

ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ

2022, ТОМ 22, НОМЕР 6 (НОЯБРЬ–ДЕКАБРЬ)

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

ИЗБРАННЫЕ МАТЕРИАЛЫ XXXII ШКОЛЫ ПО ГОЛОГРАФИИ

Часть II

Импульсная запись динамических голограмм в кристалле силиката висмута при изменении длины волны лазерного излучения

Даденков И.Г., Толстик А.Л., Миксюк Ю.И., Саечников К.А. 1025

Гибридный эндоскоп с телевизионной и многоспектральной обработкой изображений для диагностики рака внутренних органов

Зайченко К.В., Гуревич Б.С., Беляев А.В., Святкина В.И. 1031

Моделирование композитного волноводного голографического дисплея

Харитонов Д.Ю., Ахметов Д.М., Муслимов Э.Р., Гильфанов А.Р., Павлычева Н.К. 1037

ОПТИЧЕСКИЕ СИСТЕМЫ И ТЕХНОЛОГИИ

Применение методов инфракрасной спектроскопии в исследовании составов для проклейки бумаги

Бабаханова Х.А., Садриддинов А.А., Абдуназаров М.М., Бабаханова М.А., Громыко И.Г. 1048

Методика оптимизации распределения плотности пикселей по зоне наблюдения

Волхонский В.В., Ковалевский В.А. 1055

АВТОМАТИЧЕСКОЕ УПРАВЛЕНИЕ И РОБОТОТЕХНИКА

Оценка ошибки и разработка методики компенсации погрешности позиционирования оборудования с числовым программным управлением (на англ. яз.)

Ризки М.А., Федосов Ю.В. 1063

Компенсация внешних возмущений по выходу для класса линейных систем с запаздыванием в канале управления

Буй В.Х., Маргун А.А. 1072

НОВЫЕ МАТЕРИАЛЫ И НАНОТЕХНОЛОГИИ

Люминесцентный метод исследования роста квантовых точек AgInS_2 (на англ. яз.)

Ахмад А.К., Мохаммед А.Х., Скапцов А.А. 1078

Особенности импульсного лазерного напыления тонких пленок InGaAsN в атмосфере активного фоновго газа

Девяцкий О.В. 1085

Нахождение распределения электронов в сверхрешетках AlGaAs/GaAs с узкими барьерами методом вольт-фарадного профилирования

Василькова Е.И., Пирогов Е.В., Соболев М.С., Баранов А.И., Гудовских А.С., Буравлев А.Д. 1092

Спектральные и кинетические свойства квантовых точек сульфида серебра во внешнем электрическом поле

Дайбаге Д.С., Амброзович С.А., Перепелица А.С., Захарчук И.А., Осадченко А.В., Безверхняя Д.М., Авраменко А.И., Селюков А.С. 1098

Влияние наноразмерных горизонтальных неоднородностей на послойный анализ поверхности методом рентгеновской фотоэлектронной спектроскопии

Лукьянцев Д.С., Лубенченко А.В., Иванов Д.А., Лубенченко О.И., Федотов А.С. 1104

Органические светоизлучающие диоды с новыми красителями на основе кумарина

Осадченко А.В., Ващенко А.А., Захарчук И.А., Дайбаге Д.С., Амброзович С.А., Володин Н.Ю., Чепцов Д.А., Долотов С.М., Травень В.Ф., Авраменко А.И., Семенова С.Л., Селюков А.С. 1112

Изготовление и характеристика гибридного композита Al6082/SiC /порошок рисовой шелухи, получаемого методом фрикционного перемешивания (на англ. яз.)

Кумар Н., Гупта П., Сингх Р.К. 1119

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

Многопутевая безопасная маршрутизация для обнаружения атаки с захватом узла в беспроводной сенсорной сети (на англ. яз.)

Колангаппан Д., Сентил Кумар А. 1127

Метод документирования архитектурных решений вычислительных платформ

Горбачев Я.Г. 1136

Повышение точности распознавания внесловарных слов для интегральной системы автоматического распознавания русской речи (на англ. яз.)

