Fluid Flow*, 2016, vol. 57, pp. 150–172. <https://doi.org/10.1016/j.ijheatfluidflow.2015.10.006>
19. Mohanan A.K., Prasad B.V., Vengadesan S. Flow and heat transfer characteristics of a cross-flow heat exchanger with elliptical tubes. *Heat Transfer Engineering*, 2021, vol. 42, no. 21, pp. 1846–1860. <https://doi.org/10.1080/01457632.2020.1826742>
20. Tugunov P.I., Novoselov V.F., Korshak A.A., Shammazov A.M. *Standard Calculations for the Design and Operation Of Petroleum Storage Depots and Oil Pipelines*. Moscow, Dizajn Poligraf Servis Publ., 2002, 234 p. (in Russian)
21. Kirillov P.L., Iurev Iu.S., Bobkov V.P. *Handbook of Thermal-Hydraulic Calculations (Nuclear-Power Reactors, Heat Exchangers, Steam Generating Units)*. Moscow, Jenergoatomizdat, 1990, 360 p. (in Russian)
22. Petukhov B.S., Genin L.G., Kovalev S.A. *Heat transfer in nuclear power plants*. Moscow, Atomizdat Publ., 1974, 408 p. (in Russian)
23. Kutateladze S.S. *Heat Transfer and Hydrodynamical Resistance*. Moscow, Jenergoatomizdat Publ., 1990, 367 p. (in Russian)
24. 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>
25. Menter F., Esch T., Kubacki S. Transition modelling based on local variables. *Engineering Turbulence Modelling and Experiments*, Elsevier, 2002, pp. 555–564. <https://doi.org/10.1016/b978-008044114-6/50053-3>
26. Langtry R.B., Menter F.R. Correlation-based transition modeling for unstructured parallelized computational fluid dynamics codes. *AIAA Journal*, 2009, vol. 47, no. 12, pp. 2894–2906. <https://doi.org/10.2514/1.42362>
27. Coder J.G., Maughmer M.D. Computational fluid dynamics compatible transition modeling using an amplification factor transport equation. *AIAA Journal*, 2014, vol. 52, no. 11, pp. 2506–2512. <https://doi.org/10.2514/1.j052905>
28. Menter F.R., Smirnov P.E., Liu T., Avancha R. A one-equation local correlation-based transition model. *Flow, Turbulence and Combustion*, 2015, vol. 95, no. 4, pp. 583–619. <https://doi.org/10.1007/s10494-015-9622-4>
29. Gorji S., Seddighi M., Ariyaratne C., Vardy A.E., O'Donoghue T., Pokrajac D., He S. A comparative study of turbulence models in a transient channel flow. *Computers and Fluids*, 2014, vol. 83, pp. 111–123. <https://doi.org/10.1016/j.compfluid.2013.10.037>
30. Volkov K. Numerical analysis of Navier–Stokes equations on unstructured meshes. *Handbook on Navier–Stokes Equations: Theory and Analysis*. Nova Science, 2016, pp. 365–442.
31. Volkov K. Multigrid and preconditioning techniques in CFD applications. *CFD Techniques and Thermo-Mechanics Applications*, Springer, 2018, pp. 83–149. [https://doi.org/10.1007/978-3-319-70945-1\\_6](https://doi.org/10.1007/978-3-319-70945-1_6)

### Авторы

**Курманова Динара Есентаевна** — докторант, Евразийский национальный университет им. Л.Н. Гумилева, Астана, 010008, Республика Казахстан, <https://orcid.org/0009-0009-9787-7426>, [dikonya89\\_29@mail.ru](mailto:dikonya89_29@mail.ru)

**Джайчибеков Нурболат Жумабекович** — доктор физико-математических наук, профессор, профессор, Евразийский национальный университет им. Л.Н. Гумилева, Астана, 010008, Республика Казахстан, [sc 57195809348, https://orcid.org/0000-0002-3053-8288](https://orcid.org/0000-0002-3053-8288), [jaich@mail.ru](mailto:jaich@mail.ru)

*Статья поступила в редакцию 10.03.2023*  
*Одобрена после рецензирования 11.05.2023*  
*Принята к печати 16.07.2023*

### Authors

**Dinara E. Kurmanova** — Doctoral Student, L.N. Gumilyov Eurasian National University, Astana, 010008, Kazakhstan, <https://orcid.org/0009-0009-9787-7426>, [dikonya89\\_29@mail.ru](mailto:dikonya89_29@mail.ru)

**Nurbolat Zh. Jaichibekov** — D.Sc. (Physics & Mathematics), Full Professor, L.N. Gumilyov Eurasian National University, Astana, 010008, Kazakhstan, [sc 57195809348, https://orcid.org/0000-0002-3053-8288](https://orcid.org/0000-0002-3053-8288), [jaich@mail.ru](mailto:jaich@mail.ru)

*Received 10.03.2023*  
*Approved after reviewing 11.05.2023*  
*Accepted 16.07.2023*



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-828-835

УДК 629.7.017.1

## Подход к формированию обобщенных параметров технического состояния сложных технических систем с использованием нейросетевых структур

Георгий Николаевич Мальцев<sup>1</sup>, Виктор Леонидович Якимов<sup>2</sup>✉

<sup>1,2</sup> Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация

<sup>1</sup> [vka@mil.ru](mailto:vka@mil.ru), <http://orcid.org/0000-0002-6755-5700>

<sup>2</sup> [vka@mil.ru](mailto:vka@mil.ru) ✉, <http://orcid.org/0000-0001-9721-2453>

### Аннотация

**Введение.** Рассмотрен процесс формирования обобщенных параметров технического состояния сложных технических систем. Данный подход актуален для современных робототехнических комплексов, оснащенных встроенными средствами телеизмерений. **Метод.** Предложенный подход к формированию обобщенных параметров основан на взвешенном суммировании разнородных телеметрируемых параметров с использованием информации о структурных и функциональных связях в сложной технической системе с последующей цифровой низкочастотной фильтрацией результатов взвешенного суммирования. Такое решение позволяет повысить достоверность оценки технического состояния сложных технических систем по значениям обобщенных параметров во внешнем контуре управления. **Основные результаты.** Выбрана форма представления обобщенного параметра в виде градаций технического состояния, которое соответствует штатному функционированию, нештатным ситуациям и частично-работоспособному состоянию сложной технической системы. Разработана многоуровневая иерархическая модель формирования обобщенных параметров технического состояния сложной технической системы по данным телеизмерений на основе множества нейросетевых структур, позволяющих учесть нелинейный характер телеметрируемых параметров и взаимное влияние между ними. В модели применены множества цифровых низкочастотных фильтров, снижающих уровень возмущений во временных рядах обобщенных параметров. Возникновение возмущений связано с неопределенностью изменения значений телеметрируемых и обобщенных параметров вблизи границ допусков, задаваемых экспертами при расширенном количестве градаций технического состояния. Сведения о границах допусков характеризуют не только ситуации штатного и нештатного функционирования, но и частично-работоспособное состояние сложных технических систем. Представлены результаты формирования обобщенных параметров технического состояния бортовой системы космического аппарата с использованием многослойных нейронных сетей, полиномов Колмогорова–Габора и методов цифровой фильтрации. Показаны преимущества использования в разработанной модели многослойных нейронных сетей и медианных фильтров. **Обсуждение.** Использование обобщенных параметров позволит значительным образом снизить информационную нагрузку на каналы передачи телеметрической информации, а также средства ее обработки и анализа во внешнем контуре управления. Предлагаемые решения на основе базовых операций взвешенного суммирования и нелинейного преобразования могут быть эффективно реализованы на перспективных векторно-матричных и тензорных процессорах, поддерживающих их выполнение на аппаратном уровне.

### Ключевые слова

сложная техническая система, техническое состояние, телеметрируемые параметры, обобщенный параметр, нейронные сети, цифровая фильтрация

**Ссылка для цитирования:** Мальцев Г.Н., Якимов В.Л. Подход к формированию обобщенных параметров технического состояния сложных технических систем с использованием нейросетевых структур // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 828–835. doi: 10.17586/2226-1494-2023-23-4-828-835

## Approach to the generalized parameters formation of the complex technical systems technical condition using neural network structures

Georgy N. Maltsev<sup>1</sup>, Victor L. Yakimov<sup>2</sup>✉

<sup>1,2</sup> Mozhasky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation

<sup>1</sup> vka@mil.ru, <http://orcid.org/0000-0002-6755-5700>

<sup>2</sup> vka@mil.ru✉, <http://orcid.org/0000-0001-9721-2453>

### Abstract

A variant of the generalized parameters process forming of the complex technical systems technical condition is considered. This approach is relevant for modern robotic systems equipped with built-in telemetry tools. The proposed approach to the generalized parameters formation is based on heterogeneous telemetry parameters weighted summation using information about structural and functional relationships in a complex technical system, followed by digital low-frequency filtering of the weighted summation results. This solution makes it possible to increase the complex technical systems technical condition reliability assessment by the values of generalized parameters in the external control loop. The form of generalized parameter representation in the technical condition gradations, which corresponds to the normal functioning, emergency situations and complex technical system partially operational state, is chosen. A multilevel hierarchical model of the generalized parameters formation of the complex technical systems technical condition on the basis of telemeasurements based on a variety of neural network structures allowing to take into account the nonlinear nature of the parameters being telemetered and the mutual influence between them has been developed. The model uses a variety of digital low-frequency filters that reduce the level of disturbances in the generalized parameters time series. The occurrence of disturbances is associated with the uncertainty of changes in the values of the telemetered and generalized parameters near the tolerance limits set by experts with expanded gradations according to technical condition. Information about the limits of tolerances characterizes not only the situations of regular and non-standard functioning, but also the complex technical systems partially operational state. The results of the generalized parameters formation of the spacecraft onboard system technical condition using multilayer neural networks, Kolmogorov–Gabor polynomials and digital filtering methods are presented. The advantages of using multilayer neural networks and median filters in the developed model are shown. The use of generalized parameters will significantly reduce the information load on the transmission channels of telemetry information as well as the means of its processing and analysis in the external control loop. The proposed solutions based on the basic operations of weighted summation and nonlinear transformation can be effectively implemented on promising vector-matrix and tensor processors that support their execution at the hardware level.

### Keywords

technical system, technical condition, telemetric parameters, generalized parameter, neural networks, digital filtering

**For citation:** Maltsev G.N., Yakimov V.L. Approach to the generalized parameters formation of the complex technical systems technical condition using neural network structures. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 828–835 (in Russian). doi: 10.17586/2226-1494-2023-23-4-828-835

### Введение

Важнейшими задачами, решаемыми комплексами управления сложными техническими системами (СТС) на этапе их эксплуатации, являются задачи технического диагностирования, осуществляемого по результатам обработки телеметрируемых параметров [1]. На практике задачи технического диагностирования СТС характеризуются рядом проблемных вопросов, среди которых можно выделить значительную размерность, вследствие большого количества телеметрируемых параметров и возможных состояний элементов и подсистем СТС. Также оказывает влияние сложность выявления негативных тенденций изменения технического состояния для его средне- и долгосрочного прогнозирования. Последнее обстоятельство связано с двумя проблемами: сложностью прогнозирования большинства нелинейных телеметрируемых параметров на длительное время упреждения прогноза; потребностью представления результатов оценивания и прогнозирования технического состояния СТС в виде, позволяющем операторам внешней системы управления принимать решения по управлению СТС.

Один из способов решения указанных проблем — использование обобщенных параметров техническо-

го состояния СТС, удовлетворяющих требованиям как разработчиков, так и операторов внешних систем управления. Существует множество подходов к формированию обобщенных параметров на основе совокупности телеметрируемых параметров [2–6]. При этом в качестве обобщенных параметров могут выступать: характеристики аппаратных модулей СТС, результаты различного рода преобразований значений телеметрируемых параметров; непосредственно номер вида технического состояния. Однако большинство известных подходов к формированию обобщенных параметров не применимы на всех уровнях СТС, которые, как правило, имеют многоуровневую иерархическую структуру. Отметим, что формируемые обобщенные параметры не всегда поддаются прогнозированию на интересующих интервалах. Так, для бортовой аппаратуры некоторых космических аппаратов существует потребность прогнозирования технического состояния на интервалах до десятка суток и более [7].

В настоящей работе предложен подход к формированию обобщенных параметров в рамках многоуровневой иерархической модели, отражающей взаимосвязи на различных уровнях иерархии СТС [8]. При этом обобщенные параметры СТС на каждом уровне иерархии формируются с использованием нелинейных функци-

аналов в виде многослойных нейронных сетей, учитывающих характер телеметрируемых и обобщенных параметров элементов СТС предшествующих уровней иерархии и взаимосвязи между ними. Для повышения достоверности значений обобщенных параметров использованы цифровые фильтры. Адаптация весовых коэффициентов многослойных нейронных сетей и цифровых фильтров осуществлена при решении задачи оптимизации целевой функции, определяющей правильность нахождения обобщенных параметров в пределах допусков, заданных экспертами для каждого вида технического состояния на обучающей, проверочной и тестовой выборках данных телеизмерений. Важной особенностью решаемой задачи формирования обобщенных параметров является рассмотрение расширенного количества видов технического состояния, в том числе частично-работоспособного состояния СТС [7].

**Формализованное описание задачи формирования обобщенных параметров технического состояния сложных технических систем**

Использование обобщенных параметров для решения задач анализа технического состояния соответствует современной концепции построения СТС как робототехнических комплексов, для которых характерно сокращение интенсивности информационного обмена с внешними средствами управления. В табл. 1 представлены результаты сравнительного анализа способов формирования обобщенных параметров на основе телеметрируемых параметров [2–6]. Отметим необходимые свойства формируемых обобщенных параметров: соответствие значений реальному техническому состоянию СТС; медленно меняющийся характер, допускающий возможность прогнозирования значений во времени на длительное время упреждения прогноза; распределение значений в пределах допусков, соответствующих техническому состоянию СТС, близкое к нормальному; возможность определения по значениям моментов времени, соответствующих изменению технического состояния СТС. В ходе предварительных исследований выбран способ формирования обобщенных параметров на основе взвешенного суммирования (табл. 1), предполагающий использование операций нелинейного пре-

образования и цифровой фильтрации для обеспечения требуемой формы обобщенных параметров.

На рис. 1 представлена модель формирования обобщенных параметров, соответствующая многоуровневой иерархической структуре СТС, включающей  $m$  уровней иерархии.

На рис. 1 использованы следующие обозначения: ЦФНЧ — модули цифровой низкочастотной фильтрации; ФОПТС — модули формирования обобщенных параметров технического состояния;  $\mathbf{A}_{l,j_l}$  — вектора параметров модулей фильтрации телеметрируемых и обобщенных параметров;  $\mathbf{V}_{l,j_l}$  — вектора параметров модулей формирования обобщенных параметров;  $\mathbf{X}_{l,j_l}$  — вектора нефильтрованных телеметрируемых и обобщенных параметров технического состояния сложной технической системы на различных уровнях иерархии;  $\mathbf{X}_{l,j_l}^*$  — вектора фильтрованных телеметрируемых и обобщенных параметров технического состояния сложной технической системы на различных уровнях иерархии;  $l$  — номер уровня,  $l = 1, \dots, m$ ;  $s_l$  — номер вершины  $l$ -го уровня;  $N_l$  — количество вершин  $l$ -го уровня.

Проведенный анализ показал целесообразность использования следующих уровней: телеметрируемых параметров; аппаратных модулей; подсистем; функциональных систем и уровня СТС [8].

Вершины уровня 1 соответствуют телеметрируемым параметрам. Каждая вершина, начиная с уровня 2, характеризует соответствующее оборудование СТС. Для значений  $N_l$  должно выполняться условие:  $N_2 > N_3 > \dots > N_l > \dots > N_m$ , при этом, как правило,  $N_m = 1$ .

Уровень 1 модели формирования обобщенных параметров включает модули ЦФНЧ, в которых реализованы нормализация и цифровая фильтрация входных сигналов, представляющих собой отсчеты телеметрируемых параметров и признаки режимов функционирования аппаратных модулей от встроенных средств контроля в пределах «окна наблюдения» длиной  $L_\Phi - x_{1,j_1,k}, x_{1,j_1,k-1}, \dots, x_{1,j_1,k-L_\Phi+1}$  образующих вектора  $\mathbf{X}_{1,j_1,k}$ , где  $j_1$  — номер телеметрируемого параметра;  $k$  — номер дискретного момента времени. На выходах уровня 1 сформированы отфильтрованные значения телеметрируемых параметров  $x_{1,j_1,k}^* = \mathbf{F}_1(\mathbf{X}_{1,j_1,k}, \mathbf{A}_{1,j_1})$ ,

Таблица 1. Сравнительный анализ способов формирования обобщенных параметров технического состояния СТС  
Table 1. Comparative analysis of CTS technical condition generalized parameters formation methods

Способ	Применимость	Примеры обобщенных параметров	Недостатки
Аналитический	Уровень аппаратных модулей	Параметр деградации	Трудности получения аналитической модели
Взвешенного суммирования	На всех уровнях	Параметр деградации	Трудности определения весовых коэффициентов
Биективных преобразований	Уровень аппаратных модулей	Результат амплитудного уплотнения сигнальных параметров	Не применим к функциональным параметрам
Логических преобразований	На всех уровнях	Номер вида технического состояния	Трудности применения к динамическим системам
На основе описания аттрактора	Уровень систем и СТС	Показатель Херста, корреляционная энтропия	Применим к динамическим системам

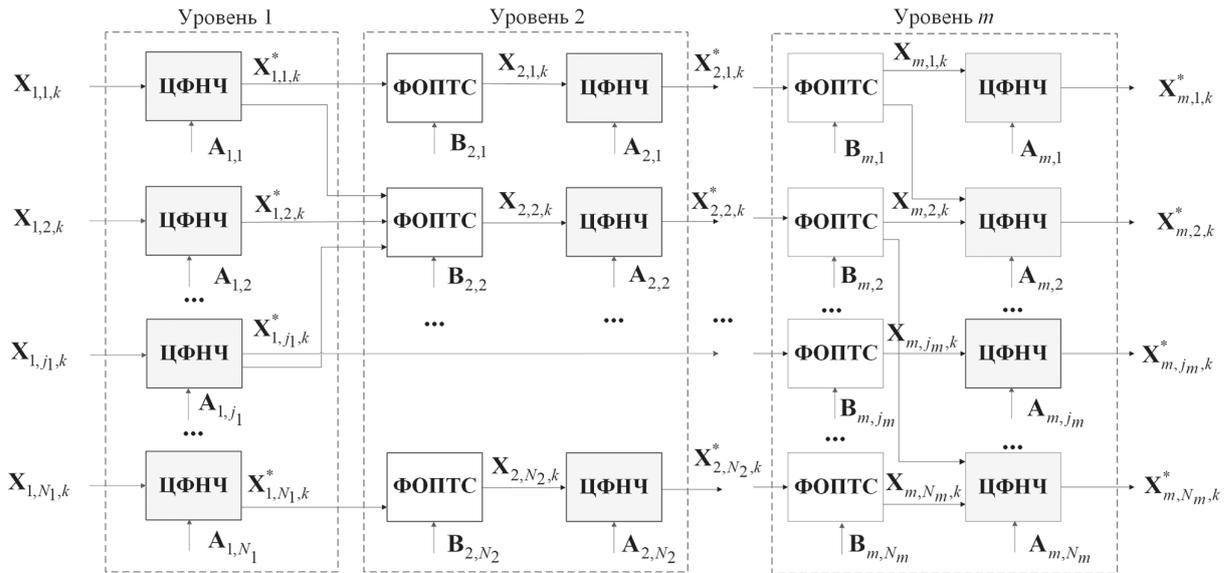


Рис. 1. Многоуровневая иерархическая модель формирования обобщенных параметров технического состояния сложных технических систем

Fig. 1. A multi-level hierarchical model of the technical condition generalized parameters formation for the complex technical systems

где  $F_1$  — оператор ЦФНЧ;  $A_{1,j_1} = [a_{1,j_1}]$  — вектор параметров модулей ЦФНЧ уровня 1.

Значения  $x_{1,j_1,k}^*$  рассмотрены в пределах «окна наблюдения» размерностью  $L$  и образуют вектора  $X_{1,j_1,k}^* = [x_{1,j_1,k}^*]$ . Все уровни иерархии модели включают модули ФОПТС и ЦФНЧ. Линии связи между модулями с первого уровня по уровень  $m$ , а также между модулями смежных уровней модели задаются экспертом.

Вектора отфильтрованных значений телеметрируемых параметров  $X_{1,j_1,k}^*$  и обобщенных параметров  $X_{l-1,j_{l-1},k}^*$  предшествующих уровней иерархии поступают в модули ФОПТС последующих уровней иерархии модели  $M_\Sigma$  с формированием на их выходе значений обобщенных параметров  $x_{l,j_1,k} = F_2(X_{l-1,j_{l-1},k}^*, \dots, X_{1,j_1,k}^*, B_{l,j_1})$ , где  $B_{l,j_1} = [b_{l,j_1}]$  — вектор параметров модулей ФОПТС  $l$ -го уровня;  $F_2$  — функционал нелинейного преобразования. Значения  $x_{l,j_1,k}$  рассматриваются в пределах «окна наблюдения» размерностью  $L_\Phi$  и образуют вектора  $X_{l,j_1,k} = [x_{l,j_1,k}]$ , которые поступают в модули ЦФНЧ, где происходит их фильтрация. Отфильтрованные значения обобщенных параметров  $x_{l,j_1,k}^* = F_1(X_{l,j_1,k}, A_{l,j_1})$ , где  $A_{l,j_1} = [a_{l,j_1}]$  — вектора параметров модулей ЦФНЧ  $l$ -го уровня рассматриваются в пределах «окна наблюдения» размерностью  $L$ . Представленная модель должна разрабатываться во внешнем контуре управления СТС и периодически перезаключаться во внутренний контур [8].

На рис. 2, а показан один из возможных форматов представления обобщенных параметров в виде зависимости его значений  $x_{l,j_1,k}^*$  от номера временного отсчета  $k$  с переходом из области штатного функционирования (ШФ) через область частично-работоспособного состояния (ЧРС) в область нештатной ситуации (НШС) [8, 9].

На рис. 2 использованы следующие обозначения:  $L_{ш}$  — размер шкалы обобщенного параметра;  $N_{тс}$  —

количество видов технического состояния СТС;  $E_\Phi$  — среднеквадратическая погрешность формирования обобщенного параметра технического состояния СТС.

Допуска на значения каждого обобщенного параметра  $D_{l,j_1,1}, \dots, D_{l,j_1,N_{тс}}$  принадлежат множеству  $\{D_{l,j_1}\}$ , характеризуют степень деградации СТС и могут пересекаться на границах ЧРС. Увеличение количества данных областей  $N_{тс}$  позволяет расширить возможности диагностирования СТС [9]. В случае непрогнозируемых отказов в СТС значения обобщенных параметров будут скачком переходить из области ШФ в область НШС.