Андрусенко А.Ю., Романенко А.Н. 1143

Метод мониторинга состояния элементов киберфизических систем на основе анализа временных рядов

Семенов В.В. 1150

Применение волновой модели текста к задаче сентимент-анализа

Груздева А.С., Юрьев Р.Н., Бессмертный И.А. 1159

Автоматизированная оценка параметров электрокардиограмм в условиях пандемии COVID-19

Ватьян А.С., Гусарова Н.Ф., Добренко Н.В., Змиевский Д.А., Кабышев М.В., Полевая Т.А., Татаринова А.А., Томилов И.В. 1166

Мультиагентная адаптивная маршрутизация агентами-клонами на основе многоголового внутреннего внимания с использованием обучения с подкреплением

Грибанов Т.А., Фильченков А.А., Азаров А.А., Шалыто А.А. 1178

Совместное обучение агентов и векторных представлений графов в задаче управления конвейерными лентами

Рыбкин К.Е., Фильченков А.А., Азаров А.А., Забашта А.С., Шалыто А.А. 1187

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ

Моделирование процессов переноса излучения в газожидкостных пенах

Исаева А.А., Исаева Е.А., Пантюков А.В. 1197

Статистическая оценка влияния величины сигнал/помеха на погрешность измерения параметров акустической эмиссии

Федоров А.В., Алтай Е., Степанова К.А., Кузичанов Д.О. 1205

Моделирование процесса стационарного термоотражения для измерения теплопроводности материалов

Тукмакова А.С., Демченко П.С., Тхоржевский И.Л., Новотельнова А.В., Ходзицкий М.К. 1216

Математическое и компьютерное моделирования однорядных и двухрядных шестипластных винтокольцевых движителей

Дудников С.Ю., Булат М.П., Вокин Л.О., Кузнецов П.Н., Чернышов П.С. 1226

КРАТКИЕ СООБЩЕНИЯ

Дифференциально-разностная модель теплопереноса в твердых телах с использованием метода параметрической идентификации

Пилипенко Н.В., Колодийчук П.А., Заричняк Ю.П. 1237

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. — к.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)
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, IBIS 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

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

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 — PhD, Associate Professor (Russia)
Denis V. Efimov — PhD (France)
Danil 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)
Nikolay V. Nikonorov — D.Sc, Professor (Russia)
Grigori 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

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



Подписано к печати 01.12.2022 Тираж 350 экз. Заказ № 90 (142)

Издание Университета ИТМО

Отпечатано в ООО «Университетские телекоммуникации»
Адрес: 199034, Санкт-Петербург, Биржевая линия, д. 16

© Университет ИТМО

ИЗБРАННЫЕ МАТЕРИАЛЫ XXXII ШКОЛЫ ПО ГОЛОГРАФИИ. ЧАСТЬ II PROCEEDINGS OF THE XXXII SCHOOL ON HOLOGRAPHY. PART II

doi: 10.17586/2226-1494-2022-22-6-1025-1030

УДК 535.343

Импульсная запись динамических голограмм в кристалле силиката висмута при изменении длины волны лазерного излучения

Иван Геннадьевич Даденков¹✉, Алексей Леонидович Толстик²,
Юрий Иванович Миксюк³, Константин Алексеевич Саечников⁴

^{1,2} Белорусский государственный университет, Минск, 220030, Республика Беларусь

^{3,4} Белорусский государственный педагогический университет имени Максима Танка, Минск, 220050, Республика Беларусь

¹ dadenkov.ivan@gmail.com✉, <https://orcid.org/0000-0003-0722-4683>

² tolstik@bsu.by, <https://orcid.org/0000-0003-4953-4890>

³ miksiuk@mail.ru, <https://orcid.org/0000-0002-3969-0454>

⁴ ska-bspu@tut.by, <https://orcid.org/0000-0001-8591-9312>

Аннотация

Предмет исследования. Исследован процесс формирования динамических голограмм в фоторефрактивном кристалле силиката висмута наносекундными лазерными импульсами при изменении длины волны излучения. Предложена оригинальная схема записи голограмм для сохранения периода решетки при работе на разных длинах волн. **Метод.** Применен метод импульсной записи динамических решеток на основе предложенной оптической схемы, который обеспечивает выполнение условия Брэгга для зондирующего излучения. Оригинальность схемы заключается в использовании первых порядков дифракции пропускающей дифракционной решетки в качестве опорной и сигнальной волн, а также телескопической системы, примененной для фиксации периода записываемой решетки независимо от используемой длины волны. **Основные результаты.** Получены кинетические зависимости дифракционной эффективности динамических голограмм в фоторефрактивном кристалле силиката висмута при изменении длины волны записываемого излучения в актуальном спектральном диапазоне (от 450 до 600 нм). Установлен эффект проявления конкурирующих механизмов записи коротко- (сотни микросекунд) и долгоживущих (секунды) решеток, вклад которых зависит от длины волны записываемого голограмму излучения. Определена оптимальная длина волны для получения наибольшей дифракционной эффективности голограмм. Показано, что излучение в сине-зеленой области спектра приводит к преимущественной записи короткоживущей решетки, в то время как в красной области спектра доминируют долгоживущие решетки. **Практическая значимость.** Необходимость исследований фоторефрактивных кристаллов семейства силленитов определена их применением для мультиплексной записи динамических голограмм и реализации метода адаптивной интерферометрии для отслеживания изменений объектов в реальном времени.