При разработке диагностических комплексов существует необходимость максимизации апостериорной вероятности (достоверности) принятия решения о техническом состоянии СТС. Зависимость достоверности определения технического состояния СТС  $D$  от значения нормально распределенной в пределах допусков  $D_{l,j_1}$  погрешности формирования обобщенных параметров с параметрами матожидания  $\mu = 0$  и среднеквадратического отклонения  $\sigma$  представлена на рис. 2, б и может быть определена аналитически [10]:

$$D = 1 - \gamma(E_\Phi) = 1 - \left( \Phi\left(\frac{\mu + E_\Phi}{\sigma}\right) - \Phi\left(\frac{\mu - E_\Phi}{\sigma}\right) \right) = 1 - 2\Phi\left(\frac{6E_\Phi N_{тс}}{L_{ш}}\right), \quad (1)$$

где  $\gamma(E_\Phi)$  — компонента, характеризующая снижение достоверности определения вида технического состояния СТС по значениям обобщенных параметров вследствие наличия погрешности формирования  $E_\Phi$ ;  $\Phi(\bullet)$  — функция Лапласа. Из выражения (1) следует, что при формировании обобщенных параметров может быть

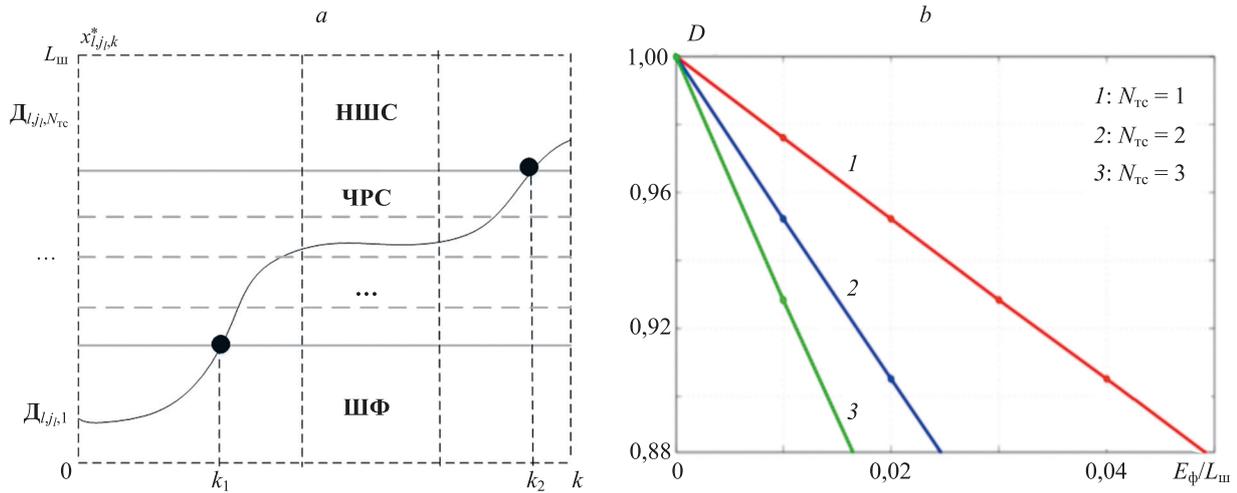


Рис. 2. Зависимости: значений обобщенных параметров от номера дискретного отсчета времени (а); достоверности оценки технического состояния сложных технических систем по обобщенным параметрам от нормированной к шкале погрешности их формирования при нормальном законе ее распределения (б)

Fig. 2. Dependences: of generalized parameters values vs. the number of discrete time reference (a); of reliability of the technical condition assessment for the complex technical systems according to the generalized parameters vs. the error of their formation normalized to the scale under the normal law of its distribution (b)

использован критерий минимума среднеквадратической погрешности формирования [11]:

$$E_{\phi}(\mathbf{A}_{l,j_l} \mathbf{B}_{l,j_l} \mathbf{X}_{1,j_1,k}) = \sqrt{\frac{1}{mMN} \sum_{l=2}^m \sum_{j_l=1}^{N_l} \sum_{k=1}^M (x_{l,j_1,k}^*(\mathbf{A}_{l,j_l} \mathbf{B}_{l,j_l} \mathbf{X}_{1,j_1,k}) - x_{l,j_1,k,тр}^*)^2}, \quad (2)$$

$$\{\mathbf{A}_{l,j_l} \mathbf{B}_{l,j_l}\}_{opt} = \arg \min_{\mathbf{A}_{l,j_l} \mathbf{B}_{l,j_l}} E_{\phi} \left| \begin{array}{l} \mathbf{X}_{1,j_1,k} \in \{D_{1,j_1}\}, \mathbf{X}_{1,j_1,k} \in \{D_{1,j_1}\}, \\ l \in [2, m], j_l \in [1, N_l], \end{array} \right.$$

где  $\mathbf{X}_{1,j_1,k,тр}^* = [x_{l,j_1,k,тр}^*]$  — вектор требуемых значений обобщенных параметров;  $\{\mathbf{A}_{l,j_l} \mathbf{B}_{l,j_l}\}_{opt}$  — множества векторов оптимальных значений параметров модулей ЦФНЧ и ФОПТС;  $\{D_{1,j_1}\}$  — множество допусков на  $j_1$ -е телеметрируемые параметры;  $M$  — количество отсчетов обобщенных параметров. Погрешность формирования обобщенных параметров  $E_{\phi}$  имеет закон распределения, отличный от нормального, ввиду наличия аномальных отсчетов в значениях обобщенных параметров, что связано с изменением значений телеметрируемых и обобщенных параметров вблизи допусков и их пересечением. При этом допуски на значения телеметрируемых параметров в различных технических состояниях не пересекаются только для ситуаций ШФ и НШС. Для ЧРС допуски могут пересекаться, особенно с увеличением количества градаций обобщенных параметров. В этой связи становится актуальным поиск адекватных методов формирования и фильтрации обобщенных параметров технического состояния.

Для реализации модулей ФОПТС предложено использовать многослойные нейронные сети с двумя скрытыми слоями, аппроксимационные возможности которых теоретически обоснованы и рассмотрены в работах [11–13]. На рис. 3 представлена структурная схема такой сети, имеющая  $n_1, n_2, n_3$  и  $n_4 = 1$  нейронов во входном, скрытых и выходном слое.

Рассмотрим связь между моделью на рис. 1 и схемой на рис. 3. Каждый вектор параметров  $\mathbf{B}_{l,j_l}$  модулей ФОПТС включает следующие параметры нейронной сети: весовые коэффициенты матриц  $\mathbf{W}_1 = [w_{1,i_1,i_2}]$ ,  $\mathbf{W}_2 = [w_{2,i_2,i_3}]$ ,  $\mathbf{W}_3 = [w_{3,i_3,i_4}]$ ; векторов постоянных смещений  $\mathbf{C}_1 = [c_{1,i_2}]$ ,  $\mathbf{C}_2 = [c_{2,i_3}]$ ,  $\mathbf{C}_3 = [c_{3,i_4}]$ , где  $i_1, i_2, i_3$  — номера нейронов в каждом слое нейронной сети, а также размер «окна наблюдения»  $L$ . Если вектор вход-

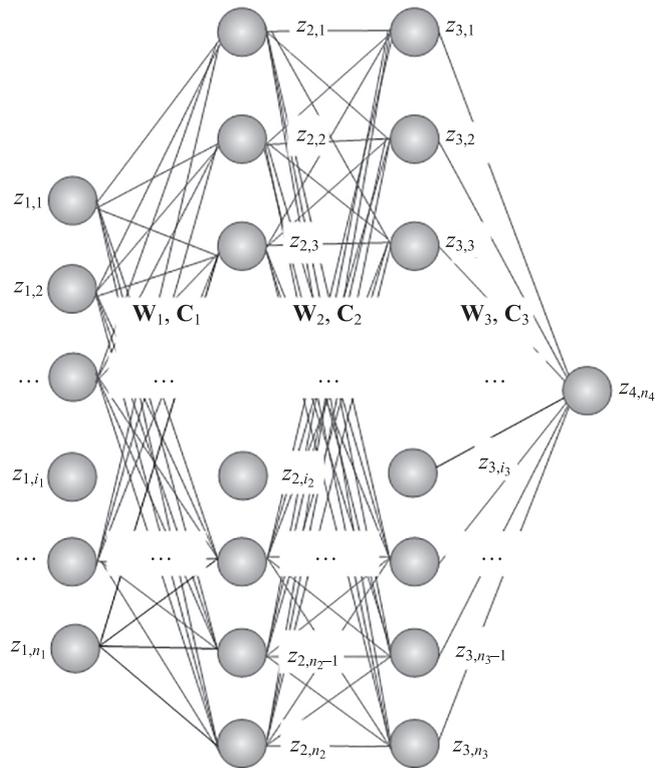


Рис. 3. Структурная схема нейронной сети  
Fig. 3. Structural diagram of the neural network

ных значений сети обозначить как  $\mathbf{Z}_{1,i_1} = [z_{1,i_1}]$ , где элементы  $z_{1,i_1}$  представляют собой элементы  $x_{l-1,j_{l-1},k}^*$  векторов входных значений  $\mathbf{X}_{l-1,j_{l-1},k}^*$  модулей ФОПТС, то значения нейронов на выходе каждого слоя сети [7, 11] имеют вид:

$$\begin{aligned} z_{2,i_2} &= f\left(\sum_{i_1=1}^{n_1} w_{1,i_1,i_2} z_{1,i_1} + c_{1,i_2}\right), \\ z_{3,i_3} &= f\left(\sum_{i_2=1}^{n_2} w_{2,i_2,i_3} z_{2,i_2} + c_{2,i_3}\right), \\ z_{4,i_4} &= f\left(\sum_{i_3=1}^{n_3} w_{3,i_3,i_4} z_{3,i_3} + c_{3,i_4}\right), \end{aligned}$$

где  $z_{4,i_4} = x_{l,j_{l-1},k}$  — выходное значение сети;  $f$  — нелинейная функция активации нейронов.

Весовые коэффициенты сети определим в процессе ее обучения и решения задачи (2) на обучающих, проверочных и тестовых данных, формируемых с использованием допусков  $\{D_{1,j_1}\}$  и  $\{D_{l,j_l}\}$ . Начальные значения весовых коэффициентов нейронной сети заданы эвристически случайным образом до начала обучения в диапазоне значений  $[-0,5; 0,5]$ . С целью сравнительного анализа результативности предлагаемого подхода при реализации модулей ФОПТС рассмотрены полиномы Колмогорова–Габора второго порядка [14]. Для нахождения весовых коэффициентов нейронной сети использован алгоритм наискорейшего спуска, для нахождения весовых коэффициентов полиномов — метод псевдоинверсии, а для определения оптимальных структур полиномов — алгоритм стохастической оптимизации [7, 11, 12, 15].

При реализации модулей ЦФНЧ каждый вектор параметров  $\mathbf{A}_{l,j_l}$  включает коэффициенты соответствующего цифрового фильтра (при наличии) и размер «окна наблюдения»  $L_\phi$ . Для реализации модулей ЦФНЧ исследованы следующие фильтры: с конечной импульсной характеристикой —  $x_{l,j_l,k}^* = \sum_{p=0}^{L_\phi-1} a_{l,j_l} x_{l,j_l,k-p}$ ; на основе

скользящего среднего —  $x_{l,j_l,k}^* = \frac{1}{L_\phi} \sum_{p=0}^{L_\phi-1} a_{l,j_l} x_{l,j_l,k-p}$ ; медианные фильтры —  $x_{l,j_l,k}^* = \text{med}(x_{l,j_l,k}, \dots, x_{l,j_l,k-L_\phi+1})$ .

### Результаты моделирования

На этапе моделирования в качестве СТС рассмотрена система приема и передачи информации космического аппарата, для которой был известен перечень телеметрируемых параметров и значения допусков в состояниях НС, ШФ и ЧРС.

При формировании обучающих пар векторов рандомизированным способом выбраны значения телеметрируемых и соответствующие им средние значения обобщенных параметров из заданных допусков в пределах «окна наблюдения». На рис. 4 показаны результаты формирования обобщенных параметров одного из аппаратных модулей — бортового передатчика, с использованием полиномов Колмогорова–Габора (рис. 4, а) и многослойной нейронной сети (рис. 4, б).

Кривые 1 представляют собой временные ряды сформированных, а кривые 2 — желаемых значений обобщенных параметров. Для приближения формы обобщенных параметров передатчика к требуемой выполнена цифровая фильтрация. Результаты формирования обобщенных параметров показаны на рис. 5 и приведены в табл. 2. Получены зависимости среднеквадратической погрешности формирования обобщенных параметров от номера цикла обучения нейронной сети на проверочной выборке ( $N_{ц}$ ) (рис. 5, а), также исходных (кривая 1) и отфильтрованных (кривая 2) значений обобщенных параметров передатчика на тестовой выборке от номера отсчета, полученная с использованием наилучшей нейронной сети и медианного фильтра (рис. 5, б).

Использование многослойных нейронных сетей и медианных фильтров позволило уменьшить погрешность формирования и количество аномальных отсче-

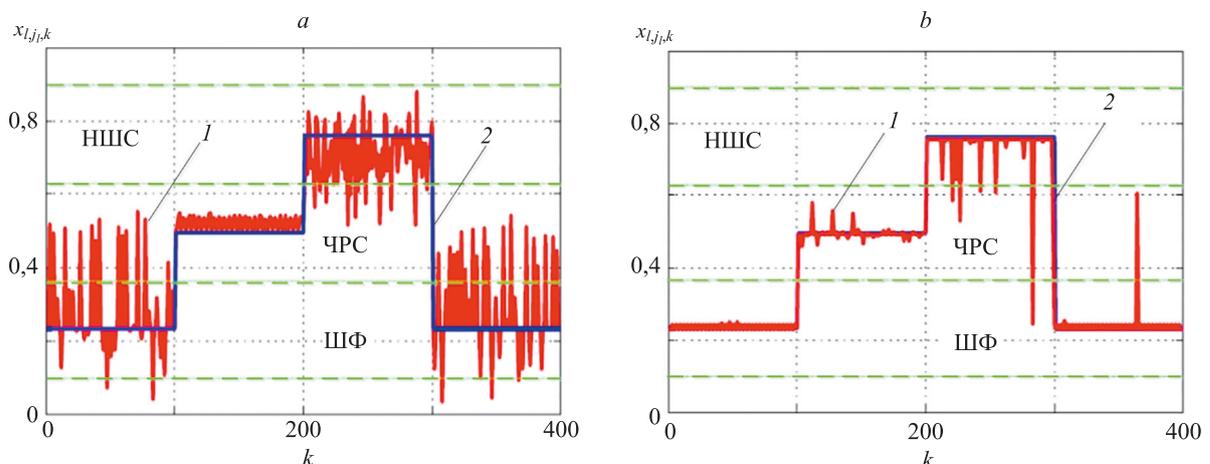


Рис. 4. Результаты формирования обобщенных параметров бортового передатчика при использовании: полиномов Колмогорова–Габора (а) и нейронных сетей (б)

Fig. 4. The result of onboard transmitter generalized parameters formation when using: Kolmogorov-Gabor polynomials (а) and neural networks (б)

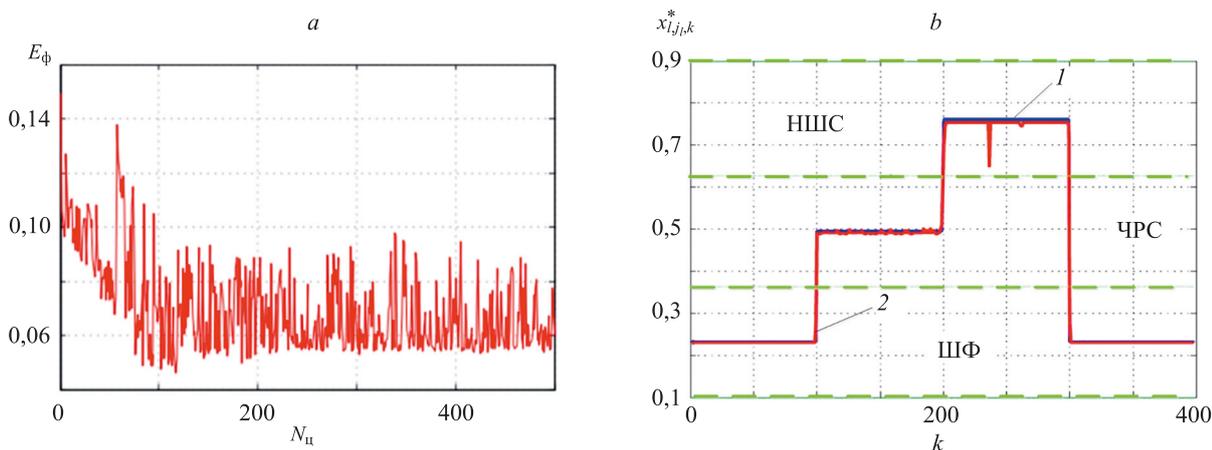


Рис. 5. Результаты обучения нейронной сети (а) и формирования значений обобщенных параметров с использованием нейронной сети и медианного фильтра (b)

Fig. 5. Results: of neural network training (a); of generalized parameters forming values using neural network and median filter (b)

Таблица 2. Результаты формирования обобщенных параметров на тестовой выборке

Table 2. The results of methods study for implementing the FTCPG modules and DLPF modules on the test samples

Модуль ФОПТС	Модуль ЦФНЧ	Среднеквадратическая погрешность, $E_\phi$	Количество аномальных отсчетов
Многослойные нейронные сети	Без фильтрации	0,0418	8
	Скользящее среднее	0,0294	7
	Фильтр с конечной импульсной характеристикой	0,0264	5
	Медианный фильтр	<b>0,0123</b>	<b>1</b>
Полином Колмогорова–Габора	Без фильтрации	0,0939	49
	Скользящее среднее	0,0625	13
	Фильтр с конечной импульсной характеристикой	0,0577	9
	Медианный фильтр	0,0463	9

тов в значениях обобщенных параметров при выбранной форме их представления.

### Заклучение

Задача формирования обобщенных параметров технического состояния сложных технических систем является многопараметрической нелинейной задачей оптимизации и фильтрации, что обуславливает ее решение с использованием аппарата нейронных сетей и медианных фильтров. В результате исследований уменьшена вероятность ошибок определения вида технического состояния сложных технических систем по

значениям обобщенных параметров с 0,0225 до 0,0025 по сравнению с полиномами Колмогорова–Габора. При этом синтез обобщенных параметров целесообразно начинать с разработки модели их формирования и определения допусков на значения телеметрируемых параметров и обобщенных параметров, тем самым обеспечив возможность уточнения и коррекции модели на всех этапах жизненного цикла сложных технических систем. В основе предложенного решения лежат однородные вычисления, которые могут быть эффективно реализованы на перспективных векторно-матричных и тензорных процессорах.

## Литература

1. Поспелов Д.А. Ситуационное управление: теория и практика. М.: URSS, 2021. 288 с.
2. Song Y., Peng Y., Liu D. Model-based health diagnosis for lithium-ion battery pack in space applications // *IEEE Transactions on Industrial Electronics*. 2021. V. 68. N 12. P. 12375–12384. <https://doi.org/10.1109/TIE.2020.3045745>
3. Ritter N., Straub J. Implementation of hardware-based expert systems and comparison of their performance to software-based expert systems // *Machines*. 2021. V. 9. N 12. P. 361. <https://doi.org/10.3390/machines9120361>
4. Лоскутов А.И., Клыков В.А. Идентификация и техническое диагностирование бортовой аппаратуры автономных космических аппаратов на основе биективного преобразования множества диагностических признаков // *Контроль. Диагностика*. 2016. № 4. С. 57–63. <https://doi.org/10.14489/td.2016.04.pp.057-063>
5. Мозгалевский А.В., Гаскаров Д.В. Техническая диагностика (непрерывные объекты): учебное пособие для вузов. М.: Высшая школа, 1975. 207 с.
6. Якимов В.Л., Панкратов А.В. Алгоритм формирования диагностических признаков бортовых динамических систем на основе показателя Херста // *Труды МАИ*. 2015. № 83. С. 23.
7. Якимов В.Л., Мальцев Г.Н. Гибридные сетевые структуры и их использование при диагностировании сложных технических систем // *Информатика и автоматизация*. 2022. Т. 21. № 1. С. 126–160. <https://doi.org/10.15622/ia.2022.21.5>
8. Мальцев Г.Н., Назаров А.В., Якимов В.Л. Исследование процесса диагностирования бортовой аппаратуры автоматических космических аппаратов с использованием дискретно-событийной имитационной модели // *Труды СПИИРАН*. 2018. № 1(56). С. 95–121. <https://doi.org/10.15622/sp.56.5>
9. Duer S., Valicek J., Paš J., Stawowy M., Bernatowicz D., Duer R., Walczak M. Neural networks in the diagnostics process of low-power solar plant devices // *Energies*. 2021. V. 14. N 9. P. 2719. <https://doi.org/10.3390/en14092719>
10. Кузьмин А.Б. Достоверность допусковых методов контроля сложных технических систем // *Автоматика и телемеханика*. 1987. № 10. С. 59–65.
11. Горбань А. Обобщенная аппроксимационная теорема и вычислительные возможности нейронных сетей // *Сибирский журнал вычислительной математики*. 1998. Т. 1. № 1. С. 12–24.
12. Marushko E.E., Doudkin A.A. Ensembles of neural networks for forecasting of time series of spacecraft telemetry // *Optical Memory and Neural Networks*. 2017. V. 26. N 1. P. 47–54. <https://doi.org/10.3103/S1060992X17010064>
13. Fayemi O., Di Q., Zhen Q., Wang Y. Adaptive processing for EM telemetry signal recovery: Field data from sichuan province // *Energies*. 2020. V. 13. N 22. P. 5873. <https://doi.org/10.3390/en13225873>
14. Liu W., Dou Z., Wang W., Liu Y., Zou H., Zhang B., Hou S. Short-term load forecasting based on elastic net improved GMDH and difference degree weighting optimization // *Applied Sciences*. 2018. V. 8. N 9. P. 1603. <https://doi.org/10.3390/app8091603>
15. Jung J., Sael L. Fast and accurate pseudoinverse with sparse matrix reordering and incremental approach // *Machine Learning*. 2020. V. 109. N 12. P. 2333–2347. <https://doi.org/10.1007/s10994-020-05920-5>

## Авторы

**Мальцев Георгий Николаевич** — доктор технических наук, профессор, профессор, Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация, [vka@mil.ru](mailto:vka@mil.ru), [sc 66002715766](https://orcid.org/0000-0002-6755-5700), <https://orcid.org/0000-0002-6755-5700>, [vka@mil.ru](mailto:vka@mil.ru)

**Якимов Виктор Леонидович** — доктор технических наук, доцент, профессор, Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация, [vka@mil.ru](mailto:vka@mil.ru), [sc 57194512292](https://orcid.org/0000-0001-9721-2453), <https://orcid.org/0000-0001-9721-2453>, [vka@mil.ru](mailto:vka@mil.ru)

Статья поступила в редакцию 13.12.2022  
Одобрена после рецензирования 21.05.2023  
Принята к печати 26.07.2023

## References

1. Pospelov D.A. *Situational Management: Theory and Practice*. Moscow, URSS Publ., 2021, 288 p. (in Russian)
2. Song Y., Peng Y., Liu D. Model-based health diagnosis for lithium-ion battery pack in space applications. *IEEE Transactions on Industrial Electronics*, 2021, vol. 68, no. 12, pp. 12375–12384. <https://doi.org/10.1109/TIE.2020.3045745>
3. Ritter N., Straub J. Implementation of hardware-based expert systems and comparison of their performance to software-based expert systems. *Machines*, 2021, vol. 9, no. 12, pp. 361. <https://doi.org/10.3390/machines9120361>
4. Loskutov A.I., Klykov V.A. Identification and technical diagnostics of onboard equipment of autonomous spacecrafts based on bijective transformations of the set of diagnostic features. *Testing. Diagnostics*, 2016, no. 4, pp. 57–63. (in Russian). <https://doi.org/10.14489/td.2016.04.pp.057-063>
5. Mozgalevskii A.V., Gaskarov D.V. *Technical Diagnostics (Continuous Objects)*. Moscow, Vysshaja shkola Publ., 1975, 207 p. (in Russian)
6. Yakimov V. L., Pankratov A. V. Algorithm of diagnostic signs formation for onboard dynamic systems based on Hurst index. *Trudy MAI*, 2015, no. 83, pp. 23. (in Russian)
7. Yakimov V., Maltsev G. Hybrid network structures and their use in diagnosing complex technical systems. *Informatics and Automation*, 2022, vol. 21, no. 1, pp. 126–160. (in Russian). <https://doi.org/10.15622/ia.2022.21.5>
8. Maltsev G.N., Nazarov A.V., Yakimov V.L. The study of the diagnosing of automatic spacecraft on-board equipment using a discrete-event simulation model. *SPIIRAS Proceedings*, 2018, no. 1(56), pp. 95–121. (in Russian). <https://doi.org/10.15622/sp.56.5>
9. Duer S., Valicek J., Paš J., Stawowy M., Bernatowicz D., Duer R., Walczak M. Neural networks in the diagnostics process of low-power solar plant devices. *Energies*, 2021, vol. 14, no. 9, pp. 2719. <https://doi.org/10.3390/en14092719>
10. Kuz'min A.B. Credibility of tolerance complex industrial system monitoring methods. *Avtomatika i Telemekhanika*, 1987, no. 10, pp. 59–65. (in Russian)
11. Gorban A. Generalized approximation theorem and computational capabilities of neural networks. *Sibirskij zhurnal vychislitel'noj matematiki*, 1998, vol. 1, no. 1, pp. 12–24. (in Russian)
12. Marushko E.E., Doudkin A.A. Ensembles of neural networks for forecasting of time series of spacecraft telemetry. *Optical Memory and Neural Networks*, 2017, vol. 26, no. 1, pp. 47–54. <https://doi.org/10.3103/S1060992X17010064>
13. Fayemi O., Di Q., Zhen Q., Wang Y. Adaptive processing for EM telemetry signal recovery: Field data from sichuan province. *Energies*, 2020, vol. 13, no. 22, pp. 5873. <https://doi.org/10.3390/en13225873>
14. Liu W., Dou Z., Wang W., Liu Y., Zou H., Zhang B., Hou S. Short-term load forecasting based on elastic net improved GMDH and difference degree weighting optimization. *Applied Sciences*, 2018, vol. 8, no. 9, pp. 1603. <https://doi.org/10.3390/app8091603>
15. Jung J., Sael L. Fast and accurate pseudoinverse with sparse matrix reordering and incremental approach. *Machine Learning*, 2020, vol. 109, no. 12, pp. 2333–2347. <https://doi.org/10.1007/s10994-020-05920-5>

## Authors

**Georgy N. Maltsev** — D.Sc., Full Professor, Mozhasky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation, [vka@mil.ru](mailto:vka@mil.ru), [sc 66002715766](https://orcid.org/0000-0002-6755-5700), <https://orcid.org/0000-0002-6755-5700>, [vka@mil.ru](mailto:vka@mil.ru)

**Victor L. Yakimov** — D.Sc., Associate Professor, Professor, Mozhasky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation, [vka@mil.ru](mailto:vka@mil.ru), [sc 57194512292](https://orcid.org/0000-0001-9721-2453), <https://orcid.org/0000-0001-9721-2453>, [vka@mil.ru](mailto:vka@mil.ru)

Received 13.12.2022  
Approved after reviewing 21.05.2023  
Accepted 26.07.2023



doi: 10.17586/2226-1494-2023-23-4-836-842

УДК 621.45.053.4

## Численное моделирование газодинамики при работе широкодиапазонного ракетного сопла с пористой вставкой

Михаил Владимирович Михайлов<sup>1</sup>✉, Николай Васильевич Продан<sup>2</sup>,  
Максим Евгеньевич Ренев<sup>3</sup>

<sup>1,2,3</sup> Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация

<sup>1</sup> [mixailovmish@yandex.ru](mailto:mixailovmish@yandex.ru)✉, <https://orcid.org/0009-0003-3840-203X>

<sup>2</sup> [kolinti@mail.ru](mailto:kolinti@mail.ru), <https://orcid.org/0000-0002-6383-039X>

<sup>3</sup> [renevme@mail.ru](mailto:renevme@mail.ru), <https://orcid.org/0000-0002-0583-0607>

### Аннотация

**Введение.** При движении ракеты в плотных слоях атмосферы Земли классические сопла работают в режиме перерасширения реактивной струи. В этом режиме происходит частичное снижение величины удельного импульса. Вследствие этого увеличивается количество расходуемого ракетным двигателем топлива. Повышение эффективности работы сопел можно достигнуть использованием конструкций широкодиапазонных сопел. В этом случае замена сплошной стенки сопла на перфорированную позволяет компенсировать потери удельного импульса. В работе представлено исследование влияния пористой вставки на режимы работы сопла. **Метод.** Выполнено численное моделирование в пакете программ Ansys Fluent. На первом этапе исследования создана геометрическая модель расчетной зоны: двумерная модель сопла ракетного двигателя РД-107 и окружающая сопло область, в которой задана внешняя среда (воздушная атмосфера). Произведен расчет истечения продуктов сгорания через построенное сопло при различном давлении атмосферного воздуха. Классическое сопло заменено на сопло с пористой вставкой и проведен расчет при аналогичных значениях атмосферного давления. Выполнено сравнение значений удельного импульса, полученных при расчетах с классическим и пористым соплом. Определено количество топлива, сэкономленного при замене классического сопла на пористое с помощью определения разницы площадей, ограниченных кривыми на графике зависимости удельного импульса от рассматриваемой высоты над уровнем моря. **Основные результаты.** Результаты сравнения величин удельного импульса сопел с непроницаемой стенкой и пористой вставкой позволили сделать вывод, что до высоты 5,4 км удельный импульс сопла с перфорированной стенкой превышает значения удельного импульса классического сопла. Оценка эффективности применения газопроницаемой вставки в конструкции сопла при его работе в плотных слоях атмосферы Земли показала, что с началом работы на высоте 0 км над уровнем моря и до высоты, на которой сопло работает в расчетном режиме, величина компенсируемого удельного импульса составляет 2,2%. **Обсуждение.** Результаты исследования могут быть применены при конструировании сопловых устройств современных ракетных двигателей, работающих в плотных слоях атмосферы.

### Ключевые слова

численное моделирование, ракетный двигатель, широкодиапазонные сопла, пористая вставка, режим перерасширения

### Благодарности

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации в ходе реализации проекта «Создание опережающего научно-технического задела в области разработки передовых технологий малых газотурбинных, ракетных и комбинированных двигателей сверхлегких ракет-носителей, малых космических аппаратов и беспилотных воздушных судов, обеспечивающих приоритетные позиции российских компаний на формируемых глобальных рынках будущего», № FZWF-2020-0015.

**Ссылка для цитирования:** Михайлов М.В., Продан Н.В., Ренев М.Е. Численное моделирование газодинамики при работе широкодиапазонного ракетного сопла с пористой вставкой // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 836–842. doi: 10.17586/2226-1494-2023-23-4-836-842

© Михайлов М.В., Продан Н.В., Ренев М.Е., 2023

## Numerical simulation of gas dynamics during operation of a wide-range rocket nozzle with a porous insert

Mikhail V. Mikhailov<sup>1</sup>✉, Nikolai V. Prodan<sup>2</sup>, Maxim E. Renev<sup>3</sup>

<sup>1,2,3</sup> Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation

<sup>1</sup> mixailovmish@yandex.ru✉, <https://orcid.org/0009-0003-3840-203X>

<sup>2</sup> kolinti@mail.ru, <https://orcid.org/0000-0002-6383-039X>

<sup>3</sup> renevme@mail.ru, <https://orcid.org/0000-0002-0583-0607>

### Abstract

When the rocket moves in the dense layers of the Earth’s atmosphere, classical nozzles operate in the jet overexpansion mode. In this mode, there is a partial decrease in the magnitude of the specific impulse. As a result, the amount of fuel consumed by the rocket engine increases. An increase in the efficiency of nozzle operation can be achieved by using designs of wide-range nozzles, in which case the replacement of a solid nozzle wall with a perforated one makes it possible to compensate for the loss of specific impulse. The paper presents a study of the effect of a porous insert on the operating modes of the nozzle. Numerical simulation was performed in the Ansys Fluent software package. At the first stage of the study, a geometric model of the computational zone is created which includes a two-dimensional model of the RD-107 rocket engine nozzle and a computational domain that simulates the external environment (air atmosphere). The calculation of the outflow of combustion products through the constructed nozzle at different pressures of atmospheric air is carried out. In the future, the classical nozzle is replaced by a nozzle with a porous insert, and the calculation is carried out at the same values of atmospheric pressure. The values of the specific impulse obtained in calculations with a classical and porous nozzle are compared. The amount of fuel saved when replacing a classic nozzle with a porous one is determined by the difference in the areas bounded by the curves on the plot of specific impulse versus the considered height above the Earth’s surface. Comparison of the values of the specific impulse of nozzles with an impenetrable wall and a porous insert made it possible to conclude that up to a height of 5.4 km the specific impulse of the nozzle with a perforated wall exceeds the values of the specific impulse of the classical nozzle. Evaluation of the effectiveness of the use of a gas-permeable insert in the nozzle design when the nozzle operates in dense layers of the Earth’s atmosphere showed that with the start of operation at a height of 0 km above sea level and up to the height at which the nozzle operates in the design mode – the value of the compensated specific impulse is 2.2 %. The results of the study can be applied in the design of nozzle devices of modern rocket engines operating in dense layers of the atmosphere.

### Keywords

numerical simulation, rocket engine, wide-range nozzles, porous insert, overexpansion mode

### Acknowledgements

The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation during the implementation of the project “Creation of advanced scientific and technical groundwork in the development of advanced technologies of small gas turbine, rocket and combined engines of ultralight launch vehicles, small spacecraft and unmanned aircraft, ensuring the priority positions of Russian companies in the emerging global markets of the future”, no. FZWF-2020-0015.

**For citation:** Mikhailov M.V., Prodan N.V., Renev M.E. Numerical simulation of gas dynamics during operation of a wide-range rocket nozzle with a porous insert. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 836–842 (in Russian). doi: 10.17586/2226-1494-2023-23-4-836-842

### Введение

В настоящее время в области конструирования ракетных двигателей актуальным является вопрос создания широкодиапазонных и авторегулируемых сопловых устройств. Часто сопла ракетных двигателей в плотных слоях атмосферы работают в режиме перерасширения, далее наступает непродолжительный этап полета, когда сопло работает в расчетном режиме. Показатель нерасчетности работы сопла для определенной высоты полета задается на этапе проектирования, так как зависит от давления рабочего тела в выходном сечении сопла и от давления окружающей среды. При работе сопла в режиме перерасширения сопловые газы в выходном сечении расширяются до давления, меньшего чем атмосферное. Из-за этого от сечения, в котором газ расширился до давления в атмосфере, и до выходного сечения — происходит потеря тяги сопла. Возможность авторегулирования давления на срезе сопла, во время полета ракеты в плотных слоях атмосферы, позволила

бы до определенной высоты поддерживать расчетный режим работы.

Создание регулируемого соплового устройства возможно несколькими методами, одним из них является применение в конструкции сверхзвуковой части сопла — отверстий (щелей) [1]. В таком случае связь рабочего тела с атмосферным воздухом в сопле будет реализована через отверстия в стенках сопла. Сквозь щели внешнее давление воздействует на сопловые газы, ограничивая перерасширение потока. При прохождении первого участка траектории первая щель закрывается и поток расширяется до следующего сечения, в котором давление потока совпадает с внешним давлением. Щелевые сопла имеют высотные характеристики, близко прилегающие к идеально регулируемому соплу [2].

В работах [3, 4] проведена оценка эффективности работы двигателя первой ступени ракеты-носителя с двумя различными вариантами степени расширения сопла. Давление на срезе варьировалось в диапазоне 10–5 кПа. Степень расширения при этом изменялась

посредством установки в сверхзвуковой части двух кольцевых щелей. В результате численного расчета и проведенных огневых испытаний на участке с давлением сопловых газов меньше атмосферного организованы щели в контуре сопла. Полученный результат оказывает существенное влияние на прирост тяги двигателя и способствует организации направленного раннего отрыва потока внутри сопла. По сравнению с гладким коническим соплом прирост тяги достиг 5–8 %.

Наряду со щелевыми отверстиями активно исследовано применение газопроницаемых пористых материалов в конструкции сопловых устройств и различных аэродинамических профилей. В работах [5–7] приведены результаты исследования обтекания цилиндра с газопроницаемой пористой вставкой, установленной на торце цилиндра. При малых скоростях течения коэффициент инерционных потерь уменьшился на порядок, но при этом, значительно повысился коэффициент сопротивления трения на переднем участке пористой вставки. Данное наблюдение иллюстрирует одну из возможных проблем, которая может возникнуть при установке пористой вставки в конструкции сопел ракетных двигателей, так как при работе сопла — температура на стенках может достигать 2000–3000 К. В качестве материалов для изготовления пористой структуры возможно использование жаропрочных сплавов, либо композитных материалов из углерода<sup>1</sup>. В случае применения углеродистых материалов необходимо обеспечивать антиокислительную защиту. В работе [8] описан результат проведения огневых испытаний обтекания топливно-окислительной смесью образцов углерод-углеродного боросилицированного материала при температуре 2400 К. После проведения испытания образец материала не подвергся эрозии.

В работе [9] вставка из пористо-ячеистого материала установлена на внешнюю поверхность центрального тела клино-воздушного ракетного двигателя. В результате численного моделирования, при условиях, соответствующих земной атмосфере на уровне моря, выявлено, что из-за появления газопроницаемой структуры, процесс отрыва потока со стенки центрального тела изменился значительно. При этом перераспределение давления по поверхности центрального тела привело к увеличению тяги двигателя на 8 % по сравнению тем же соплом, но без пористой вставки.

В настоящей работе исследована возможность авторегулируемости значения удельного импульса сопла, работающего в плотных слоях атмосферы. Приведена оценка компенсации потерь удельного импульса классического сопла двигателя РД-107 при замене сплошной стенки сопла на перфорированную. Представлены результаты численного моделирования истечения реактивной струи для классического и пористого сопел, позволяющие оценить эффективность использования газопроницаемых материалов при работе ракетного сопла.

<sup>1</sup> Углерод-углеродные композиционные материалы. Описание свойств и возможностей применения в авиационной технике [Электронный ресурс]. URL: [https://bstudy.net/810004/tehnika/uglerod\\_uglerodnye\\_kompozitsionnye\\_materialy?ysclid=lfibz5ps4n321303151](https://bstudy.net/810004/tehnika/uglerod_uglerodnye_kompozitsionnye_materialy?ysclid=lfibz5ps4n321303151) (дата обращения: 13.03.2023).

## Постановка задачи

Жидкостный ракетный двигатель РД-107<sup>2</sup> используется на первых ступенях ракет-носителей «Союз-ФГ» и «Союз-2». Выбор данного двигателя для проведения численного моделирования связан с тем, что для него известны параметры работы и геометрия соплового устройства. Геометрическая модель расчетной области представлена на рис. 1.

Границы расчетной области определены для 5 калибров по высоте и 10 — по длине. За значение калибра выбран радиус выходного сечения сопла, чтобы снизить влияние граничных условий на процесс истечения газа из сопла. Задача моделирования двумерная и осесимметричная.

На первом этапе выполнено численное моделирование с применением классического сопла и определены точки с давлением, соответствующим внешней среде. Численное моделирование проведено для 5 высот с шагом 1,5 км и началом работы на высоте 0 км. Крайней является высота, на которой сопло работает в расчетном режиме. После определения искомых точек классическое сопло было заменено на пористое. При этом пористая вставка закрывалась сплошной стенкой по мере увеличения рассматриваемой высоты.

Учет спутного потока воздуха, натекающего на внешнюю стенку сопла и имитирующего движение ракеты, не проводился. Отметим, что не было исследовано взаимодействие сопел при многосопловой компоновке. Причиной отказа от включения в решение задачи спутного потока является то, что выполненная серия расчетов позволяет исследовать взаимодействие потока с перфорированной стенкой независимо от того, с какой скоростью может двигаться ракета-носитель в атмосфере и на какой конкретно ракете оно установлено.

## Расчетная модель

В качестве уравнений движения в математической модели использованы уравнения Навье-Стокса —

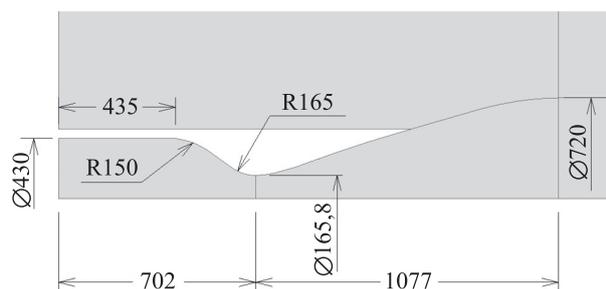


Рис. 1. Геометрическая модель сопла РД-107.

Размеры указаны в мм

Fig. 1. Geometric model of the RD-107 nozzle (dimensions are in mm)

<sup>2</sup> Схема, устройство и некоторые параметры ракетного двигателя РД-107 [Электронный ресурс]. URL: <http://www.lpre.de/energomash/RD-107/index.htm> (дата обращения: 16.03.2023).

осредненные по Рейнольдсу и замкнутые моделью турбулентности Shear Stress Transport (SST)  $k-\omega$ , которая применима для расчетов сверхзвуковых струйных течений [10]. Данная модель является более точной и надежной по сравнению с обычной  $k-\omega$ . Решатель выбран на основе давления [11].

Параметры, задаваемые на входе в сопло, соответствовали параметрам в реальном двигателе: полному давлению — 5850 кПа и температуре — 3520 К. В качестве рабочего тела использованы продукты сгорания после реакции сгорания керосина Т-1 и жидкого кислорода. Внешняя среда — воздушная атмосфера. Газовый поток реализован с помощью подключения модели газовой смеси (Species Transport). Через входное сечение в сопло подавалась газовая смесь, состоящая из  $H_2O$  и  $CO_2$ , при этом внешняя среда представляет собой воздух, состоящий преимущественно из  $N_2$  и  $O_2$ . Количество компонента задано массовой долей в смеси.

Свойства смеси определялись составом с использованием закона смешивания. Теплопроводность, диффузия и вязкость изменялись по закону смешивания идеальных газов. Согласно данной модели — происходит усреднение свойств среды отдельно для каждого компонента с учетом их массовых долей. Плотность компонентов смеси изменялась по закону совершенного газа.

**Модель пористого тела.** Для исследования несжимаемого течения в пористой среде использовано уравнение Дарси [12]:

$$\nabla p = K_1 \mu v_{vol},$$

где  $\nabla p$  — градиент давления, под действием которого газ движется внутри поры;  $\mu$  — динамическая вязкость газа;  $K_1$  — коэффициент вязкостного сопротивления пористой среды. Приведенная скорость  $v_{vol}$  (рассчитана по объемному расходу газа сквозь пористое тело) связана со скоростью в порах  $v$  соотношением:

$$v_{vol} = \gamma_{пор} v,$$

где  $\gamma_{пор}$  — коэффициент пористости, определяемый как отношение объема, занятого газом к общему объему:

$$\gamma_{пор} = \Omega_f / (\Omega_f + \Omega_s),$$

где  $\Omega_f + \Omega_s$  — объем, занятый пористым телом;  $\Omega_f$  и  $\Omega_s$  — объемы, занятые газом и материалом пор в пористом теле. В уравнении Дарси перепад давления вдоль пористого канала линейно зависит от скорости потока.

Пористая среда в программе Ansys Fluent реализована включением функции «Porous zone». В этом случае участки стенки с отверстиями заданы проницаемыми областями, через которые газ с большим давлением (воздушная среда) проникает в область с меньшим давлением (продукты сгорания). Для определения параметров пористой вставки в Ansys Fluent, введен коэффициент вязкостного сопротивления газовой среды. Значение коэффициента определено по результатам серии численных расчетов в Ansys Fluent, при этом коэффициент варьировался. Выбранное значение ко-

эффициента  $1 \cdot 10^7 \text{ м}^{-2}$  получено из предположения, что газ из внешней среды, проходя через перфорированную стенку — на выходе из нее будет иметь скорость в два раза меньшую, чем при входе. При этом поры направляются параллельно оси сопла. Вдоль оси поперечного направления к оси сопла — значение коэффициента принято равным  $1 \cdot 10^{11} \text{ м}^{-2}$ .

## Результаты численного моделирования

**Расчет классического сопла.** Для определения точек установки пористой вставки выполнен расчет классического сопла при изменяющемся атмосферном давлении. Сопло двигателя РД-107 в выходном сечении имеет значение давления — 0,039 МПа, при старте ракеты и некоторое время после — сопло работает в режиме перерасширения потока. Исходя из этого, при работе сопла на уровне моря определено сечение сопла, в котором давление потока совпадает с внешним давлением и с точкой установки первой пористой вставки. Остальные участки установки пористых вставок определены аналогичным образом для рассматриваемых высот.

По результатам численных расчетов установлены сечения сопла, в которых возможно организовать ранний отрыв потока от стенки. Рассмотренные высоты, значения атмосферного давления<sup>1</sup>, а также значения тяги сопла для соответствующих высот в сопле с непроницаемой стенкой приведены в табл. 1.

Заметим, что на уровне моря пористая вставка полностью открыта. По мере увеличения высоты полета до 5,4 км — вставка закрывается и к моменту выхода сопла на расчетный режим полностью закрыта.