Ключевые слова

динамические голограммы, фоторефрактивные кристаллы, силикат висмута, импульсная голография, дифракционная эффективность, перестройка длины волны

Ссылка для цитирования: Даденков И.Г., Толстик А.Л., Миксюк Ю.И., Саечников К.А. Импульсная запись динамических голограмм в кристалле силиката висмута при изменении длины волны лазерного излучения // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1025–1030. doi: 10.17586/2226-1494-2022-22-6-1025-1030

Pulse recording of dynamic holograms in bismuth silicate crystal in a broad wavelength range

Ivan G. Dadenkov^{1✉}, Alexei L. Tolstik², Yuri I. Miksiuk³, Konstantin A. Saechnikov⁴

^{1,2} Belarussian State University, Minsk, 220030, Republic of Belarus

^{3,4} Belarussian State Pedagogical University named after Maxim Tank, Minsk, 220050, Republic of Belarus

¹ dadenkov.ivan@gmail.com✉, <https://orcid.org/0000-0003-0722-4683>

² tolstik@bsu.by, <https://orcid.org/0000-0003-4953-4890>

³ miksiuk@mail.ru, <https://orcid.org/0000-0002-3969-0454>

⁴ ska-bspu@tut.by, <https://orcid.org/0000-0001-8591-9312>

Abstract

The formation of dynamic holograms in a photorefractive bismuth silicate crystal by nanosecond laser pulses with a change in the radiation wavelength is studied. An original scheme for recording holograms is proposed to preserve the grating period when operating at different wavelengths. The method of pulsed recording of dynamic gratings based on the proposed optical scheme is applied, which ensures the fulfillment of the Bragg condition for probing radiation. The originality of the scheme lies in the use of the first orders of diffraction of a transmitting diffraction grating as the reference and signal waves, as well as a telescopic system used to fix the period of the recorded grating, regardless of the wavelength used. Kinetic dependences of the diffraction efficiency of dynamic holograms in a photorefractive bismuth silicate crystal are obtained with a change in the wavelength of recording radiation in the actual spectral range (from 450 nm to 600 nm). The effect of the manifestation of competing recording mechanisms of short-lived (hundreds of microseconds) and long-lived (seconds) gratings, the contribution of which depends on the wavelength of the radiation recording the hologram, has been established. The optimal wavelength for obtaining the highest diffraction efficiency of holograms is determined. It is shown that radiation in the blue-green region of the spectrum leads to the predominant recording of a short-lived grating, while long-lived gratings dominate in the red region of the spectrum. The need to study photorefractive crystals of the sillenite family is determined by their use for multiplex recording of dynamic holograms and the implementation of the adaptive interferometry method to track changes in objects in real time.

Keywords

dynamic holograms, photorefractive crystals, bismuth silicate, pulse holography, diffraction efficiency, wavelength tuning

For citation: Dadenkov I.G., Tolstik A.L., Miksiuk Yu.I., Saechnikov K.A. Pulse recording of dynamic holograms in bismuth silicate crystal in a broad wavelength range. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1025–1030 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1025-1030

Введение

Одно из существенных преимуществ фоторефрактивных кристаллов — возможность записи и перезаписи в них динамических голограмм, время жизни которых может изменяться от микросекунд до секунд и даже часов. Это определяет их использование в голографических системах записи, хранения и обработки информации, адаптивных интерферометрах и др. [1–5]. Физические процессы записи голограмм в фоторефрактивных кристаллах основаны на пространственном перераспределении зарядов в поле интерферирующих световых пучков по многочисленным центрам, имеющим разнообразную природу возникновения и характеристики. По своей энергетической структуре фоторефрактивные кристаллы семейства силленитов, в том числе исследуемый в настоящей работе кристалл силиката висмута, являются широкозонными полупроводниками. При этом существенное влияние на формирование голограмм оказывают примеси и структурные дефекты кристаллической решетки, приводящие к возникновению в запрещенной зоне донорных и акцепторных энергетических уровней [6–9]. Особенность кристаллов силленитов — одновременное существование долго- (секунды, часы) и короткоживущих ловушечных уровней (микро- и миллисекунды), которые могут быть заселены в результате поглощения лазерного излучения кристаллом в широком спектральном диапазоне [10–12]. При этом формирование светоиндуцированного отклика, выражающегося в локальном изменении

показателя преломления и коэффициента поглощения, может быть реализовано с использованием как непрерывного лазерного излучения, так и под действием нано- и пикосекундных лазерных импульсов [13, 14]. В работах [15, 16] показано, что, проявление короткоживущих ловушечных уровней существенно зависит от интенсивности записывающего излучения. При небольших интенсивностях записываются долгоживущие решетки с временем жизни на уровне секунд, в то время как с увеличением интенсивности проявляются быстрые компоненты в диапазоне сотен микросекунд. Отметим, что исследования выполнены на одной длине волны 532 нм. Важный параметр для записи динамических голограмм — использование длины волны записывающего излучения, поскольку при использовании другой длины волны возможно проявление различных особенностей формирования и релаксации динамических решеток.

Цель данной работы — оптимизация спектральных условий импульсной записи голограмм в фоторефрактивном кристалле силиката висмута. Для достижения поставленной цели определены следующие задачи: измерить зависимости интенсивности дифрагированного сигнала от времени при использовании для записи импульсного лазерного излучения на разных длинах волн; определить время формирования и релаксации динамических решеток; установить оптимальные длины волн для записи коротко- и долгоживущих решеток. С целью оптимизации условий голографической записи выполнен анализ спектральных особенностей форми-

рования динамических голограмм при использовании лазерных импульсов на различных длинах волн в диапазоне от 450 до 600 нм.

Методика проведения эксперимента

Для исследования зависимости амплитудных и временных характеристик динамических голограмм в кристалле силиката висмута в качестве источника излучения, формирующего голограмму, использован параметрический генератор с длительностью лазерных импульсов 10 ± 1 нс и шириной линии генерации $\delta\lambda$ менее 0,1 нм в видимой области спектра. Параметрический генератор возбужден излучением третьей гармоники Nd:YAG лазера. Зондирование области записи голограммы проведено излучением непрерывного гелий-неонового лазера мощностью в несколько милливатт. Использование данного лазера в качестве зонда способствовало непрерывному контролю за процессами формирования и релаксации записываемых голограмм. Разработана оригинальная схема записи (рис. 1), основанная на использовании дифракционной решетки с известным периодом и телескопической системы, фокусирующей первые порядки дифракции на исследуемый кристалл. Отличительная особенность данной схемы: при изменении длины волны записывающего излучения изменяется угол между пучками, дифрагированными в плюс/минус первые порядки, что приводит к автоматической подстройке и сохранению периода формируемой голографической решетки независимо от длины волны падающего излучения. В этом случае имеет место автоматическое выполнение условия Брэгга для восстанавливающего светового пучка при изменении длины волны записывающего излучения. Наблюдение

за кинетикой релаксации записанной голограммы осуществлено при помощи *p-i-n* фотодиода и цифрового осциллографа. Параметрический генератор позволил получить импульсное лазерное излучение в диапазоне от 400 нм до 2 мкм с шагом изменения 0,1 нм, что дает возможность работать во всем спектральном диапазоне, актуальном для кристаллов семейства силленитов (450–600 нм). В данном интервале для кристалла силиката висмута имеет место эффект фотоиндуцированного поглощения, однако, при использовании для записи голограмм одиночных лазерных импульсов с фиксированием энергии на уровне 1 мДж, формируемая этим эффектом амплитудная решетка оказывает пренебрежимо малое влияние на общую дифракционную эффективность голограммы. Выбор коротковолновой границы диапазона 450 нм обусловлен высоким поглощением излучения в кристалле силиката висмута при приближении к длине волны, отвечающей прямым межзонным переходам. В этом случае замечено нарушение условия объемности формируемой дифракционной решетки, и голограмма записана не в объеме кристалла, а в приграничном слое. Такое нарушение связано с тем, что эффективная глубина проникновения излучения в кристалл силиката висмута, согласно закону Бугера, меньше фактической толщины кристалла, и это приводит к тому, что параметр Клейна Q меньше 10. При этом голограмма записана не во всем объеме кристалла, а в приповерхностном слое. Длинноволновая граница выбранного диапазона обусловлена снижением дифракционной эффективности и чувствительности кристаллов при записи голограмм в красной области спектра.