По результатам численного моделирования с истечением продуктов сгорания через классическое сопло с непроницаемой стенкой — до высоты 5,4 км сопло работает в режиме перерасширения. На высоте 5,4 км по всей длине стенки сопла в сверхзвуковой части, вплоть до выходного сечения — давление газа больше или соответствует давлению окружающей среды. Распределение статического давления ( $p$ ) для двух случаев представлено на рис. 2.

Область, в которой установлена пористая вставка представляет собой объем, в который поступает воздух из внешней среды. Параметры воздуха при прохождении через объем изменяются, и воздух начинает взаимодействовать с потоком рабочего тела в сопле. По мере увеличения высоты, части проницаемых пористых вставок 1–4 закроются стенкой вплоть до высоты, на которой сопло работает в расчетном режиме. Области установок пористых вставок показаны на рис. 3.

**Расчет сопла с пористой вставкой.** При численном моделировании течения с установленной пористой вставкой — на высоте 0 км вставка полностью открыта. По мере увеличения высоты полета с шагом в 1,5 км пористая вставка постепенно закрывается и полностью непроницаема на высоте 5,4 км. Распределение скорости ( $u$ ) в сопле представлено на рис. 4.

<sup>1</sup> ГОСТ 4401–81 Атмосфера стандартная. Параметры атмосферы. Введ. 01.07.83. М.: Изд-во стандартов, 1981. 181 с.

Таблица 1. Параметры в сопле с непроницаемой стенкой  
Table 1. Parameters in a nozzle with an impermeable wall

Высота, км	Давление внешней среды, кПа	Расстояние от критического сечения до рассматриваемого сечения, мм	Тяга, кН	Удельный импульс, м/с
0	101	552,5	185	2497
1,5	85	606,3	192	2590
3,0	70	654,8	198	2669
4,5	58	733,4	203	2737
5,4	51	1077	207	2786

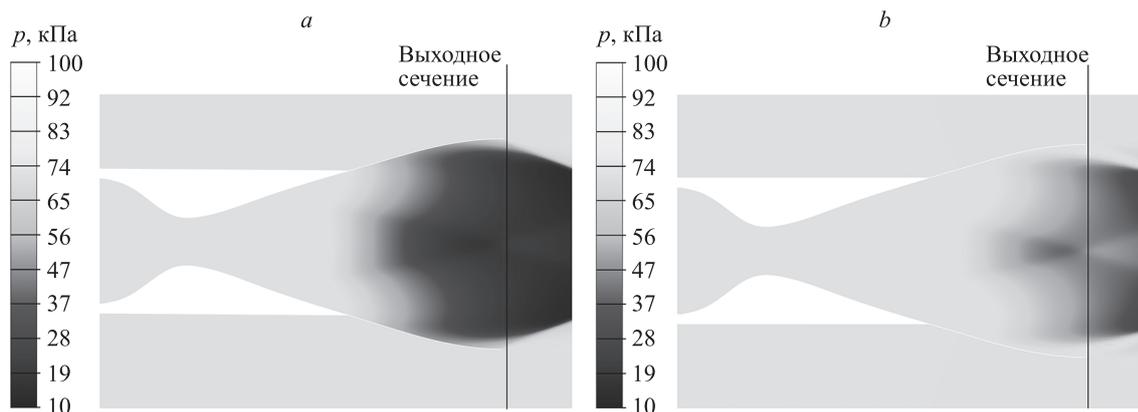


Рис. 2. Распределение статического давления в сопле с непроницаемой стенкой: 0 км (а) и 5,4 км (б)  
Fig. 2. Distribution of static pressure in a nozzle with an impermeable wall: 0 km (a) and 5,4 km (b)

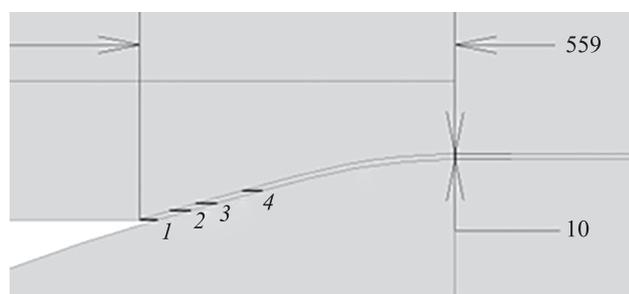


Рис. 3. Области установок пористых вставок, в зависимости от рассматриваемой высоты (размеры в мм): 1 — 0–1,5 км; 2 — 1,5–3 км; 3 — 3–4,5 км; 4 — 4,5–5,4 км

Fig. 3. Installation areas of the porous insert vs. the considered height (dimensions in mm): 1 — 0–1,5 km; 2 — 1,5–3 km; 3 — 3–4,5 km; 4 — 4,5–5,4 km

На высоте 0 км пористая вставка максимально компенсировала величину потерь удельного импульса сопла. По полученным результатам (рис. 4), видно, что сквозь вставку просачивается газ из внешней среды и воздействует на рабочее тело в сопле (рис. 4, а). Воздух способствует отрыву потока от стенки, тем самым ограничивая перерасширение потока. Когда вставка закрыта сплошной стенкой (рис. 4, б) — взаимодействие потока сопловых газов и внешней среды происходит за выходным сечением сопла.

Расход рабочего тела через двигатель без пористой вставки при истечении продуктов сгорания — 74 кг/с. В табл. 2 приведены значения основных параметров сопла, полученные в ходе численного моделирования. Изменение массового расхода через выходное сечение сопла зависит от количества проходящего через вставку атмосферного воздуха. На высоте 5,4 км газопроница-

Таблица 2. Параметры сопла с пористой вставкой  
Table 2. Nozzle parameters with porous insert

Высота, км	Давление внешней среды, кПа	Массовый расход газа через выходное сечение с использованием пористой вставки, кг/с	Удельный импульс сопла, м/с		Отношение удельного импульса сопла со вставкой и без вставки, %
			без вставки	со вставкой	
0	101	118	2497	2603	4,2
1,5	85	109	2590	2674	3,2
3,0	70	104	2669	2715	1,7
4,5	58	92	2737	2753	0,6
5,4	51	74	2786	2784	0

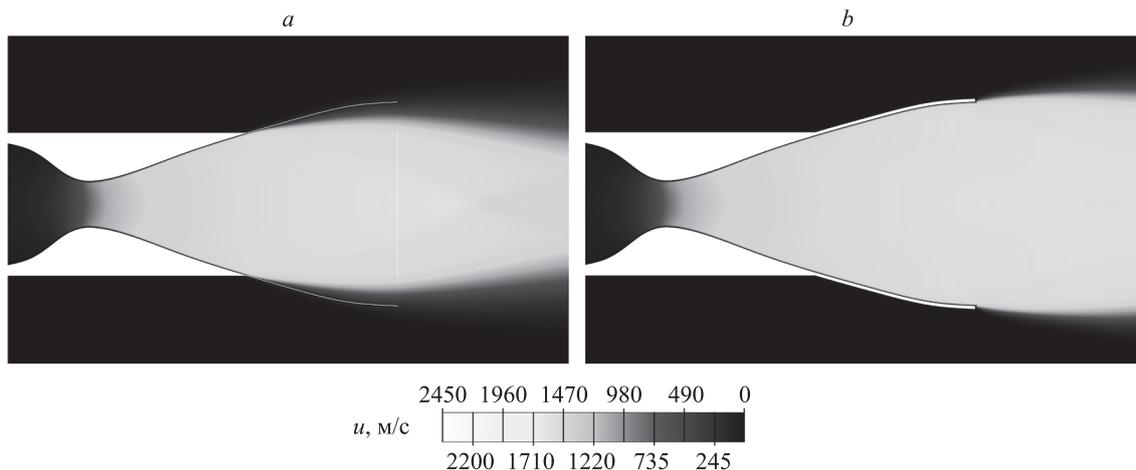


Рис. 4. Распределение скорости в сверхзвуковой части сопла с установленной пористой, газопроницаемой вставкой: 0 км (а) и 5,4 км (б)

Fig. 4. Velocity distribution in the supersonic part of the nozzle with a porous, gas-permeable insert: 0 km (a) and 5,4 km (b)

емая вставка полностью закрыта сплошной стенкой и расход в выходном сечении сопла зависит только от количества газа, подаваемого в камеру сгорания двигателя.

Исходя из результатов (табл. 2), заметно, что в сопле с пористой вставкой, при реализации режима перерасширения сопловых газов — компенсация потерь удельного импульса происходит в диапазоне 4,2–0 %. На рис. 5 приведен график зависимости удельного импульса ( $I$ ) от высоты полета ( $H$ ).

Расчетный режим работы сопла реализован на высоте 5,4 км. На больших высотах использование пористой вставки не позволило компенсировать потери удельного

импульса, так как точка, за которой давление потока меньше давления окружающей среды, находится за выходным сечением сопла и сопло работает в расчетном режиме. Разница площадей под кривыми на рис. 5 составила 2,2 %, что соответствует величине компенсируемого удельного импульса на всем рассматриваемом участке полета.

### Заключение

Предложен метод, оптимальный для анализа эффективности использования пористой вставки в сопловом устройстве ракетного двигателя. Одним из результатов работы является выработка последовательных этапов численного моделирования.

Для определения области на стенке сопла, подходящей для установки газопроницаемой пористой вставки — на заданной высоте достаточно провести расчет истечения газа через сопло, работающее в режиме перерасширения. По результатам расчета получены значения статического давления сопловых газов на стенку сопла. Начиная с точки, где давление газов равно значению давления внешней среды, найден участок, на котором происходит снижение величины удельного импульса. Полученная координата — оптимальная точка для установки пористой вставки.

При работе сопла ракетного двигателя в режиме перерасширения возможно проведение численного моделирования с использованием пористой вставки, которая выравнивает окружающее давление и давление на участке внутренней стенки сопла с перерасширением газов. Такое сопло является авторегулируемым и позволяет снизить потери удельного импульса при условии, что пористая вставка закрывается непроницаемой стенкой в соответствии с изменяющимся атмосферным давлением.

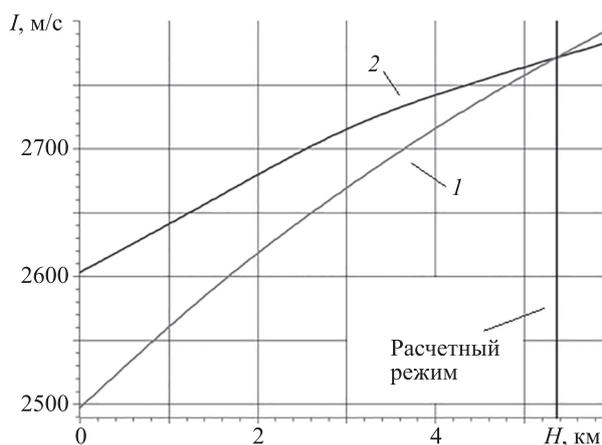


Рис. 5. Графики зависимостей удельных импульсов от высоты полета для непроницаемого сопла (кривая 1) и сопла с перфорированной стенкой (кривая 2)

Fig. 5. Graph of specific impulse versus flight altitude: 1 — tight nozzle specific impulse; 2 — perforated wall nozzle specific impulse

## Литература

1. Семенов В.В., Иванов И.Э., Крюков И.А., Федоров В.В. Повышение эффективности ракетного двигателя второй ступени // Вестник Пермского национального исследовательского политехнического университета. Аэрокосмическая техника. 2015. № 43. С. 5–17. <https://doi.org/10.15593/2224-9982/2015.43.01>
2. Васильев А.П., Кудрявцев В.М., Кузнецов В.А. и др. Основы теории и расчета жидкостных ракетных двигателей. В 2 кн. Кн. 1: учебник для авиац. спец. вузов / 4-е изд., перераб. и доп. М.: Высшая школа, 1993. 382 с.
3. Семенов В.В., Иванов И.Э., Крюков И.А. Регулирование высоты сопла с большой степенью расширения // Вестник Пермского национального исследовательского политехнического университета. Аэрокосмическая техника. 2015. № 40. С. 5–21. <https://doi.org/10.15593/2224-9982/2015.40.01>
4. Семенов В.В., Иванов И.Э., Крюков И.А. Двухконтурное щелевое сопло ракетного двигателя // Вестник Пермского национального исследовательского политехнического университета. Аэрокосмическая техника. 2016. № 46. С. 56–71. <https://doi.org/10.15593/2224-9982/2016.46.03>
5. Кириловский С.В., Маслов А.А., Миронов С.Г., Поплавская Т.В. Применение скелетной модели высокопористого ячеистого материала для моделирования сверхзвукового обтекания цилиндра с передней газопроницаемой вставкой // Известия РАН. Механика жидкости и газа. 2018. № 3. С. 78–86. <https://doi.org/10.7868/S0568528118030088>
6. Maslov A.A., Mironov S.G., Poplavskaya T.V., Kirilovskiy S.V. Supersonic flow around a cylinder with a permeable high-porosity insert: experiment and numerical simulation // Journal of Fluid Mechanics. 2019. V. 867. P. 611–632. <https://doi.org/10.1017/jfm.2019.165>
7. Mironov S.G., Kirilovskiy S.V., Militsina T.S., Poplavskaya T.V., Tsyryulnikov I.S., Valiullin I.R. Effect of the structure of skeleton models of cellular materials on the drag of a cylinder with a frontal gas-permeable insert in a supersonic flow // Journal of Physics: Conference Series. 2019. V. 1404. N 1. P. 012032. <https://doi.org/10.1088/1742-6596/1404/1/012032>
8. Гуняев Г.М., Гофин М.Я. Углерод-углеродные композиционные материалы // Авиационные материалы и технологии. 2013. № 1S. С. 62–90.
9. Каун Ю.В., Чернышов М.В. Влияние газопроницаемых конструкционных материалов на тягу сопла внешнего расширения широкодиапазонного ракетного двигателя // Известия Российской академии ракетных и артиллерийских наук. 2022. № 4(124). С. 117–123. [https://doi.org/10.53816/20753608\\_2022\\_4\\_117](https://doi.org/10.53816/20753608_2022_4_117)
10. Лойцианский Л.Г. Механика жидкости и газа. М.: Дрофа, 2003. 840 с.
11. 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>
12. Запрягаев В.И., Кавун И.Н., Солотчин А.В. Структура течения при взаимодействии сверхзвуковой струи с пористой преградой // Прикладная механика и техническая физика. 2015. Т. 56. № 3. С. 73–81. <https://doi.org/10.15372/PMTF20150309>

## Авторы

**Михайлов Михаил Владимирович** — младший научный сотрудник, Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация, <https://orcid.org/0009-0003-3840-203X>, [mikhailovmish@yandex.ru](mailto:mikhailovmish@yandex.ru)

**Продан Николай Васильевич** — кандидат технических наук, старший научный сотрудник, Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация, [sc 56177927100](https://orcid.org/0000-0002-6383-039X), <https://orcid.org/0000-0002-6383-039X>, [kolinti@mail.ru](mailto:kolinti@mail.ru)

**Ренев Максим Евгеньевич** — младший научный сотрудник, Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация, [sc 57211271545](https://orcid.org/0000-0002-0583-0607), <https://orcid.org/0000-0002-0583-0607>, [renewme@mail.ru](mailto:renewme@mail.ru)

## References

1. Semenov V.V., Ivanov I.E., Kryukov I.A., Fedorov V.V. Improvement in the efficiency of a second stage rocket engine. *PNRPU Aerospace Engineering Bulletin*, 2015, no. 43, pp. 5–17. (in Russian). <https://doi.org/10.15593/2224-9982/2015.43.01>
2. Vasilev A.P., Kudriavtsev V.M., Kuznetsov V.A. et al. *Fundamentals of the Theory and Calculation Of Liquid Rocket Engines. Vol. 2*. Moscow, Vysshaja Shkola Publ., 1993, 382 p. (in Russian)
3. Semenov V.V., Ivanov I.E., Kryukov I.A. Control of nozzle design altitude with high expansion ratio. *PNRPU Aerospace Engineering Bulletin*, 2015, no. 40, pp. 5–21. (in Russian). <https://doi.org/10.15593/2224-9982/2015.40.01>
4. Semenov V.V., Ivanov I.E., Kryukov I.A. Dual bell slot nozzle of a rocket engine. *PNRPU Aerospace Engineering Bulletin*, 2016, no. 46, pp. 56–71. (in Russian). <https://doi.org/10.15593/2224-9982/2016.46.03>
5. Kirilovskiy S.V., Maslov A.A., Mironov S.G., Poplavskaya T.V. Application of the skeleton model of a highly porous cellular material in modeling supersonic flow past a cylinder with a forward gas-permeable insert. *Fluid Dynamics*, 2018, vol. 53, no. 3, pp. 409–416. <https://doi.org/10.1134/s0015462818030102>
6. Maslov A.A., Mironov S.G., Poplavskaya T.V., Kirilovskiy S.V. Supersonic flow around a cylinder with a permeable high-porosity insert: experiment and numerical simulation. *Journal of Fluid Mechanics*, 2019, vol. 867, pp. 611–632. <https://doi.org/10.1017/jfm.2019.165>
7. Mironov S.G., Kirilovskiy S.V., Militsina T.S., Poplavskaya T.V., Tsyryulnikov I.S., Valiullin I.R. Effect of the structure of skeleton models of cellular materials on the drag of a cylinder with a frontal gas-permeable insert in a supersonic flow. *Journal of Physics: Conference Series*, 2019, vol. 1404, no. 1, pp. 012032. <https://doi.org/10.1088/1742-6596/1404/1/012032>
8. Gunyaev G.M., Gofin M.Ya. Carbon-carbon composite materials. *Aviation Materials and Technologies*, 2013, no. 1S, pp. 62–90. (in Russian)
9. Kaun Yu.V., Chernyshov M.V. Influence of gas permeable structural materials on the thrust of the external expansion nozzle of wide-range rocket engine. *Izvestiya Rossijskoj akademii raketnyh i artillerijskih nauk*, 2022, no. 4(124), pp. 117–123. (in Russian). [https://doi.org/10.53816/20753608\\_2022\\_4\\_117](https://doi.org/10.53816/20753608_2022_4_117)
10. Loitcianskii L.G. *Mechanics of Liquid and Gas*. Moscow, Drofa Publ., 2003, 840 p. (in Russian)
11. 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>
12. Zapryagaev V.I., Kavun I.N., Solotchin A.V. Flow structure formed due to interaction of a supersonic jet with a porous obstacle. *Journal of Applied Mechanics and Technical Physics*, 2015, vol. 56, no. 3, pp. 406–413. <https://doi.org/10.1134/s0021894415030098>

## Authors

**Mikhail V. Mikhailov** — Junior Researcher, Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation, <https://orcid.org/0009-0003-3840-203X>, [mikhailovmish@yandex.ru](mailto:mikhailovmish@yandex.ru)

**Nikolai V. Prodan** — PhD, Senior Researcher, Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation, [sc 56177927100](https://orcid.org/0000-0002-6383-039X), <https://orcid.org/0000-0002-6383-039X>, [kolinti@mail.ru](mailto:kolinti@mail.ru)

**Maxim E. Renev** — Junior Researcher, Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation, [sc 57211271545](https://orcid.org/0000-0002-0583-0607), <https://orcid.org/0000-0002-0583-0607>, [renewme@mail.ru](mailto:renewme@mail.ru)

doi: 10.17586/2226-1494-2023-23-4-843-849

УДК 532.529

## Точное решение задачи отражения ударной волны от стенки, экранированной слоем газозвеси

Дмитрий Викторович Садин<sup>1</sup>, Елена Николаевна Широкова<sup>2</sup><sup>1,2</sup> Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация<sup>1</sup> [sadin@yandex.ru](mailto:sadin@yandex.ru), <https://orcid.org/0000-0001-5335-4847><sup>2</sup> [shirokhelen-78@mail.ru](mailto:shirokhelen-78@mail.ru), <https://orcid.org/0000-0002-8188-2003>

### Аннотация

**Введение.** Представлено решение задачи отражения ударной волны от стенки, экранированной слоем газозвеси. Динамика газозвеси описана в двухскоростной двухтемпературной формулировке. **Метод.** В отличие от известных приближенных моделей запыленного газа, основанных на применении классических автомодельных решений путем коррекции газодинамических параметров и физических постоянных, получено асимптотически точное решение. Аналитическое решение поставленной задачи построено в виде композиции элементарных распадов разрывов. Неравновесное решение сходится к точному при уменьшении характерных времен динамической и тепловой релаксаций несущего газа и взвешенных частиц произвольной концентрации. Расчеты по неравновесной модели выполнены гибридным методом крупных частиц второго порядка аппроксимации по пространству и времени. **Основные результаты.** Приведены точные и расчетные по неравновесной модели профили относительных величин давления и плотности смеси, нормированной скорости дисперсной фазы. Изучено влияние интенсивности падающей ударной волны, а также концентрации частиц в слое газозвеси на параметры воздействия ударно-волнового импульса на стенку. Наличие экранирующего слоя привело к повышению давления отражения от стенки по сравнению с отражением ударной волны в чистом газе. Выполнен анализ влияния релаксационных свойств слоя газозвеси при изменении размеров частиц от 1 до 8 мкм. Для достаточно мелких частиц 1 мкм и принятых масштабов задачи неравновесное решение хорошо воспроизводит ударно-волновую структуру и соответствует асимптотике. С увеличением размеров дисперсных включений возрастают пространственные зоны релаксации, сглаживающие профили параметров. Погрешность расчета скорости и других параметров для неравновесной газозвеси с частицами 1 мкм по сравнению с точным решением находится в интервале от  $10^{-7}$  до  $10^{-5}$ . **Обсуждение.** Полученные результаты имеют прикладное значение при обосновании влияния примесей инертных частиц на динамическое нагружение конструкций. Аналитическое решение задачи может быть востребовано при тестировании различных численных схем.