Результат проведенного эксперимента — получение набора осциллограмм, снятых при различных времен-

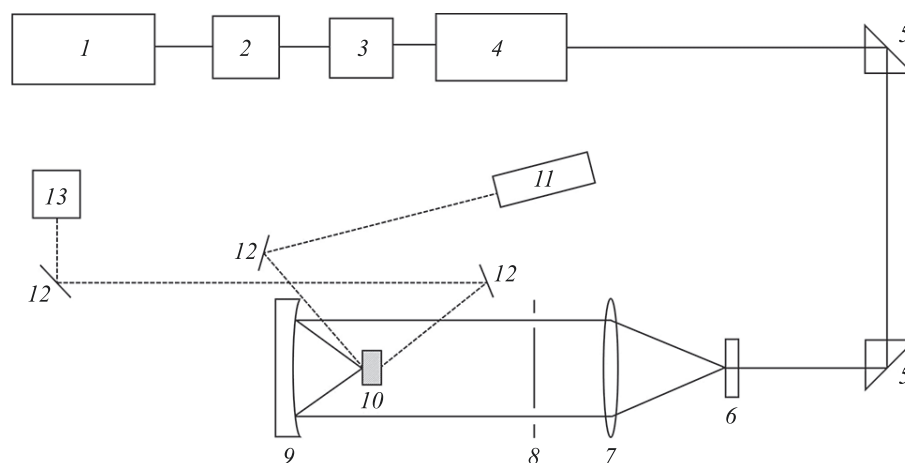


Рис. 1. Схема экспериментальной установки:

1 — импульсный лазер на иттрий-алюминиевом гранате с ионами неодима; 2 — генератор второй гармоники на кристалле KDP; 3 — генератор третьей гармоники на кристалле LBO; 4 — параметрический генератор на кристалле BBO; 5 — призма полного внутреннего отражения; 6 — дифракционная решетка; 7 — собирающая линза с $f = 350$ мм; 8 — диафрагма; 9 — сферическое зеркало с $f = 80$ мм; 10 — кристалл силиката висмута с линейными размерами $10 \times 10 \times 2$ мм; 11 — непрерывный гелий-неоновый лазер; 12 — зеркало; 13 — фотоприемное устройство

Fig. 1. Scheme of the experimental setup, where 1 — pulsed yttrium-aluminum garnet laser with neodymium ions; 2 — second harmonic generator on a KDP crystal; 3 — third harmonic generator on an LBO crystal; 4 — parametric generator on a BBO crystal; 5 — prism total internal reflection; 6 — diffraction grating; 7 — converging lens with $f = 350$ mm; 8 — diaphragm; 9 — spherical mirror with $f = 80$ mm; 10 — bismuth silicate crystal with linear dimensions $10 \times 10 \times 2$ mm; 11 — continuous helium-neon laser; 12 — mirror; 13 — photodetector

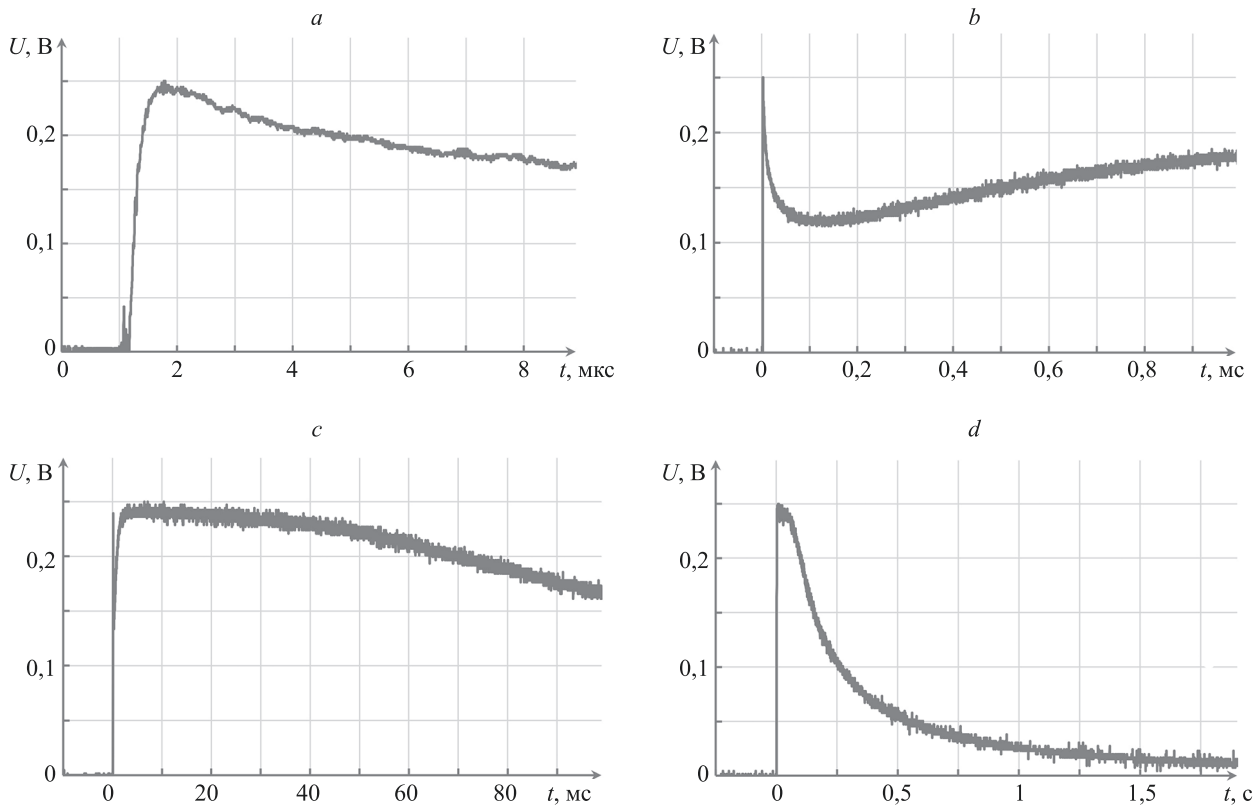


Рис. 2. Осциллограммы формирования и релаксации записанной голограммы, полученные путем наблюдения за дифрагированным лучом гелий-неонового лазера, с временными разрешениями: 250 нс (а), 50 мкс (б), 5 мс (с), 100 мс (д), при освещении кристалла лазерным импульсом параметрического генератора с длиной волны 510 нм и интенсивностью 1 МВт/см^2 , где U — напряжение, t — время

Fig. 2. Oscillograms of the formation and relaxation of the recorded hologram obtained by observing the diffracted beam of a helium-neon laser, with time resolutions: 250 ns (a), 50 ms (b), 5 ms (c), 100 ms (d), when the crystal is illuminated by a laser pulse of a parametric generator with a wavelength of 510 nm and an intensity of 1 MW/cm^2 , where U is the voltage, t is the time

ных разрешений (рис. 2). Изменение разрешений необходимо, так как при импульсной записи формируются коротко- и долгоживущие голограммы с временами жизни, отличающимися на несколько порядков.

Результаты и обсуждение

Дифракционные эффективности (рис. 3, а, б) нормированы на максимальное значение, полученное в ходе экспериментов. Видно, что дифракционная эффективность (η) долгоживущей решетки практически не зависит от длины волны излучения (λ), формирующей голограмму, в то время как короткоживущая решетка имеет максимум на длине волны 470 нм. Это связано с тем, что при толщине исследуемого кристалла, равной 2 мм, на этой длине волны происходит наиболее эффективное поглощение падающего излучения в объеме кристалла. При использовании излучения с меньшей длиной волны голограмма будет записываться в слое с меньшей толщиной, чем исследуемый кристалл, а при использовании излучения с большей длиной волны — заметная часть излучения будет проходить через кристалл, не испытывая поглощение.

На рис. 3, с, д представлены зависимости времени формирования коротко- и долгоживущей решеток от длины волны записывающего излучения. Видно, что время записи короткоживущей решетки находится на уровне нескольких сотен наносекунд, в то время как для долгоживущей решетки это время составляет несколько миллисекунд. Из представленных зависимостей также можно сделать вывод, что в коротковолновой области спектра время формирования короткоживущей решетки максимальное, в то время как долгоживущей — минимальное. При продвижении в длинноволновую область спектра время формирования стабилизируется примерно на одном уровне для обеих решеток.