### Ключевые слова

точное решение, отражение, ударная волна, стенка, слой газозвеси

**Ссылка для цитирования:** Садин Д.В., Широкова Е.Н. Точное решение задачи отражения ударной волны от стенки, экранированной слоем газозвеси // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 843–849. doi: 10.17586/2226-1494-2023-23-4-843-849

## The exact solution of a shock wave reflection problem from a wall shielded by a gas suspension layer

Dmitry V. Sadin<sup>1</sup>, Elena N. Shirokova<sup>2</sup><sup>1,2</sup> Mozhaisky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation<sup>1</sup> [sadin@yandex.ru](mailto:sadin@yandex.ru), <https://orcid.org/0000-0001-5335-4847><sup>2</sup> [shirokhelen-78@mail.ru](mailto:shirokhelen-78@mail.ru), <https://orcid.org/0000-0002-8188-2003>

### Abstract

The paper is devoted to solving the shockwave reflection problem from a wall shielded by a gas suspension layer. The dynamics of the gas suspension are described in a two-speed two-temperature formulation. In contrast to the known

approximate models of dusty gas based on the application of classical self-similar solutions by correcting gas dynamic parameters and physical constants, an asymptotically exact solution is obtained. The analytical solution to the problem is constructed in the form of a composition of elementary decays discontinuities. The nonequilibrium solution converges to the exact one with a decrease in the characteristic times of dynamic and thermal relaxation of the carrier gas and suspended particles of arbitrary concentration. Calculations based on the nonequilibrium model are performed by the hybrid large-particle method of the second-order approximation in space and time. Both for the exact and calculate profiles of the relative values of the pressure and density of the mixture, the normalized velocity of the dispersed phase obtained from the nonequilibrium model are given. The influence of the intensity of the incident shock wave, as well as the concentration of particles in the gas suspension layer on the parameters of the impact of the shock wave pulse on the wall, is studied. The presence of a shielding layer leads to an increase in the reflection pressure from the wall compared to the reflection of the shock wave in a pure gas. The analysis of the influence of the relaxation properties of the gas suspension layer with a change in particle sizes from 1 to 8  $\mu\text{m}$  is carried out. For sufficiently small particles of 1 micron and the accepted scales of the problem, the nonequilibrium solution reproduces the shock-wave structure well and corresponds to the asymptotics. With the increase in the size of dispersed inclusions, the spatial relaxation zones, smoothing the profiles of the parameters, increase. The error in calculating the velocity and other parameters for a nonequilibrium gas suspension with particles of 1  $\mu\text{m}$  compared to the exact solution is in the range from  $10^{-7}$  to  $10^{-5}$ . The results obtained are of practical importance in substantiating the influence of inert particle impurities on the dynamic loading of structures. The analytical solution to the problem may be in demand when testing various numerical schemes.

### Keywords

exact solution, reflection, shock wave, wall, layer of gas suspension

**For citation:** Sadin D.V., Shirokova E.N. The exact solution of a shock wave reflection problem from a wall shielded by a gas suspension layer. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 843–849 (in Russian). doi: 10.17586/2226-1494-2023-23-4-843-849

## Введение

Изучение ударно-волновых явлений в газозвесах актуально в технологиях импульсного нанесения покрытий, взрывоопасных производствах, а также представляет теоретический интерес при исследовании структурных и релаксационных свойств смесей газа с частицами. Данным вопросам посвящено большое количество научных работ, например работы, в которых изучены: релаксация ударной волны в газе с мелкими взвешенными частицами [1]; распространение и отражение ударных волн в смесях газа и частиц микронных размеров [2]; передача ударно-волновой нагрузки экранируемой плоской стенке через слой порошкообразной среды [3]; влияния свойств газовой составляющей взвеси твердых частиц на разлет сжатого объема газозвеси в двухкомпонентной среде [4]; взаимодействие плоской ударной волны с областями различной формы и плотности в мелкодисперсной газозвеси [5] и с плотным слоем частиц [6].

Важное место в волновой динамике газозвесей занимают аналитические методы исследования. Присутствие в газе взвешенных частиц усложняет построение решений. Работы в этом направлении основаны на модели, так называемого запыленного газа [1, 2, 7–9]. Этот подход заключается во введении эффективного показателя адиабаты в сочетании с имеющими определенный физический смысл ограничениями, при этом система нестационарных уравнений движения дисперсной системы упрощается до уравнений совершенного газа. В [10] доказано, что такая аналогия возможна, если объем газа мало отличается от объема смеси (концентрация частиц мала).

В случае большой объемной доли частиц дополнительной искомой переменной служит массовая или объемная концентрация какой-либо из фаз при одновременном увеличении числа уравнений на единицу. Аналогия становится неполной ввиду существенного отличия

уравнения состояния совершенного газа и двухфазной среды. Следовательно, применение известных автомодельных решений газовой динамики для двухфазных равновесных течений с произвольной концентрацией частиц путем изменения плотности, теплоемкостей и параметров подобия является некорректным.

Трудность построения точного решения волновой динамики газозвесей произвольной концентрации дисперсной фазы заключалась в получении инвариантов Римана, которые удалось разрешить в явном виде путем приведения к интегралу от дифференциального бинорма [11].

Настоящая работа посвящена асимптотически точному решению задачи отражения ударной волны от стенки, экранированной слоем газозвеси. Цель работы — анализ влияния ряда факторов (концентрации и размеров частиц, интенсивности начального скачка уплотнения) на параметры преломленного и отраженного ударно-волнового импульса, а также подтверждения численной сходимости неравновесного решения к точному.

## Неравновесная модель динамики газозвеси

Рассмотрим динамику газозвеси в двухскоростной двухтемпературной формулировке [12]:

$$\begin{aligned} \frac{\partial \rho_i}{\partial t} + \frac{\partial \rho_i u_i}{\partial x} &= 0, \quad \frac{\partial \rho_1 u_1}{\partial t} + \frac{\partial \rho_1 u_1^2}{\partial x} + \alpha_1 \frac{\partial p}{\partial x} = -F_\mu, \\ \frac{\partial \rho_2 u_2}{\partial t} + \frac{\partial \rho_2 u_2^2}{\partial x} + \alpha_2 \frac{\partial p}{\partial x} &= F_\mu, \quad \frac{\partial \rho_2 e_2}{\partial t} + \frac{\partial \rho_2 e_2 u_2}{\partial x} = Q_T, \quad (1) \\ & \frac{\partial}{\partial t} (\rho_1 E_1 + \rho_2 E_2) + \\ & + \frac{\partial}{\partial x} (\rho_1 E_1 u_1 + \rho_2 E_2 u_2 + p(\alpha_1 u_1 + \alpha_2 u_2)) = 0, \\ \rho_i &= \rho_i^\circ \alpha_i \quad (i = 1, 2), \quad \alpha_1 + \alpha_2 = 1, \quad E_i = e_i + u_i^2/2, \end{aligned}$$

где нижние индексы «1» и «2» — параметры несущей и дисперсной фаз; верхний индекс «°» — истинные значения плотности;  $\alpha_i, \rho_i, u_i, E_i, e_i$  и  $p$  — объемная доля, приведенная плотность, скорость, полная и внутренняя энергии единицы массы  $i$ -ой фазы, давление газа;  $F_\mu$  и  $Q_T$  — вязкая составляющая силы межфазного взаимодействия и мощность теплообмена между газом и частицами в единице объема;  $x$  и  $t$  — координата и время.

Замыкающими соотношениями системы (1) являются уравнения состояния идеального calorически совершенного газа и несжимаемых твердых частиц:  $p = (\gamma_1 - 1)\rho_1^\circ e_1, e_1 = c_v T_1, e_2 = c_2 T_2, \{\gamma_1, c_v, c_2, \rho_2^\circ\} \equiv \text{const}$ , где  $T_1$  и  $T_2$  — температуры несущей фазы и частиц;  $\gamma_1$  и  $c_v$  — показатель адиабаты и удельная теплоемкость газа при постоянном объеме;  $c_2$  — удельная теплоемкость частиц. Силовое и тепловое межфазное взаимодействие  $F_\mu, Q_T$  определим из эмпирических соотношений [13]:

$$F_\mu = (3/8)(\alpha_2/r)C_\mu(\text{Re}_{12})\rho_1(u_1 - u_2)|u_1 - u_2|,$$

$$C_\mu^{(1)} = \frac{24}{\text{Re}_{12}} + \frac{4,4}{\text{Re}_{12}^2} + 0,42 (\alpha_2 < 0,08),$$

$$Q_T = (3/2)(\alpha_2/r^2)\lambda_1 \text{Nu}_1 (T_1 - T_2),$$

$$\text{Nu}_1 = \begin{cases} 2 + 0,106\text{Re}_{12}\text{Pr}_1^{1/3} & (\text{Re}_{12} \leq 200), \\ 2,274 + 0,6\text{Re}_{12}^{0,67}\text{Pr}_1^{1/3} & (\text{Re}_{12} > 200), \end{cases}$$

$$\text{Re}_{12} = 2r\rho_1^\circ|v_1 - v_2|/\mu_1 \quad \text{Pr}_1 = c_v\gamma_1\mu_1/\lambda_1,$$

где  $\text{Re}_{12}, \text{Nu}_1$  и  $\text{Pr}_1$  — числа Рейнольдса, Нуссельта и Прандтля;  $C_\mu, \mu_1, \lambda_1$  и  $r$  — коэффициент межфазного трения, динамическая вязкость, коэффициент теплопроводности газа и радиус частицы.

### Постановка задачи

Плоская ударная волна с числом Маха  $M_1 \in [1,05; 1,50]$  и начальной координатой  $x = x_0 = 3$  м движется в воздухе в направлении жесткой стенки  $x = x_w = 8$  м (рис. 1). Стенка экранирована однородным слоем газозвеси  $x_1 \leq x \leq x_w$  ( $x_1 = 4$  м). Газозвесь с объемной долей  $\alpha_2 \in [10^{-10}; 10^{-3}]$  содержит монодисперсные частицы диаметром  $d \in [1; 8]$  мкм, плотностью  $\rho_2^\circ = 2500$  кг/м<sup>3</sup> и теплоемкостью частиц —  $c_2 = 710$  Дж/(кг·К). В начальный момент времени  $t = 0$  перед падающей ударной волной среда неподвижна и находится в термодинамическом равновесии с давлением  $p^{(j)} = 10^5$  Па и температурой фаз  $T_1^{(j)} = T_2^{(j)} = 293$  К ( $j = 0, 1$  номер области на рис. 1). Краевые условия заданы в виде условий непротекания на стенке  $x = x_w$  и параметров за падающей ударной волной при  $x = 0$ . Требуется найти аналитическое решение задачи отражения ударной волны от стенки.

### Построение точного решения

Найдем асимптотически точное решение, когда времена динамической  $t_1^{(\mu)}, t_2^{(\mu)}$  и тепловой  $t_1^{(T)}, t_2^{(T)}$  релак-

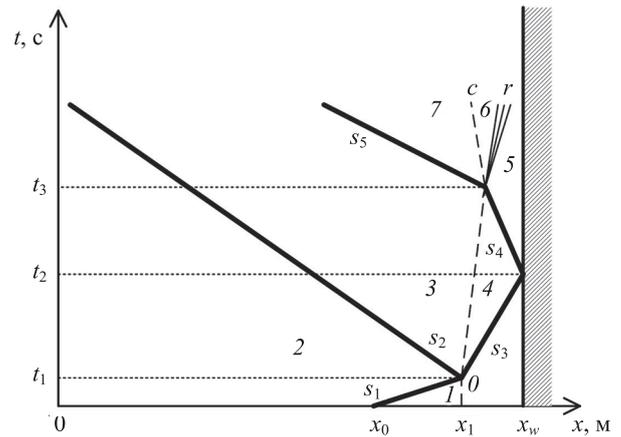


Рис. 1. Ударно-волновая диаграмма.

Траектории ударных волн (жирные линии); контактный разрыв (штриховые линии), веер волн разрежения (тонкие прямые). Области диаграммы: начальное состояние газа 1 и газозвеси 0; состояние 2 — за падающей  $s_1$ , 3 — отраженной от слоя  $s_2$  и 4 — преломленной  $s_3$  ударными волнами, 5 — за отраженной от стенки ударной волной  $s_4$  и 7 — после ее преломления  $s_5$  на границе слоя  $c$ ; 6 — зона постоянного течения за волной разрежения  $r$ ;  $t_1, t_3$  и  $t_2$  — моменты времени преломления и отражения ударных волн

Fig. 1. Shock wave diagram.

Trajectories of shock waves (bold lines); contact (dashed line); fan of rarefaction waves (thin lines).

Diagram regions: initial states of gas 1 and gas suspension 0; state 2 — behind the falling  $s_1$ , 3 — reflected from layer  $s_2$  and 4 — refracted  $s_3$  shock waves, 5 — behind the shock wave reflected from wall  $s_4$  and 7 — after its refraction  $s_5$  at the boundary of layer  $c$ ; 6 — constant flow zone behind rarefaction wave  $r$ ;  $t_1, t_3$  and  $t_2$  — moments of refraction and reflection of shock waves

саций фаз [14] по отношению к временному интервалу задачи  $t_f$  малы.

Параметры за падающей ударной волной  $s_1$  (рис. 1) при заданном числе Маха  $M_1$  определим по классическим формулам [15]. Общее решение построим с помощью композиции элементарных решений [16]: при  $t_1 \leq t < t_2$  (рис. 1) — распада разрыва с образованием отраженной от слоя газозвеси  $s_2$  и преломленной  $s_3$  ударных волн, при  $t_2 \leq t < t_3$  — отраженной от стенки ударной волны  $s_4$ , при  $t > t_3$  — преломленной на интерфейсной границе  $c$  ударной волны  $s_5$  и отраженной волны разрежения  $r$ .

Распад разрыва падающей ударной волны на границе слоя газозвеси. Рассчитаем давление  $P = p^{(3)} = p^{(4)}$  из соотношения

$$u^{(2)} - (P - p^{(2)}) \sqrt{\frac{\chi^{(2)} - 1}{\rho^{(2)}(\kappa^{(2)}P + p^{(2)})}} =$$

$$= (P - p^{(0)}) \sqrt{\frac{\chi^{(0)} - 1}{\rho^{(0)}(\kappa^{(0)}P + p^{(0)})}}.$$

Определим скорость  $U = u^{(3)} = u^{(4)}$ , плотность смеси слева  $\rho^{(3)}$  и справа  $\rho^{(4)}$  от контактного разрыва, а также скорости левой  $D_2$  и правой  $D_0$  ударных волн:

$$\begin{aligned}
 U &= u^{(2)} - (P - p^{(2)}) \sqrt{\frac{\chi^{(2)} - 1}{\rho^{(2)}(\kappa^{(2)}P + p^{(2)})}}, \\
 \rho^{(3)} &= \rho^{(2)} \frac{\kappa^{(2)}P + p^{(2)}}{\chi^{(2)}p^{(2)} + \frac{\gamma^{(2)} - 2\alpha_1^{(2)} + 1}{\gamma^{(2)} - 1} P}, \\
 \alpha_1^{(3)} &= 1 - \frac{\rho^{(3)}}{\rho^{(2)}} (1 - \alpha_1^{(2)}), \quad \alpha_2^{(3)} = 1 - \alpha_1^{(3)}, \\
 \rho^{(4)} &= \rho^{(0)} \frac{\kappa^{(0)}P + p^{(0)}}{\chi^{(0)}p^{(0)} + \frac{\gamma^{(0)} - 2\alpha_1^{(0)} + 1}{\gamma^{(0)} - 1} P}, \\
 \alpha_1^{(4)} &= 1 - \frac{\rho^{(4)}}{\rho^{(0)}} (1 - \alpha_1^{(0)}), \quad \alpha_2^{(4)} = 1 - \alpha_1^{(4)}, \quad (2)
 \end{aligned}$$

$$\begin{aligned}
 D_2 &= u^{(2)} - a^{(2)} \sqrt{\frac{\alpha_1^{(2)} \left( \frac{\gamma^{(2)} + 2\alpha_1^{(3)} - 1}{2\gamma^{(2)}} \frac{P}{p^{(3)}} + \frac{\gamma^{(2)} - 2\alpha_1^{(3)} + 1}{2\gamma^{(2)}} \right)}, \\
 D_0 &= a^{(0)} \sqrt{\frac{\alpha_1^{(0)} \left( \frac{\gamma^{(0)} + 2\alpha_1^{(4)} - 1}{2\gamma^{(0)}} \frac{P}{p^{(0)}} + \frac{\gamma^{(0)} - 2\alpha_1^{(4)} + 1}{2\gamma^{(0)}} \right)},
 \end{aligned}$$

где верхние индексы — параметры в соответствующих областях (рис. 1). Показатель политропы среды вычислим из условий неизменности концентраций фаз вдоль траекторий смеси  $\gamma = 1 + (\zeta_1 R_1)/(\zeta_1 c_v + \zeta_2 c_2)$ , где  $\zeta_i = \rho_i/\rho$  — массовые концентрации фаз,  $R_1$  — газовая постоянная. Отметим, что из  $d\zeta_i/dt = 0$  следует  $\gamma^{(1)} = \gamma^{(2)} = \gamma^{(3)} = \gamma^{(7)}$  и  $\gamma^{(0)} = \gamma^{(4)} = \gamma^{(5)} = \gamma^{(6)}$ . Скорость звука в газозвеси —  $a = \sqrt{\gamma p/(\rho \alpha_1)}$ . В (2) использованы плотность смеси  $\rho = \rho_1 + \rho_2$  и вспомогательные функции  $\chi^{(\cdot)} = (\gamma^{(\cdot)} + 2\alpha_1^{(\cdot)} - 1)/(\gamma^{(\cdot)} - 1)$  и  $\kappa^{(\cdot)} = (\gamma^{(\cdot)} + 1)/(\gamma^{(\cdot)} - 1)$ .

Полученные значения в области 4 являются исходными данными для расчета отражения ударной волны от стенки. Давление отражения  $p^{(5)}$  определим по формуле

$$u^{(4)} = (p^{(5)} - p^{(4)}) \sqrt{\frac{\chi^{(4)} - 1}{\rho^{(4)}(\kappa^{(4)}p^{(5)} + p^{(4)})}},$$

а остальные параметры из соотношений (2) при соответствующем изменении верхних индексов.

Распад разрыва отраженной от стенки ударной волны на границе слоя газозвеси. Используя рассчитанные параметры в областях 3 и 5, рассчитаем давление в зоне постоянного течения  $P = p^{(6)} = p^{(7)}$ :

$$\begin{aligned}
 u^{(3)} - u^{(5)} &+ \frac{2a^{(5)}\alpha_1^{(5)} \left( 1 - \left( \frac{P}{p^{(5)}} \right)^{\frac{\gamma^{(5)}-1}{2\gamma^{(5)}}} \right)}{\gamma^{(5)} - 1} = \\
 &= (P - p^{(3)}) \sqrt{\frac{\chi^{(3)} - 1}{\rho^{(3)}(\kappa^{(3)}P + p^{(3)})}}.
 \end{aligned}$$

Скорость  $U = u^{(6)} = u^{(7)}$  в зоне постоянного течения, плотность смеси слева  $\rho^{(7)}$  и справа  $\rho^{(6)}$  от интерфейсной границы, а также скорость левой  $D_3$  ударной волны определим из соотношений:

$$U = u^{(3)} - (P - p^{(3)}) \sqrt{\frac{\chi^{(3)} - 1}{\rho^{(3)}(\kappa^{(3)}P + p^{(3)})}},$$

$$\rho^{(7)} = \rho^{(3)} \frac{\kappa^{(3)}P + p^{(3)}}{\chi^{(3)}p^{(3)} + \frac{\gamma^{(3)} - 2\alpha_1^{(3)} + 1}{\gamma^{(3)} - 1} P},$$

$$\alpha_1^{(7)} = 1 - \frac{\rho^{(7)}}{\rho^{(3)}} (1 - \alpha_1^{(3)}), \quad \alpha_2^{(7)} = 1 - \alpha_1^{(7)},$$

$$\rho^{(6)} = \rho^{(5)} \left[ \alpha_2^{(5)} + \alpha_1^{(5)} \left( 1 - \frac{\gamma^{(5)} - 1 - U}{2\alpha_1^{(5)} a^{(5)}} \right)^{-\frac{2}{\gamma^{(5)}-1}} \right]^{-1},$$

$$\alpha_1^{(6)} = 1 - \frac{\rho^{(6)}}{\rho^{(5)}} (1 - \alpha_1^{(5)}), \quad \alpha_2^{(6)} = 1 - \alpha_1^{(6)},$$

$$D_3 = u^{(3)} - a^{(3)} \sqrt{\frac{\alpha_1^{(3)} \left( \frac{\gamma^{(3)} + 2\alpha_1^{(7)} - 1}{2\gamma^{(3)}} \frac{P}{p^{(3)}} + \frac{\gamma^{(3)} - 2\alpha_1^{(7)} + 1}{2\gamma^{(3)}} \right)}.$$

### Расчеты по неравновесной модели и сравнение с точным решением

Приведем результаты исследований поставленной задачи в неравновесной двухскоростной двухтемпературной формулировке (1) в сопоставлении с асимптотически точным аналитическим решением. Расчеты по неравновесной модели выполним гибридным методом крупных частиц второго порядка аппроксимации [17] на равномерной сетке с пространственным шагом  $h = 0,01$  м. Шаг по времени — переменный и определен числом Куранта  $CFL = 0,4$ . Для однородности вычислений в области «чистого» газа зададим объемную долю частиц пренебрежимо малой  $\alpha_2 = 10^{-10}$ . Реализуем алгоритм в виде асинхронного (параллельного) кода на языке Free Pascal с графической обработкой результатов в библиотеке Matplotlib.

На рис. 2 приведены точные и расчетные по неравновесной модели профили давления и плотности смеси, отнесенные к соответствующим значениям за падающей ударной волной (рис. 2, а) и скорости дисперсной фазы, нормированной начальной скоростью звука в «чистом» газе  $a_1^{(1)} = \sqrt{\gamma_1 p^{(1)}/\rho_1^{(1)}}$  (рис. 2, б). Результаты получены для ударной волны с числом Маха  $M_1 = 1,5$ .

Теоретический и прикладной интерес представляет влияние интенсивности падающей ударной волны  $M_1$ , а также концентрации частиц в слое газозвеси  $\alpha_2^{(0)}$  на параметры воздействия ударно-волнового импульса на стенку. Например, каково соотношение между давлениями за отраженной от стенки ударной волной, экранированной слоем газозвеси, и при его отсутствии?

Изменения давления и плотности смеси за отраженной от слоя и стенки ударных волн по отношению к соответствующим значениям за исходным скачком уплотнения показаны на рис. 3 в зависимости от начальной объемной концентрации частиц в слое  $\alpha_2^{(0)}$  и на рис. 4 при варьировании числа Маха  $M_1$ . Численные замеры параметров произведены в точках  $x_w' = x_0 - 0,01x_w$  (отражение от слоя газозвеси) и  $x_w' = 0,99x_w$  (отражение от стенки).

Из рис. 3 и 4 видно, что величины давления и плотности смеси за отраженной от стенки волной монотонно возрастают при увеличении интенсивности падающего скачка уплотнения и объемной доли дис-

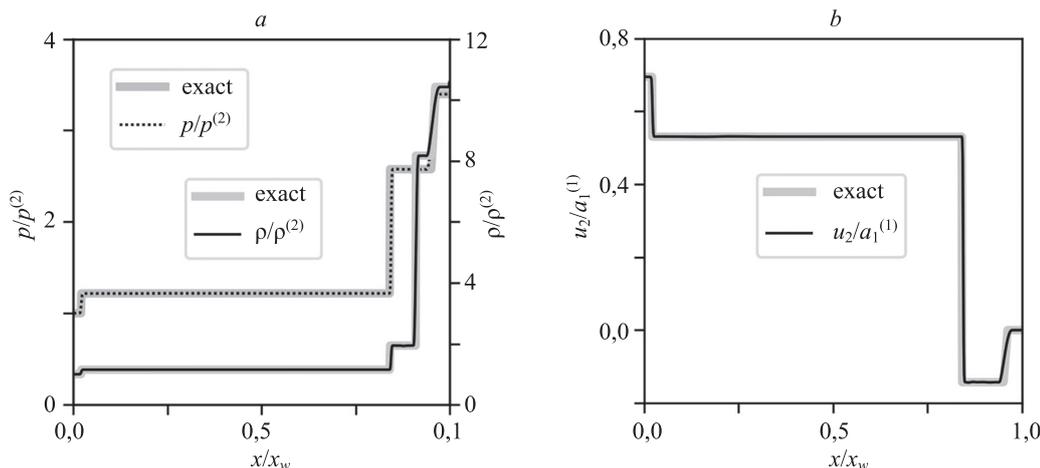


Рис. 2. Профили относительных величин давления, плотности смеси (а) и скорости частиц (б) при  $d = 1 \text{ мкм}$  в момент времени  $t_f = 22 \text{ мс}$ .

Точное решение (exact — серая сплошная линия); расчет на сетке 1/800 (тонкая сплошная линия) (плотность и скорость смеси); давление (пунктирная линия)

Fig. 2. Profiles of the relative value of pressure, mixture density (a), and particle velocity (b) at  $d = 1 \text{ }\mu\text{m}$  at a time  $t_f = 22 \text{ }\mu\text{s}$  (the exact solution is a bold solid curve; the calculations on the 1/800 grid are density and velocity of the mixture (a thin solid line) and pressure (dotted line))

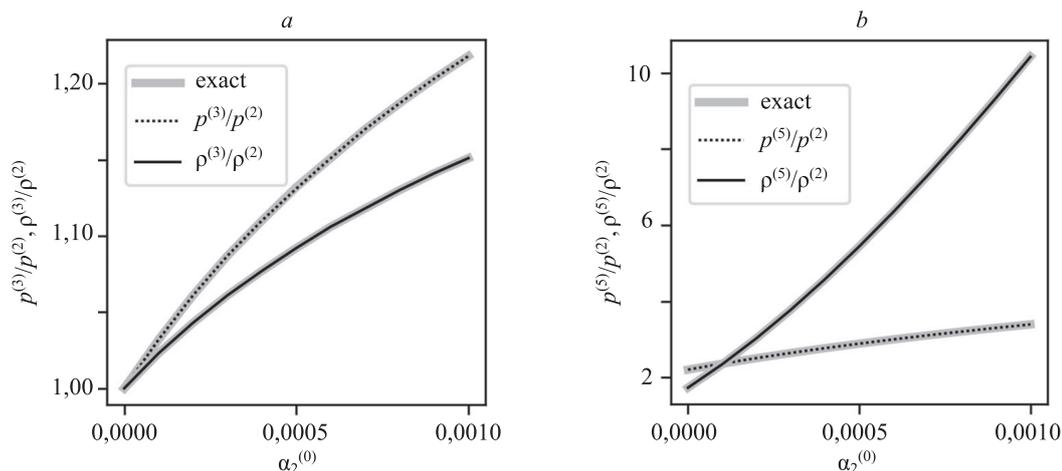


Рис. 3. Относительные значения давления и плотности смеси за отраженной от слоя (а) и стенки (б) ударных волн в зависимости от начальной объемной доли частиц  $\alpha_2^{(0)}$

Fig. 3. Relative values of the pressure and density of the mixture behind the shock waves reflected from the layer (a) and the wall (b) vs. the initial volume fraction of particles  $\alpha_2^{(0)}$

персной фазы экранирующего слоя. В практическом аспекте слой газозвеси не защищает от ударно-волнового импульса, а, напротив, увеличивает давление отражения. Например, при начальной концентрации частиц  $\alpha_2^{(0)} = 0,001$  и числе Маха  $M_1 = 1,5$  давление отражения составляет  $p^{(3)}/p^{(2)}|_{0,001} = 3,397$ , что в 1,5 раза больше случая отсутствия экранирующего слоя взвеси  $p^{(5)}/p^{(2)}|_0 = 2,207$ .

Влияние релаксационных свойств слоя газозвеси при изменении размеров частиц от 1 до 8 мкм для расчетной области в окрестности стенки показано на рис. 5, а. Для достаточно мелких частиц  $d < 1 \text{ мкм}$  и принятых масштабов задачи неравновесное решение хорошо воспроизводит ударно-волновую структуру и

соответствует асимптотике. С увеличением размеров дисперсных включений возрастают пространственные зоны релаксации, сглаживающие профили параметров (штриховая кривая для частиц диаметром  $d = 8 \text{ мкм}$ ).

Точность применяемого гибридного метода крупных частиц для рассматриваемой задачи показана на рис. 5, б. Здесь при решении задачи в неравновесной формулировке приведены численные значения модуля нормированной скорости дисперсной фазы в окрестности стенки после отражения ударной волны  $|u_2^{(5)}/a_1^{(1)}|$  (асимптотически точное значение равно нулю). Погрешность расчета скорости и других параметров для газозвеси с частицами  $d = 1 \text{ мкм}$  при числе Маха

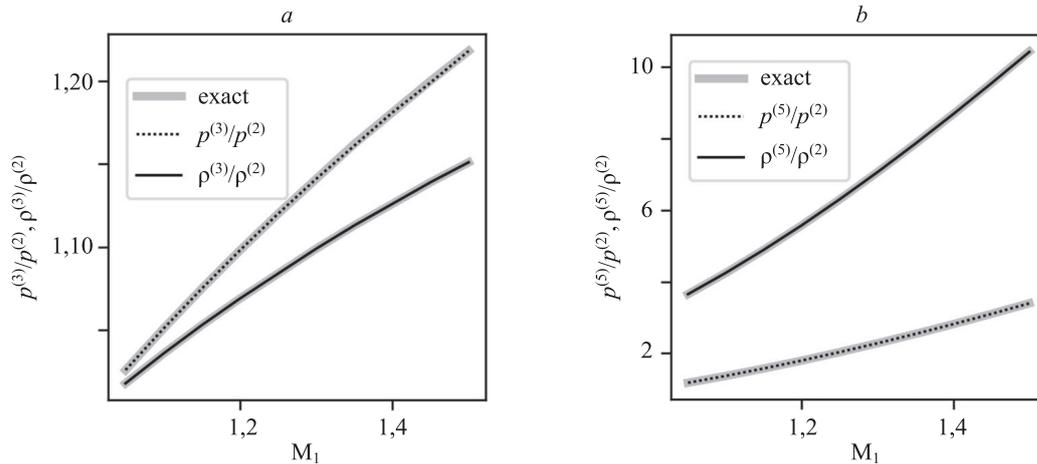


Рис. 4. Относительные значения давления и плотности смеси за отраженной от слоя (а) и стенки (б) ударных волн в зависимости от числа Маха падающей ударной волны  $M_1$

Fig. 4. Relative values of the pressure and density of the mixture behind the shock waves reflected from the layer (a) and the wall (b) vs. the Mach number of the incident shock wave  $M_1$

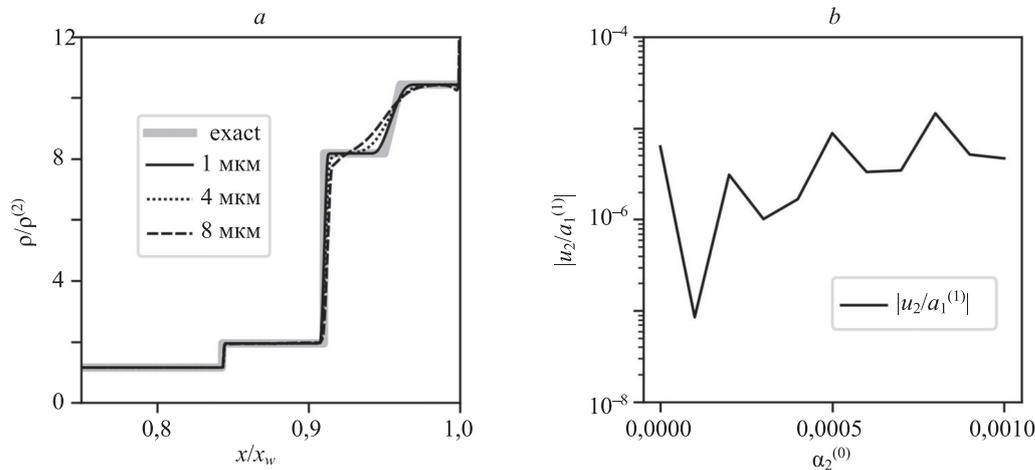


Рис. 5. Профили относительной плотности газозвеси для размеров частиц  $d = 1, 4, 8$  мкм (а) и погрешность расчета нормированной скорости частиц при  $d = 1$  мкм по неравновесной модели в окрестности стенки (б)

Fig. 5. Profiles of the relative density of the gas suspension for various particle sizes  $d = 1, 4, 8$   $\mu\text{m}$  (a) and errors in calculating the normalized particle velocity at  $d = 1$   $\mu\text{m}$  according to the nonequilibrium model in a vicinity of the wall (b)

падающей ударной волны  $M_1 = 1,5$  находится в интервале от  $10^{-7}$  до  $10^{-5}$ .

### Заключение

Построено асимптотически точное решение задачи отражения ударной волны от стенки, экранированной слоем газозвеси, как композиция элементарных решений распадов разрывов. Аналитическое решение

может представлять бенчмарк при проверке различных численных схем. Исследовано влияние интенсивности падающей ударной волны, концентрации частиц в слое газозвеси на параметры воздействия ударно-волнового импульса на стенку. Наличие экранирующего слоя привело к повышению давления отражения от стенки по сравнению с отражением ударной волны в чистом газе. Показана сходимость численных решений к точным при уменьшении размеров взвешенных частиц.

### Литература

1. Rudinger G. Some properties of shock relaxation in gas flows carrying small particles // *Physics of Fluids*. 1964. V. 7. N 5. P. 658–663. <https://doi.org/10.1063/1.1711265>
2. Sommerfeld M., Selzer M., Grönig H. Shock wave reflections in dusty-gas // *Proc of the 15<sup>th</sup> International Symposium on Shock Waves and Shock Tubes*. 1986. P. 683–689.
3. Болдырева О.Ю., Губайдуллин А.А., Дудко Д.Н., Кутушев А.Г. Численное исследование передачи ударно-волновой нагрузки

### References

1. Rudinger G. Some properties of shock relaxation in gas flows carrying small particles. *Physics of Fluids*, 1964, vol. 7, no. 5, pp. 658–663. <https://doi.org/10.1063/1.1711265>
2. Sommerfeld M, Selzer M, Grönig H. Shock wave reflections in dusty-gas. *Proc 15<sup>th</sup> International Symposium on Shock Waves and Shock Tubes*, 1986, pp. 683–689.
3. Boldyreva O.Y., Gubaidullin A.A., Dudko D.N., Kutushev A.G. Numerical study of the transfer of shock-wave loading to a screened

- экранируемой плоской стенке через слой порошкообразной среды и разделяющий их воздушный зазор // Физика горения и взрыва. 2007. Т. 43. № 1. С. 132–142.
4. Тукмаков Д.А. Численное исследование влияния свойств газовой составляющей взвеси твердых частиц на разлет сжатого объема газовой взвеси в двухкомпонентной среде // Инженерно-физический журнал. 2020. Т. 93. № 2. С. 304–310.
  5. Садин Д.В., Давидчук В.А. Взаимодействие плоской ударной волны с областями различной формы и плотности в мелкодисперсной газовой взвеси // Инженерно-физический журнал. 2020. Т. 93. № 2. С. 489–498.
  6. Волков К.Н., Емельянов В.Н., Ефремов А.В. Численное моделирование взаимодействия ударной волны с плотным слоем частиц // Инженерно-физический журнал. 2021. Т. 94. № 3. С. 658–667.
  7. Сидоркина С.И. О некоторых движениях аэрозоля // Доклады АН СССР. 1957. Т. 112. № 3. С. 398–400.
  8. Арутюнян Г.М. Термогидродинамическая теория гетерогенных систем. М.: Физматлит, 1994. 272 с.
  9. Георгиевский П.Ю., Левин В.А., Сутырин О.Г. Фокусировка ударной волны при взаимодействии ударной волны с цилиндрическим облаком пыли // Письма в Журнал технической физики. 2016. Т. 42. № 18. С. 17–24.
  10. Арутюнян Г.М. Условия применимости результатов гидродинамики совершенного газа к дисперсным средам // Известия АН СССР. Механика жидкости и газа. 1979. № 1. С. 157–160.
  11. Иванов А.С., Козлов В.В., Садин Д.В. Нестационарное истечение двухфазной дисперсной среды из цилиндрического канала конечных размеров в атмосферу // Известия РАН. Механика жидкости и газа. 1996. № 3. С. 60–66.
  12. Нигматулин Р.И. Динамика многофазных сред. Ч. 1. М.: Наука, 1987. 464 с.
  13. Ивандяев А.И., Кутушев А.Г., Рудаков Д.А. Численное исследование метания слоя порошка сжатым газом // Физика горения и взрыва. 1995. Т. 31. № 4. С. 63–70.
  14. Садин Д.В. О жесткости систем уравнений в частных производных, описывающих движения гетерогенных сред // Математическое моделирование. 2002. Т. 14. № 11. С. 43–53.
  15. Рождественский Б.Л., Яненко Н.Н. Системы квазилинейных уравнений и их приложения к газовой динамике. М.: Наука, 1978. 688 с.
  16. Садин Д.В. Тестовые задачи динамики газовой взвеси с использованием асимптотически точных решений // Математическое моделирование. 2022. Т. 34. № 12. С. 59–74. <https://doi.org/10.20948/mm-2022-12-04>
  17. Садин Д.В. Эффективная реализация гибридного метода крупных частиц // Математическое моделирование. 2022. Т. 34. № 4. С. 113–127. <https://doi.org/10.20948/mm-2022-04-08>
- flat wall through a layer of a powdered medium and a subsequent air gap. *Combustion, Explosion, and Shock Waves*, 2007, vol. 43, no. 1, pp. 114–123. <https://doi.org/10.1007/s10573-007-0016-3>
4. Tukmakov D.A. Numerical investigation of the influence of properties of the gas component of a suspension of solid particles on the spreading of a compressed gas-suspension volume in a binary medium. *Journal of Engineering Physics and Thermophysics*, 2020, vol. 93, no. 2, pp. 291–297. <https://doi.org/10.1007/s10891-020-02120-9>.
  5. Sadin D.V., Davidchuk V.A. Interaction of a plane shock wave with regions of varying shape and density in a finely divided gas suspension. *Journal of Engineering Physics and Thermophysics*, 2020, vol. 93, no. 2, pp. 474–483. <https://doi.org/10.1007/s10891-020-02143-2>
  6. Volkov K.N., Emel'yanov V.N., Efremov A.V. Numerical simulation of the interaction of a shock wave with a dense layer of particles. *Journal of Engineering Physics and Thermophysics*, 2021, vol. 94, no. 3, pp. 638–647. <https://doi.org/10.1007/s10891-021-02339-0>.
  7. Sidorkina S.I. Some aerosol motions. *Dokl. Akad. Nauk SSSR*, 1957, vol. 112, no. 3, pp. 398–400. (in Russian)
  8. Arutyunyan G.M. *Thermohydrodynamic Theory of Heterogeneous Systems*. Moscow, Fizmatlit Publ., 1994, 272 p. (in Russian)
  9. Georgievskiy P.Yu., Levin V.A., Sutyurin O.G. Shock focusing upon interaction of a shock with a cylindrical dust cloud. *Technical Physics Letters*, 2016, vol. 42, no. 9, pp. 936–939. <https://doi.org/10.1134/S1063785016090182>
  10. Arutyunyan G.M. Conditions of applicability of the results of the hydrodynamics of a perfect gas to disperse media. *Fluid Dynamics*, 1979, vol. 14, no. 4, pp. 118–121. <https://doi.org/10.1007/BF01050823>
  11. Ivanov A.S., Kozlov V.V., Sadin D.V. Unsteady flow of a two-phase disperse medium from a cylindrical channel of finite dimensions into the atmosphere. *Fluid Dynamics*, 1996, vol. 31, no. 3, pp. 386–391. <https://doi.org/10.1007/BF02030221>
  12. Nigmatulin R.I. *Dynamics of Multiphase Media. Vol. 1*. New York, USA, Hemisphere Publ. Corp., 1990, 507 p.
  13. Ivandayev A.I., Kutushev A.G., Rudakov D.A. Numerical investigation of throwing a powder layer by a compressed gas. *Combustion, Explosion, and Shock Waves*, 1995, vol. 31, no. 4, pp. 459–465. <https://doi.org/10.1007/BF00789367>.
  14. Sadin D.V. On stiff systems of partial differential equations for motion of heterogeneous media. *Matematicheskoe modelirovanie*, 2002, vol. 14, no. 11, pp. 43–53. (in Russian)
  15. Roždestvenskii B.L., Janenko N.N. *Systems of Quasilinear Equations and Their Applications to Gas Dynamics*. American Mathematical Soc., 1983, 676 p.
  16. Sadin D.V. Test problems of gas suspension dynamics using asymptotically exact solutions. *Mathematical Models and Computer Simulations*, 2023, vol. 15, no. 3, pp. 564–573. <https://doi.org/10.1134/S2070048223030158>
  17. Sadin D.V. Efficient implementation of the hybrid large particle method. *Mathematical Models and Computer Simulations*, 2022, vol. 14, no. 6, pp. 946–954. <https://doi.org/10.1134/S207004822206014X>

### Авторы

**Садин Дмитрий Викторович** — доктор технических наук, профессор, профессор, Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация, [orcid.org/0000-0001-5335-4847](https://orcid.org/0000-0001-5335-4847), [sadin@yandex.ru](mailto:sadin@yandex.ru)  
**Широкова Елена Николаевна** — кандидат химических наук, преподаватель, Военно-космическая академия имени А.Ф. Можайского, Санкт-Петербург, 197198, Российская Федерация, <https://orcid.org/0000-0002-8188-2003>, [shirokhelen-78@mail.ru](mailto:shirokhelen-78@mail.ru)

Статья поступила в редакцию 30.05.2023  
 Одобрена после рецензирования 15.06.2023  
 Принята к печати 18.07.2023

### Authors

**Dmitry V. Sadin** — D.Sc., Full Professor, Mozhaisky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation, [orcid.org/0000-0001-5335-4847](https://orcid.org/0000-0001-5335-4847), [sadin@yandex.ru](mailto:sadin@yandex.ru)

**Elena N. Shirokova** — PhD (Chemistry), Lecturer, Mozhaisky Military Aerospace Academy, Saint Petersburg, 197198, Russian Federation, <https://orcid.org/0000-0002-8188-2003>, [shirokhelen-78@mail.ru](mailto:shirokhelen-78@mail.ru)

Received 30.05.2023  
 Approved after reviewing 15.06.2023  
 Accepted 18.07.2023



Работа доступна по лицензии  
 Creative Commons  
 «Attribution-NonCommercial»

## КРАТКИЕ СООБЩЕНИЯ

## BRIEF PAPERS

doi: 10.17586/2226-1494-2023-23-4-850-853

УДК 681.5.015

**Адаптивный наблюдатель переменных состояния нелинейной  
нестационарной системы с неизвестными постоянными параметрами  
и запаздыванием в канале измерений**

**Алексей Алексеевич Бобцов<sup>1</sup>, Николай Анатольевич Николаев<sup>2</sup>✉,  
Ольга Андреевна Козачёк<sup>3</sup>, Ольга Владимировна Оськина<sup>4</sup>**

<sup>1,2,3,4</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>1</sup> bobtsov@mail.ru, <https://orcid.org/0000-0003-1854-6717>

<sup>2</sup> nanikolaev@itmo.ru✉, <https://orcid.org/0000-0002-8835-5142>

<sup>3</sup> oakozachek@itmo.ru, <https://orcid.org/0009-0008-8613-2835>

<sup>4</sup> ov\_oskina@itmo.ru, <https://orcid.org/0009-0005-5121-0432>

**Аннотация**

Рассмотрена задача оценки неизвестных постоянных параметров нелинейной нестационарной системы в условиях запаздывания в канале измерений. Целью работы является синтез адаптивного наблюдателя для нелинейной нестационарной системы, обеспечивающего асимптотическую сходимость оценок неизвестных постоянных параметров к истинным значениям. Предложен метод оценивания неизвестных постоянных параметров нелинейной нестационарной системы, базирующийся на технологии GPEBO (Generalized Parameter Estimation Based Observer). На основе технологии GPEBO выполнена параметризация исходной динамической системы к виду линейной регрессионной модели с последующей идентификацией неизвестных параметров. Для оценивания неизвестных параметров линейной регрессионной модели применен метод наименьших квадратов с фактором забывания. В рамках работы предложено расширение предыдущих результатов авторского коллектива на случай нелинейных нестационарных систем с запаздыванием в канале измерений. Предложенный алгоритм оценки параметров может использоваться для решения прикладных задач, таких как контроль технического состояния, а также в задачах синтеза систем автоматического управления.

**Ключевые слова**

идентификация параметров, линейная регрессия, запаздывание

**Благодарности**

Работа подготовлена при финансовой поддержке Российского научного фонда, грант № 22-21-00499, <https://rscf.ru/project/22-21-00499>.

**Ссылка для цитирования:** Бобцов А.А., Николаев Н.А., Козачёк О.А., Оськина О.В. Адаптивный наблюдатель переменных состояния нелинейной нестационарной системы с неизвестными постоянными параметрами и запаздыванием в канале измерений // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 850–853. doi: 10.17586/2226-1494-2023-23-4-850-853

**Adaptive observer for state variables of a time-varying nonlinear system  
with unknown constant parameters and delayed measurements**

**Alexey A. Bobtsov<sup>1</sup>, Nikolay A. Nikolaev<sup>2</sup>✉, Olga A. Kozachek<sup>3</sup>, Olga V. Oskina<sup>4</sup>**

<sup>1,2,3,4</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>1</sup> bobtsov@mail.ru, <https://orcid.org/0000-0003-1854-6717>

<sup>2</sup> nanikolaev@itmo.ru✉, <https://orcid.org/0000-0002-8835-5142>

<sup>3</sup> oakozachek@itmo.ru, <https://orcid.org/0009-0008-8613-2835>

<sup>4</sup> ov\_oskina@itmo.ru, <https://orcid.org/0009-0005-5121-0432>

© Бобцов А.А., Николаев Н.А., Козачёк О.А., Оськина О.В., 2023

**Abstract**

Unknown constant parameters estimation problem for a nonlinear time-varying system with delayed measurements is considered. The objective of this work is to design an adaptive observer for a nonlinear time-varying system. The observer must provide asymptotic convergence of the unknown constant parameters estimates to their true values. The main idea behind the method is to perform the parametrization of initial dynamical system based on GPEBO (Generalized Parameter Estimation Based Observer) technology and to build a linear regression model. The identification of linear regression model unknown parameters is performed using least square method with forgetting factor. This work develops the previously published approach for the case of nonlinear time-varying systems with delayed measurements. New parameters estimation algorithm can be applied for technical tasks, such as technical condition control and automatic control systems design.

**Keywords**

parameters identification, linear regression, delay

**Acknowledgements**

This work was supported by Russian Science Foundation, project no. 22-21-00499, <https://rscf.ru/project/22-21-00499>.

**For citation:** Bobtsov A.A., Nikolaev N.A., Kozachek O.A., Oskina O.V. Adaptive observer for state variables of a time-varying nonlinear system with unknown constant parameters and delayed measurements. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 850–853 (in Russian). doi: 10.17586/2226-1494-2023-23-4-850-853

Предложено развитие результата работы [1] для случая, когда в канале измерения присутствует известное постоянное запаздывание (т. е. выходная переменная измерена с некоторой временной задержкой). В рамках обобщенного подхода к синтезу наблюдателей [2], основанного на оценке постоянных параметров, предложен наблюдатель, обеспечивающий асимптотическую сходимость оценок неизвестных постоянных параметров нелинейной нестационарной системы к истинным значениям в условии наличия запаздывания в канале измерений.

Рассмотрим нелинейную нестационарную систему с одним входом и одним выходом (SISO) вида:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{A}(t)\mathbf{x}(t) + \mathbf{k}C^T(t)\mathbf{x}(t) + \mathbf{b}u(t) + \mathbf{w}(C^T(t)\mathbf{x}(t), t), \\ y(t) &= C^T(\varphi(t))\mathbf{x}(\varphi(t)), \end{aligned} \quad (1)$$

где  $\mathbf{x}(t) \in \mathbb{R}^n$  — неизвестный вектор состояния;  $u(t) \in \mathbb{R}$  — известный входной сигнал;  $y(t) \in \mathbb{R}$  — измеряемый выходной сигнал;  $\mathbf{A}(t) \in \mathbb{R}^{n \times n}$  и  $C^T(t) \in \mathbb{R}^n$  — известные матрицы с ограниченными во времени нестационарными параметрами;  $\mathbf{k} \in \mathbb{R}^n$  и  $\mathbf{b} \in \mathbb{R}^n$  — постоянные и неизвестные векторы;  $\mathbf{w}(C^T(t)\mathbf{x}(t), t)$  — частично неизвестная нелинейная вектор-функция;  $\varphi(t)$  — известная неотрицательная функция, определяющая запаздывание в канале измерений

$$\varphi(t) = t - d, \varphi(t) \geq 0, \quad (2)$$

где  $d > 0$  — постоянное запаздывание.

В отношении рассматриваемой системы при решении поставленной задачи применим следующие типовые допущения (например, [3–5]).

**Допущение 1.** Нелинейная вектор-функция  $\mathbf{w}(C^T(t)\mathbf{x}(t), t)$  может быть представлена в виде  $\mathbf{w}(C^T(t)\mathbf{x}(t), t) = \mathbf{m}f(C^T(t)\mathbf{x}(t))$ , где  $f(C^T(t)\mathbf{x}(t))$  — известная скалярная нелинейная функция, а  $\mathbf{m} \in \mathbb{R}^n$  — вектор неизвестных постоянных параметров.

**Допущение 2.** Предполагается, что сигнал  $u(t)$  ограничен и при этом траектории вектора переменных состояния  $\mathbf{x}(t)$  также ограничены.

**Допущение 3.** Пара матриц  $\mathbf{A}(t)$  и  $C^T(t)$  обнаруживаема, т. е. существует вектор обратной связи  $\mathbf{L}(t)$  такой, что автономная система

$$\dot{\mathbf{x}}(t) = [\mathbf{A}(t) - \mathbf{L}(t)C^T(t)]\mathbf{x}(t)$$

является асимптотически устойчивой.

**Допущение 4.** Автономная система  $\dot{\mathbf{x}}(t) = \mathbf{A}_0(t)\mathbf{x}(t)$ , где  $\mathbf{A}_0(t) = \mathbf{A}(t) - \mathbf{L}(t)C^T(t)$  — равномерно устойчива (uniformly stable). Другими словами, ее фундаментальная матрица удовлетворяет условию [6, теорема 6.4]:

$$\|\Phi_{A_0}(t, \tau)\| \leq c_1, \forall t \geq \tau \geq 0.$$

Для системы (1) определим задачу синтеза адаптивного наблюдателя вида:

$$\begin{aligned} \dot{\chi}(t) &= F(\chi(t), u(t), y(t)), \\ [\hat{\mathbf{x}}(t) \quad \hat{\mathbf{k}}(t) \quad \hat{\mathbf{b}}(t) \quad \hat{\mathbf{m}}(t)]^T &= S(\chi(t), u(t), y(t)), \end{aligned}$$

где  $\chi(t) \in \mathbb{R}^{n_\chi}$  — все сигналы вектора  $\chi(t)$  ограничены;  $\hat{\mathbf{x}}(t)$ ,  $\hat{\mathbf{k}}(t)$ ,  $\hat{\mathbf{b}}(t)$  и  $\hat{\mathbf{m}}(t)$  — текущие оценки, соответственно,  $\mathbf{x}(t)$ ,  $\mathbf{k}$ ,  $\mathbf{b}$  и  $\mathbf{m}$ .

Адаптивный наблюдатель должен обеспечивать сходимость оценок переменных состояния и постоянных неизвестных параметров к реальным значениям, а именно:

$$\hat{\mathbf{x}}(t) = \mathbf{x}(t), \hat{\mathbf{k}}(t) = \mathbf{k}, \hat{\mathbf{b}}(t) = \mathbf{b}, \hat{\mathbf{m}}(t) = \mathbf{m},$$

для всех  $\mathbf{x}_0 \in \mathbb{R}^n$ ,  $\chi(t) \in \mathbb{R}^{n_\chi}$ .

Для решения поставленной задачи, по аналогии с [5], рассмотрим систему (1) в момент времени  $t - d$

$$\dot{\mathbf{x}}_d = \mathbf{A}_d \mathbf{x}_d + \mathbf{k}C_d^T \mathbf{x}_d + \mathbf{b}u_d + \mathbf{w}_d, y(t) = C_d^T \mathbf{x}_d, \quad (3)$$

где  $\mathbf{x}_d = \mathbf{x}(\varphi(t))$ ,  $\mathbf{A}_d = \mathbf{A}(\varphi(t))$ ,  $C_d^T = C^T(\varphi(t))$ ,  $u_d = u(\varphi(t))$ ,  $\mathbf{w}_d = \mathbf{m}f(C_d^T \mathbf{x}_d)$ .

Для оценки неизвестных параметров системы (3) используем технологию Generalized Parameter

Estimation Based Observer и по аналогии с работой [1] рассмотрим уравнения вида:

$$\dot{\xi}(t) = \mathbf{A}_{0d}\xi(t) + \mathbf{L}_d y(t), \xi(0) = \mathbf{0}_{n \times 1}, \quad (4)$$

$$\dot{\eta}(t) = \mathbf{A}_{0d}\eta(t) + \mathbf{I}y(t), \eta(0) = \mathbf{0}_{n \times n}, \quad (5)$$

$$\dot{\zeta}(t) = \mathbf{A}_{0d}\zeta(t) + \mathbf{I}u_d, \zeta(0) = \mathbf{0}_{n \times n}, \quad (6)$$

$$\dot{\rho}(t) = \mathbf{A}_{0d}\rho(t) + \mathbf{I}f_d, \rho(0) = \mathbf{0}_{n \times n}, \quad (7)$$

$$\dot{\Phi}(t) = \mathbf{A}_{0d}\Phi(t), \Phi(0) = \mathbf{I}_{n \times n}, \quad (8)$$

где  $\mathbf{A}_{0d} = \mathbf{A}_0(\varphi(t))$  и  $\mathbf{L}_d = \mathbf{L}(\varphi(t))$  — матрицы;  $\mathbf{I}$  — единичная матрица соответствующей размерности.

Таким образом, после несложных математических преобразований (по аналогии с [1]), исходную динамическую систему (1) преобразуем к линейной регрессионной модели вида:

$$\mathbf{z}(t) = \Psi(t)\Theta, \quad (9)$$

где  $\mathbf{z}(t) = y(t) - \mathbf{C}_d^T \xi(t)$  — измеряемый сигнал;  $\Psi(t) = [\mathbf{C}_d^T \Phi(t) \quad \mathbf{C}_d^T \eta(t) \quad \mathbf{C}_d^T \zeta(t) \quad \mathbf{C}_d^T \rho(t)]$  — вектор известных функций;  $\Theta = [\theta \quad \mathbf{k} \quad \mathbf{b} \quad \mathbf{m}]^T$  — вектор неизвестных постоянных параметров,  $\theta = \mathbf{x}(0)$ .

восстановлен вектор состояния исходной динамической системы (1).

Для иллюстрации работоспособности предложенного подхода выполним компьютерное моделирование. При моделировании для системы (1) были выбраны

следующие параметры:  $\mathbf{A}(t) = \begin{bmatrix} 2 - \sin t & 1 \\ -8 + \cos t & 0 \end{bmatrix}$ ,  $\mathbf{b} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ,  $\mathbf{k} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$ ,  $\mathbf{C}(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ,  $\mathbf{w}(y, t) = \mathbf{m} \sin(\mathbf{C}^T(t)\mathbf{x}(t))$ , где  $\mathbf{m} = \begin{bmatrix} -4 \\ 4 \end{bmatrix}$ .

Начальные условия вектора состояния  $\mathbf{x}(0) = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$ .  
Используя вектор обратной связи  $\mathbf{L}(t) = \begin{bmatrix} 2 - \sin(t) \\ 1 + \cos(t) \end{bmatrix}$ ,

получим  $\mathbf{A}_0(t) = \begin{bmatrix} 0 & 1 \\ -9 & 0 \end{bmatrix}$ .

Для проверки работоспособности предложенного алгоритма (4)–(8) совместно с методом наименьших квадратов с фактором забывания для оценки параметров модели (1) применим следующие параметры  $\alpha = 10^5$ ,  $M = 10^6$ ,  $\beta = 1$ ,  $f_0 = 0,1$ . При этом на вход системы был подан синусоидальный сигнал  $u(t) = \sin(t)$ .

Результаты моделирования (рисунок) подтвердили достижение поставленной цели.

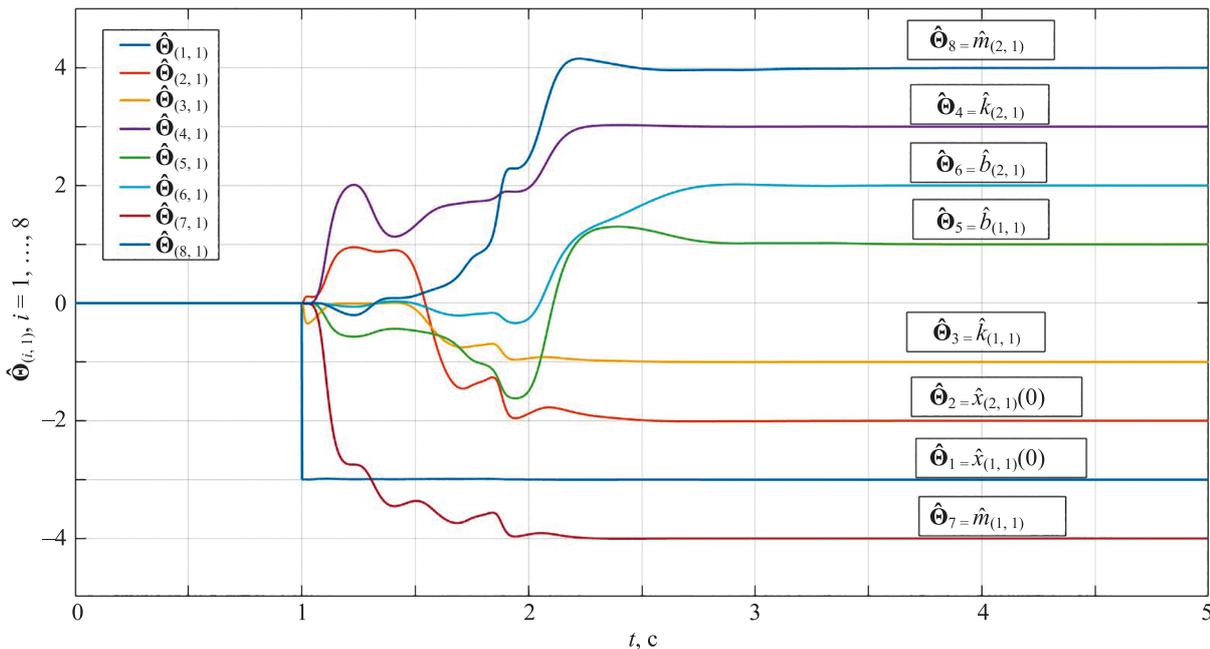


Рисунок. Переходные процессы по оценкам неизвестных параметров

Figure. Transients of unknown parameters estimates

Для оценки неизвестных постоянных параметров модели (9) применим метод наименьших квадратов с фактором забывания (forgetting factor) [7, 8]. После получения оценок неизвестных параметров может быть

В работе предложено развитие результата [1] для случая, когда выходная переменная динамической системы доступна измерению с известным постоянным запаздыванием вида (2). Результаты моделирования продемонстрировали работоспособность предложенного алгоритма.

### Литература

1. Козачёк А.А., Бобцов А.А., Николаев Н.А. Адаптивный наблюдатель переменных состояния нелинейной нестационарной системы с неизвестными постоянными параметрами // arXiv. 2023. arXiv:2305.15504. <https://doi.org/10.48550/arXiv.2305.15504>
2. Ortega R., Bobtsov A., Nikolaev N., Schiffer J., Dochain D. Generalized parameter estimation-based observers: Application to power systems and chemical-biological reactors // *Automatica*. 2021. V. 129. P. 109635. <https://doi.org/10.1016/j.automatica.2021.109635>
3. Tranninger M., Seeber R., Zhuk S., Steinberger M., Horn M. Detectability analysis and observer design for linear time varying systems // *IEEE Control Systems Letters*. 2020. V. 4. N 2. P. 331–336. <https://doi.org/10.1109/lcsys.2019.2927549>
4. Tranninger M., Zhuk S., Steinberger M., Fridman L., Horn M. Non-uniform stability, detectability, and, sliding mode observer design for time varying systems with unknown inputs // arXiv. 2018. arXiv:1809.06460. <https://doi.org/10.48550/arXiv.1809.06460>
5. Bobtsov A., Nikolaev N., Slita O., Kozachek O., Oskina O. Adaptive observer for a LTV system with partially unknown state matrix and delayed measurements // *Proc. of the 14<sup>th</sup> International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. 2022. P. 165–170. <https://doi.org/10.1109/ICUMT57764.2022.9943429>
6. Rugh W.J. *Linear System Theory*. Prentice-Hall, Inc., 1996. 581 p.
7. Ljung L. *System identification // Signal Analysis and Prediction*. Birkhäuser, Boston, MA, 1998. P. 163–173. [https://doi.org/10.1007/978-1-4612-1768-8\\_11](https://doi.org/10.1007/978-1-4612-1768-8_11)
8. Sastry S., Bodson M. *Adaptive Control: Stability, Convergence and Robustness*. New Jersey: Prentice-Hall, 1989. 377 p.

### Авторы

**Бобцов Алексей Алексеевич** — доктор технических наук, профессор, директор мегафакультета, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 8046819200](https://orcid.org/0000-0003-1854-6717), <https://orcid.org/0000-0003-1854-6717>, bobtsov@mail.ru

**Николаев Николай Анатольевич** — кандидат технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 13105019100](https://orcid.org/0000-0002-8835-5142), <https://orcid.org/0000-0002-8835-5142>, nanikolaev@itmo.ru

**Козачёк Ольга Андреевна** — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57219308287](https://orcid.org/0009-0008-8613-2835), <https://orcid.org/0009-0008-8613-2835>, oakozachek@itmo.ru

**Оськина Ольга Владимировна** — студент, инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57353555800](https://orcid.org/0009-0005-5121-0432), <https://orcid.org/0009-0005-5121-0432>, ov\_oskina@itmo.ru

Статья поступила в редакцию 12.05.2023  
Одобрена после рецензирования 19.06.2023  
Принята к печати 24.07.2023

### References

1. Kozachek O., Bobtsov A., Nikolaev N. Adaptive observer of state variables of a nonlinear time varying system with unknown constant parameters. *arXiv*, 2023, arXiv:2305.15504. (in Russian). <https://doi.org/10.48550/arXiv.2305.15504>
2. Ortega R., Bobtsov A., Nikolaev N., Schiffer J., Dochain D. Generalized parameter estimation-based observers: Application to power systems and chemical-biological reactors. *Automatica*, 2021, vol. 129, pp. 109635. <https://doi.org/10.1016/j.automatica.2021.109635>
3. Tranninger M., Seeber R., Zhuk S., Steinberger M., Horn M. Detectability analysis and observer design for linear time varying systems. *IEEE Control Systems Letters*, 2020, vol. 4, no. 2, pp. 331–336. <https://doi.org/10.1109/lcsys.2019.2927549>
4. Tranninger M., Zhuk S., Steinberger M., Fridman L., Horn M. Non-uniform stability, detectability, and, sliding mode observer design for time varying systems with unknown inputs. *arXiv*, 2018, arXiv:1809.06460. <https://doi.org/10.48550/arXiv.1809.06460>
5. Bobtsov A., Nikolaev N., Slita O., Kozachek O., Oskina O. Adaptive observer for a LTV system with partially unknown state matrix and delayed measurements. *Proc. of the 14<sup>th</sup> International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, 2022, pp. 165–170. <https://doi.org/10.1109/ICUMT57764.2022.9943429>
6. Rugh W.J. *Linear System Theory*. Prentice-Hall, Inc., 1996, 581 p.
7. Ljung L. *System identification. Signal Analysis and Prediction*. Birkhäuser, Boston, MA, 1998, pp. 163–173. [https://doi.org/10.1007/978-1-4612-1768-8\\_11](https://doi.org/10.1007/978-1-4612-1768-8_11)
8. Sastry S., Bodson M. *Adaptive Control: Stability, Convergence and Robustness*. New Jersey, Prentice-Hall, 1989, 377 p.

### Authors

**Alexey A. Bobtsov** — D.Sc., Professor, Director of School of Computer Technologies and Control, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 8046819200](https://orcid.org/0000-0003-1854-6717), <https://orcid.org/0000-0003-1854-6717>, bobtsov@mail.ru

**Nikolay A. Nikolaev** — PhD, Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 13105019100](https://orcid.org/0000-0002-8835-5142), <https://orcid.org/0000-0002-8835-5142>, nanikolaev@itmo.ru

**Olga A. Kozachek** — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57219308287](https://orcid.org/0009-0008-8613-2835), <https://orcid.org/0009-0008-8613-2835>, oakozachek@itmo.ru

**Olga V. Oskina** — Student, Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57353555800](https://orcid.org/0009-0005-5121-0432), <https://orcid.org/0009-0005-5121-0432>, ov\_oskina@itmo.ru

Received 12.05.2023  
Approved after reviewing 19.06.2023  
Accepted 24.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2023-23-4-854-857

## RuLegalNER: a new dataset for Russian legal named entities recognition Zein Shaheen<sup>1</sup>, Dmitry I. Mouromtsev<sup>2</sup>✉, Ignat Postny<sup>3</sup>

<sup>1,2</sup> ITMO University, Saint Petersburg, 197101, Russian Federation

<sup>3</sup> T.A.G. Consulting, Moscow, 119119, Russian Federation

<sup>1</sup> shaheen@itmo.ru, <https://orcid.org/0000-0001-6802-2896>

<sup>2</sup> mouromtsev@itmo.ru✉, <https://orcid.org/0000-0002-0644-9242>

<sup>3</sup> ipostny@gmail.com, <https://orcid.org/0009-0005-9249-4160>

### Abstract

We address the scarcity of datasets specifically tailored for legal NER in the Russian language and investigate the generalization capabilities of models towards unseen named entities. A rule-based program developed by legal experts at Tag-Consulting Company was employed to automatically annotate legal texts and create the RuLegalNER dataset. Part of the named entities only exists in the development and test splits, and they are unseen in the training set. RuBERT was utilized as the base architecture for experimental evaluation. Two different architectural extensions were explored: RuBERT with CRF and RuBERT with adapters. These architectures were used to train and evaluate NER models on the RuLegalNER dataset. Utilize RuLegalNER to train and evaluate legal NER models, enhancing performance in the legal domain and studying generalization on unseen entities. A published version of RuLegalNER is presented with detailed statistics and demonstration of the usefulness of RuLegalNER by evaluating modern architectures.

### Keywords

legal named entity recognition, natural language processing, information extraction, low-resource languages, transfer learning, transformers

**For citation:** Shaheen Z., Mouromtsev D.I., Postny I. RuLegalNER: a new dataset for Russian legal named entities recognition. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2023, vol. 23, no. 4, pp. 854–857. doi: 10.17586/2226-1494-2023-23-4-854-857

УДК 004.912

## RuLegalNER: новый датасет для распознавания именованных юридических сущностей на русском языке

Зейн Шахин<sup>1</sup>, Дмитрий Ильич Муромцев<sup>2</sup>✉, Игнат Постный<sup>3</sup>

<sup>1,2</sup> Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

<sup>3</sup> T.A.G. Consulting, Москва, 119119, Российская Федерация

<sup>1</sup> shaheen@itmo.ru, <https://orcid.org/0000-0001-6802-2896>

<sup>2</sup> mouromtsev@itmo.ru✉, <https://orcid.org/0000-0002-0644-9242>

<sup>3</sup> ipostny@gmail.com, <https://orcid.org/0009-0005-9249-4160>

### Аннотация

Представлен новый датасет RuLegalNER, разработанный для обучения моделей распознавания именованных юридических сущностей на русском языке. Выполнена оценка способности моделей к обобщению при появлении в тексте ранее не встречавшихся именованных сущностей. Для автоматической разметки юридических текстов и создания набора данных RuLegalNER разработана программа на основе правил. Часть именованных сущностей в датасете была выделена в набор данных для валидации и тестирования и не встречается в обучающем наборе. Экспериментальная проверка датасета основана на базовой архитектуре RuBERT. Исследовано два расширения архитектуры: RuBERT с использованием CRF (Conditional Random Fields) и адаптеров. На основе архитектур выполнено обучение и оценка модели распознавания именованных сущностей на наборе данных RuLegalNER. Предложенный набор данных RuLegalNER может быть использован для создания новых моделей распознавания

© Shaheen Z., Mouromtsev D.I., Postny I., 2023

именованных сущностей в юридических текстах, что позволит автоматизировать контент-анализ юридических документов. Опубликована версия RuLegalNER с подробной статистикой и демонстрацией полезности набора данных RuLegalNER путем оценки на основе современных архитектур.

#### Ключевые слова

распознавание именованных юридических сущностей, обработка естественного языка, извлечение информации, языки с ограниченными ресурсами, передаточное обучение, трансформеры

**Ссылка для цитирования:** Шахин З., Муромцев Д.И., Постный И. RuLegalNER: новый датасет для распознавания именованных юридических сущностей на русском языке // Научно-технический вестник информационных технологий, механики и оптики. 2023. Т. 23, № 4. С. 854–857 (на англ. языке). doi: 10.17586/2226-1494-2023-23-4-854-857

Recognizing named entities in legal texts is a crucial task in natural language processing, with applications ranging from information extraction [1] to legal research [2]. However, the availability of resources specifically tailored for legal Named Entity Recognition (NER) in the Russian language is limited. This scarcity poses a significant challenge for researchers and practitioners working in the legal domain. Additionally, annotating legal datasets with expert human annotators is an expensive process. Furthermore, even with expert annotations, noisy labels can be present due to the complexity and ambiguity of legal texts [3].

In the Russian language, the availability of datasets specifically focused on legal named entity recognition is scarce. Among the limited options [4–6], one notable dataset is NEREL [7] which includes a total of 20 classes, with only three classes specifically related to the legal domain: LAW, CRIME, and PENALTY. However, out of the 56,000 annotated named entity instances in NEREL, only 1,679 pertain to legal named entities. This highlights the need for more comprehensive and domain-specific resources in the legal domain for Russian.

To address these challenges, this paper introduces RuLegalNER, a rule-based annotated legal dataset for the Russian language. The dataset was created using a rule-based program developed by legal experts in Tag-Consulting Company, enabling the automatic annotation of legal named entities in a large collection of Russian legal texts. This rule-based approach alleviates the need for extensive manual annotations by experts, speeding up the dataset creation process. Although rule-based annotation may introduce some noise, it serves as a starting point for subsequent refinement and iterative improvement.

To study the generalization ability of a named entity recognition model trained on an automatically annotated dataset, we developed RuLegalNER, a dataset of Russian legal documents. This dataset is annotated with more than 20 classes of named entities. However, it is important to note that not all legal named entities present in the documents are annotated, and the annotation coverage is sparse. From the initial set of classes, we specifically selected five classes for inclusion in this dataset: Individual person, legal entity, Penalty, Crime, and Law. The annotation process was performed using a rule-based system provided by TAG Consulting Company.

To ensure the evaluation of the model performance on unseen named entities, we incorporated low frequency entities into the dataset. These entities were treated as unseen during the training process, and they were exclusively reserved for the validation and test stages.

The RuLegalNER dataset consists of a sample of 100,000 Russian legal documents. Within this dataset, there are a total of 860 unique named entities. Notably, 289 of these entities appear only in the test set, resulting in a total of 777 occurrences of unseen entities in the test set.

For detailed statistics on the distribution of named entities in each split of the dataset, please refer to Table 1. The table shows the number of unique entities and their frequencies within each portion of the dataset as well as statistics for both seen and unseen entities in the test set. Additionally, Figure provides samples from the dataset, showcasing the variety of named entities present. The dataset is publicly available<sup>1</sup>.

We evaluated our models ability to predict unseen named entities, handle misspellings, and different grammatical cases. Objective metrics, such as precision, recall, and f1 score, were employed to assess their prediction power. We employed additional two objective evaluation metrics: Count of Predicted Unseen Named Entities (CP-UNE) and Count of Unique Predicted Unseen Named Entities (CUP-UNE), to measure the models generalization ability to unseen named entities.

In our experiments, we utilized various NER models for our research. The first model, RuBERT-NER, is based on RuBERT [8], a Russian text feature extraction model trained using the Russian version of Wikipedia and multilingual-BERT as the base checkpoint. We fine-tuned RuBERT using legal documents and employed a token classifier to generate probabilities for different classes. An extension of RuBERT-NER, RuBERT-NER-CRF, incorporated Conditional Random Fields (CRFs) to capture long-range dependencies and improve prediction accuracy. It utilized the Viterbi algorithm and a learnable state-state transition matrix for decoding the output labels. Another extension, RuBERT-NER-Adapter, employed adapters [9], a transfer learning strategy, to augment RuBERT-NER without significantly increasing the number of parameters. Lastly, we used a baseline model, BiLSTM-CRF [10], which combined a bidirectional LSTM with a CRF component for NER tasks. This model was pretrained on Russian Wikipedia and fine-tuned using legal documents.

Given the sparse nature of the annotated dataset and to evaluate the performance of these models and make comparisons, we divided the legal documents into segments, each containing 60 words, and filtered out segments that did not contain any annotated legal entities. The remaining segments were utilized for training,

<sup>1</sup> Available at: <https://github.com/zeino8/RuLegalNER> (accessed: 18.07.2023).

-	Исследовав представленные доказательства, <b>мировой судья</b> приходит к следующему.
-	Кроме того, <b>истец</b> назначен в качестве управляющей организации многоквартирными жилыми
-	Р Е Ш И Л : Взыскать с Шмыгалевой <ФИО1> пользу <b>ООО</b> «ГУК «Центр» <b>задолженность</b> за жилое помещение и коммунальные...
-	нарушение установленных законодательством о налогах и сборах сроков представления налоговой декларации в <b>налоговый орган</b> по месту учета.
-	<b>Ответчик</b> в <b>судебное заседание</b> не явился, о времени и месте ...
-	... то есть совершил <b>правонарушение</b> , <b>ответственность</b> за которое предусмотрена ч. 4 ст. 12.15 КоАП РФ.

Classes and Colors: **Individual Person** - **Legal Entity** - **Penalty** - **Law** - **Crime**

Figure. Legal texts samples from RuLegalNER with classes and their colors, and entities highlighted with the corresponding class color

Table 1. Named entity data in the RuLegalNER dataset organized by class number

Class	# unique entities					# occurrences				
	Train	Validation	Test		Dataset	Train	Validation	Test		Dataset
			Seen	Unseen				Seen	Unseen	
Individual	320	521	293	192	640	214,591	44,553	32,893	258	291,295
Legal Entity	30	55	30	26	59	28,343	6922	4599	165	40,029
Penalty	35	66	35	31	70	70,224	15,666	10,981	169	97,040
Crime	21	38	21	17	42	61,158	12,766	8897	91	82,912
Law	25	45	25	23	49	91,771	17,727	12,483	94	122,075
Total	431	725	404	289	860	466,087	97,634	68,853	777	633,351

Table 2. Comparing objective evaluation metrics (precision, recall, f1) for various models: (1) RuBERT-NER, (2) RuBERT-NER-CRF, (3) RuBERT-NER-Adapter, and (4) BiLSTM-CRF on all classes together. We also count number of previously unseen named entities (CP-UNE) and number of unique previously unseen named entities (CUP-UNE)

Model	Precision	Recall	F1-score	CP-UNE	CUP-UNE
RuBERT-NER	0.951	0.960	0.956	58	46
RuBERT-NER-CRF	0.976	0.847	0.907	57	48
RuBERT-NER-Adapter	0.905	0.940	0.922	52	40
BiLSTM-CRF	0.937	0.939	0.938	7	3

validation, and testing purposes. Evaluation results and comparison between models is presented in Table 2.

We introduced RuLegalNER dataset, a legal named entity recognition dataset in the Russian language. With its automatically annotated legal documents, RuLegalNER provides a valuable resource for training and evaluating

NER systems in the legal domain. The evaluation of various modern architectures for NER models on RuLegalNER highlights the strengths and limitations of each approach, enabling researchers to make informed decisions regarding model selection.

## References

- Weston L., Tshitoyan V., Dagdelen J., Kononova O., Trewartha A., Persson K.A., Ceder G., Jain A.. Named entity recognition and normalization applied to large-scale information extraction from the materials science literature. *Journal of Chemical Information and Modeling*, 2019, vol. 59, no. 9, pp. 3692–3702. <https://doi.org/10.1021/acs.jcim.9b00470>
- Angelidis I., Chalkidis I., Koubarakis M. Named entity recognition, linking and generation for greek legislation. *Legal Knowledge and Information Systems*, 2018, vol. 313, pp. 1–10.
- Zhu Y., Ye Y., Li M., Zhang J., Wu O. Investigating annotation noise for named entity recognition. *Neural Computing and Applications*, 2023, vol. 35, no. 1, pp. 993–1007. <https://doi.org/10.1007/s00521-022-07733-0>
- Vlasova N.A., Suleymanova E.A., Trofimov I.V. Report on Russian corpus for personal name retrieval. *Proceedings of Computational and Cognitive Linguistics, TEL*, 2014, pp. 36–40.

## Литература

- Weston L., Tshitoyan V., Dagdelen J., Kononova O., Trewartha A., Persson K.A., Ceder G., Jain A.. Named entity recognition and normalization applied to large-scale information extraction from the materials science literature // *Journal of Chemical Information and Modeling*. 2019. V. 59. N 9. P. 3692–3702. <https://doi.org/10.1021/acs.jcim.9b00470>
- Angelidis I., Chalkidis I., Koubarakis M. Named entity recognition, linking and generation for greek legislation // *Legal Knowledge and Information Systems*. 2018. V. 313. P. 1–10.
- Zhu Y., Ye Y., Li M., Zhang J., Wu O. Investigating annotation noise for named entity recognition // *Neural Computing and Applications*. 2023. V. 35. N 1. P. 993–1007. <https://doi.org/10.1007/s00521-022-07733-0>
- Vlasova N.A., Suleymanova E.A., Trofimov I.V. Report on Russian corpus for personal name retrieval // *Proceedings of Computational and Cognitive Linguistics, TEL*. 2014. P. 36–40.

5. Starostin A.S., Bocharov V.V., Alexeeva S.V., Bodrova A.A., Chuchunkov A.S., Dzhumaev S.S., Efimenko I.V., Granovsky D.V., Khoroshevsky V.F., Krylova I.V., Nikolaeva M.A., Smurov I.M., Toldova S.Y. Factrueval 2016: evaluation of named entity recognition and fact extraction systems for Russian. *Proc. of the International Conference "Dialogue 2016"*, 2016, pp. 702–720.
6. Gareev R., Tkachenko M., Solovyev V., Simanovsky A., Ivanov V. Introducing baselines for russian named entity recognition. *Lecture Notes in Computer Science*, 2013, vol. 7816, pp. 329–342. [https://doi.org/10.1007/978-3-642-37247-6\\_27](https://doi.org/10.1007/978-3-642-37247-6_27)
7. Loukachevitch N., Artemova E., Batura T., Braslavski P., Denisov I., Ivanov V., Manandhar S., Pugachev A., Tutubalina E. Nerel: A Russian dataset with nested named entities, relations and events. *Proc. of the Recent Advances in Natural Language Processing*, 2021, pp. 876–885 [https://doi.org/10.26615/978-954-452-072-4\\_100](https://doi.org/10.26615/978-954-452-072-4_100)
8. Kuratov Y., Arkhipov M. Adaptation of deep bidirectional multilingual transformers for Russian language. *Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialogue 2019"*, 2019.
9. Houslyby N., Giurgiu A., Jastrzebski S., Morrone B., De Laroussilhe Q., Gesmundo A., Attariyan M., Gelly S. Parameter-efficient transfer learning for NLP. *Proc. of the 36<sup>th</sup> International Conference on Machine Learning*, 2019, pp. 2790–2799.
10. Panchendrarajan R., Amasesan A. Bidirectional LSTM-CRF for named entity recognition. *Proc. of the 32<sup>nd</sup> Pacific Asia Conference on Language, Information and Computation*, 2018, pp. 531–540.

#### Authors

**Zein Shaheen** — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57209279132](https://orcid.org/0000-0001-6802-2896), <https://orcid.org/0000-0001-6802-2896>, [shaheen@itmo.ru](mailto:shaheen@itmo.ru)

**Dmitry I. Mouromtsev** — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 55575780100](https://orcid.org/0000-0002-0644-9242), <https://orcid.org/0000-0002-0644-9242>, [mouromtsev@itmo.ru](mailto:mouromtsev@itmo.ru)

**Ignat Postny** — Director, T.A.G. Consulting, Moscow, 119119, Russian Federation, <https://orcid.org/0009-0005-9249-4160>, [ipostny@gmail.com](mailto:ipostny@gmail.com)

Received 19.06.2023

Approved after reviewing 23.06.2023

Accepted 30.07.2023

#### Авторы

**Шахин Зейн** — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57209279132](https://orcid.org/0000-0001-6802-2896), <https://orcid.org/0000-0001-6802-2896>, [shaheen@itmo.ru](mailto:shaheen@itmo.ru)

**Муромцев Дмитрий Ильич** — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 55575780100](https://orcid.org/0000-0002-0644-9242), <https://orcid.org/0000-0002-0644-9242>, [mouromtsev@itmo.ru](mailto:mouromtsev@itmo.ru)

**Постный Игнат** — директор, Т.А.Г. Консалтинг, Москва, 119119, Российская Федерация, <https://orcid.org/0009-0005-9249-4160>, [ipostny@gmail.com](mailto:ipostny@gmail.com)

Статья поступила в редакцию 19.06.2023

Одобрена после рецензирования 23.06.2023

Принята к печати 30.07.2023



Работа доступна по лицензии  
Creative Commons  
«Attribution-NonCommercial»

**Опечатки и исправления в журнале**  
**Научно-технический вестник информационных технологий, науки и техники**  
**№ 3 2023**

Страница, строка	Напечатано	Следует читать
443, формула	$L_{\text{эфф}} = \begin{cases} L_{\text{ког}}, L_{\text{крис}} > L_{\text{ког}} \\ L_{\text{крис}}, L_{\text{крис}} > L_{\text{ког}} \end{cases}$	$L_{\text{эфф}} = \begin{cases} L_{\text{ког}}, L_{\text{крис}} > L_{\text{ког}} \\ L_{\text{крис}}, L_{\text{крис}} \leq L_{\text{ког}} \end{cases}$

## ТРЕБОВАНИЯ К ОФОРМЛЕНИЮ СТАТЕЙ

Журнал публикует научные обзоры, полнотекстовые статьи и краткие сообщения.

Полнотекстовая статья должна иметь четкую структуру, включающую в себя аннотацию, ключевые слова, введение, несколько содержательных разделов и заключение.

В аннотации, рассчитанной на самый широкий круг читателей, необходимо в объеме 200–400 слов в форме краткого реферата (Abstract) изложить научное содержание статьи: предмет, цель работы, метод или методология проведения работы, краткое описание эксперимента, полученные результаты, рекомендации по их применению.

Ключевые слова должны отражать информацию, представленную в статье. Основное ключевое слово указывается первым.

Во введении необходимо представить содержательную постановку рассматриваемого вопроса, провести краткий анализ известных из научной литературы решений (со ссылками на источники), дать критику их недостатков, показать научную новизну и преимущество (особенности) предлагаемого подхода.

В основном тексте статьи должна быть представлена строгая постановка решаемой задачи, изложены и обстоятельно разъяснены (доказаны) полученные утверждения и выводы, приведены результаты экспериментальных исследований или математического моделирования, иллюстрирующие сделанные утверждения. Основной текст статьи должен быть разбит на содержательные разделы (2–3 раздела).

В заключении необходимо кратко сформулировать основные результаты, прокомментировать их и, если возможно, указать на направления дальнейших исследований и области применения.

Пристатейный список литературы (рекомендуется): для обзорной статьи — не менее 50, для полнотекстовой статьи — не менее 15, для краткого сообщения — не менее 8 литературных источников.

Ссылки на стандарты и иные нормативные документы, а также на неперсонифицированные интернет-ресурсы в список литературы не вносятся, а оформляются в виде сносок.

Объем обзорной статьи предварительно согласовывается с редакцией.

Объем полнотекстовой статьи, включая иллюстрации, таблицы и список литературы, не должен превышать 8 страниц машинописного текста (с рисунками и таблицами), шрифт 12 pt, один интервал.

Объем краткого сообщения — до 3 страниц, шрифт 12 pt, один интервал. Рубрикация текста не требуется. Статьи принимаются в электронном виде (ntvitmo@itmo.ru). В распечатанном виде с подписями авторов материалы представляются в редакцию по запросу редакции после прохождения процедуры рецензирования.

Комплект документов должен включать:

- текст статьи с заверстанными рисунками и таблицами;
- формы сведений об авторах (на каждого автора); допускается указание нескольких мест работы автора;
- файлы с рисунками к статье в оригинальном формате (предпочтительно JPEG) с максимальным разрешением; допускается представление цветных рисунков, если в черно-белом варианте теряется полезная информация;
- лицензионное соглашение;
- согласие на обработку персональных данных.

## REQUIREMENTS FOR EXECUTION OF PAPERS

Scientific reviews, full-text and brief papers are published.

A full-text paper should have a well-defined structure, including an abstract, keywords, introduction, several substantive chapters and a conclusion.

An abstract is intended for general public. It is necessary to set forth the scientific content of the paper limited by 200–400 words in the form of a brief abstract: the subject, the work objective, the method or methodology of work, a brief description of the experiment, obtained results, recommendations for their application.

Keywords should be connected with the information presented in the paper. The main keyword is given first.

The introduction should contain a meaningful statement of the issue in question, a brief analysis of the solutions known from scientific literature (with references to sources), a criticism of their shortcomings, scientific novelty and advantages (features) of the proposed approach.

The main body of the paper should represent exact statement of the problem being solved, obtained assertions and conclusions are to be set forth and explained in detail (proved), the results of experimental studies or mathematical modeling should be given to illustrate the statements made. The main text of the paper needs to be split up into meaningful sections (2–3 sections).

In conclusion it is necessary to summarize the main results, comment on them and, if possible, to indicate the areas of future research and application.

A list of references is recommended: for a review paper — at least 50 items, for full-text paper — not less than 15, for a brief paper — at least 8 references. References to standards and other regulatory documents, as well as to Internet resources, are not included in a reference list and are documented as footnotes.

The length of a review paper is coordinated with the editorial board beforehand.

The length of a full-text paper, including illustrations, tables and references, should not exceed 8 pages of typewritten text (figures and tables included), font 12 pt, single-spaced.

The length of a brief paper is up to 3 pages, font 12 pt, single-spaced.

No rubrication is required.

Papers are accepted in electronic form (ntvitmo@itmo.ru).

Materials are submitted to the editors on request after the procedure of reviewing in a printed form signed by the authors.

A set of documents should include:

- The text of the paper with make-up figures and tables;
- The forms with information about the authors (for each author), several employers can be specified by the author;
- Files with pictures to the paper in their original format (preferably JPEG) with the maximum resolution; should useful information is lost in the black and white variant, color drawings are permitted;
- Publishers license agreement;
- Consent to the processing of personal data.

**OPTICAL ENGINEERING**

Determination of the action type of hydrate formation inhibitors by their infrared spectra

*Vorozhtsova Iu.S., Nosenko T.N., Uspenskaya M.V.* 669

Application of Raman spectroscopy to study the inactivation process of bacterial microorganisms

*Matveeva K.I., Kundalevich A.A., Kapitonova A.I., Zozulya A.S., Sukhikh S.A., Tsibulnikova A.V., Zyubin A.Yu., Samusev I.G.* 676

Numerical study of the effect of methemoglobin concentration in the blood on the absorption of light by human skin

*Belikov A.V., Chuchin V.Yu.* 685

Low-temperature cell for IR Fourier spectrometric investigation of hydrocarbon substances (in English)

*Kenbay A.A., Golikov O.Yu., Aldiyarov A.U., Yerezhep D.E.* 696

**NEW MATERIALS AND NANOTECHNOLOGIES**

Peculiarities of growing  $\text{Ga}_{1-x}\text{In}_x\text{As}$  solid solutions on GaAs substrates in the field of a temperature gradient through a thin gas zone

*Devitsky O.V., Lunin L.S., Mitrofanov D.V., Sysoev I.A., Nikulin D.A., Chapura O.M.* 703

**COMPUTER SCIENCE**

An enhanced AES-GCM based security protocol for securing the IoT communication (in English)

*Feroz Khan A.B., Kalpana Devi S., Rama Devi K.* 711

Attacks based on malicious perturbations on image processing systems and defense methods against them

*Esipov D.A., Buchaev A.Y., Kerimbay A., Puzikova Y.V., Saidumarov S.K., Sulimenko N.S., Popov I.Yu., Karmanovskiy N.S.* 720

Brain MRT image super resolution using discrete cosine transform and convolutional neural network (in English)

*Singh P., Ganotra D.* 734

Text augmentation preserving persona speech style and vocabulary

*Matveeva A.A., Makhnytkina O.V.* 743

Verification of event-driven software systems using the specification language of cooperating automata objects (in English)

*Afanasieva I.V., Novikov F.A., Fedorchenko L.N.* 750

Intelligent adaptive testing system

*Tagirova L.F., Zubkova T.M.* 757

**ARTIFICIAL INTELLIGENCE AND COGNITIVE INFORMATION TECHNOLOGIES**

Neural network-based method for visual recognition of driver's voice commands using attention mechanism

*Axyonov A.A., Ryumina E.V., Ryumin D.A., Ivanko D.V., Karpov A.A.* 767

Brain tumour segmentation in MRI using fuzzy deformable fusion model with Dolphin-SCA (in English)

*Tiple A.H., Kakade A.B.* 776

Optimization of human tracking systems in virtual reality based on a neural network approach

*Obukhov A.D., Teselkin D.V.* 786

**MODELING AND SIMULATION**

Errors in the demodulation algorithm with a generated carrier phase introduced by the low-pass filter

*Miroshnichenko G.P., Arzhanenkova A.N., Plotnikov M.Yu.* 795

Modeling of the process of spherical form correction for rotors of electrostatically suspended gyros

*Tit M.A., Belyaev S.N., Shcherbak A.G., Yulmetova O.S.* 803

Method of spatial multiplexing in multi-antenna communication systems

*Grishentsev A.Yu.* 812

Modeling and simulation of heat exchanger with strong dependence of oil viscosity on temperature

*Kurmanova D.E., Jaichibekov N.Zh.* 820

Approach to the generalized parameters formation of the complex technical systems technical condition using neural network structures

*Maltsev G.N., Yakimov V.L.* 828

Numerical simulation of gas dynamics during operation of a wide-range rocket nozzle with a porous insert

*Mikhailov M.V., Prodan N.V., Renev M.E.* 836

The exact solution of a shock wave reflection problem from a wall shielded by a gas suspension layer

*Sadin D.V., Shirokova E.N.* 843

**BRIEF PAPERS**

Adaptive observer for state variables of a time-varying nonlinear system with unknown constant parameters and delayed measurements

*Bobtsov A.A., Nikolaev N.A., Kozachek O.A., Oskina O.V.* 850

RuLegalNER: a new dataset for Russian legal named entities recognition (in English)

*Shaheen Z., Mouromtsev D.I., Postny I.* 854