Различия в поведении динамических решеток дают основания предположить разные физические механизмы их формирования: короткоживущая решетка обусловлена локальным перераспределением электронов по ловушечным уровням, в то время как долгоживущая решетка связана с диффузией электронов по объему кристалла и формированием фоторефрактивного отклика за счет генерации электрического поля внутри кристалла.

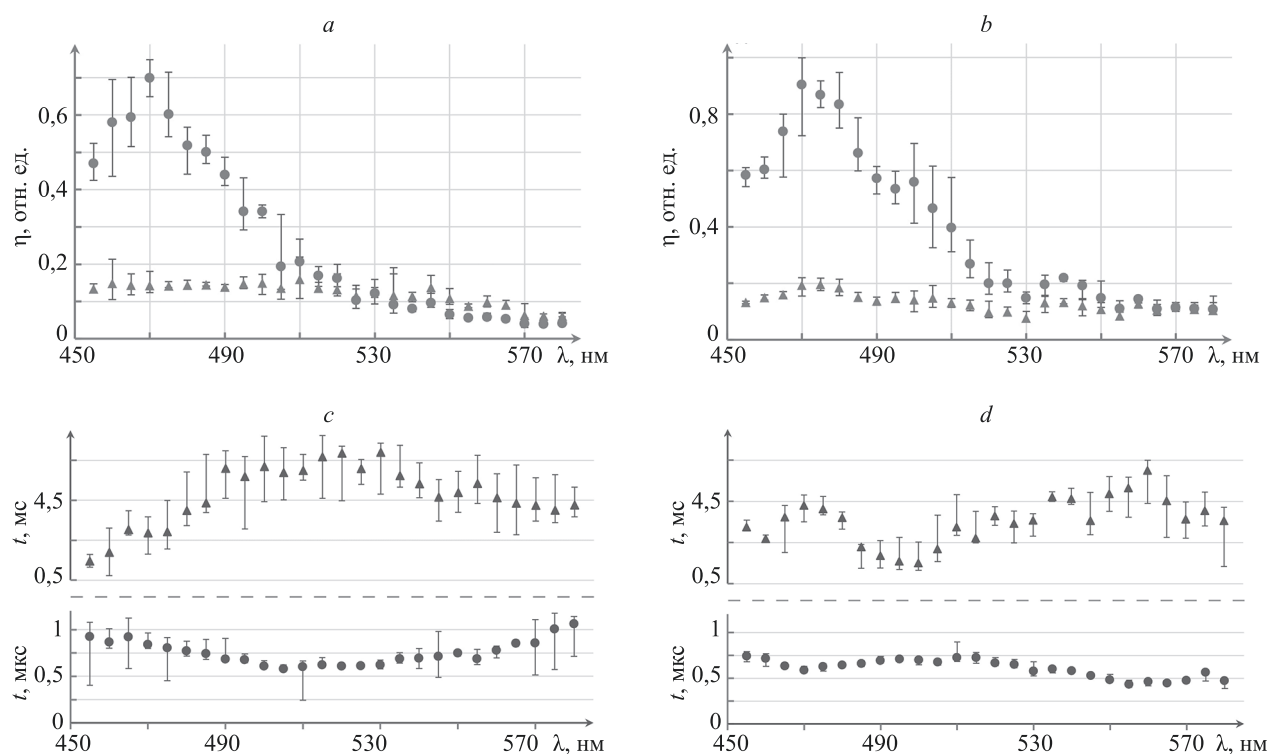


Рис. 3. Зависимости относительной дифракционной эффективности (η) (a, b) и времени формирования (t) (c, d) для коротко- и долгоживущих решеток, полученных при освещении кристалла лазерным импульсом параметрического генератора, от длины волны записывающего излучения при интенсивностях: 1 МВт/см² (a, c) и 5 МВт/см² (b, d).

Обозначения кривых: ● — коротко- и ▲ — долгоживущая решетки

Fig. 3. Dependences of the relative diffraction efficiency (η) (a, b) and formation time (t) (c, d) for short- and long-lived gratings obtained by illuminating the crystal with a laser pulse of a parametric generator on the recording radiation wavelength at intensities: 1 MW/cm² (a, c) and 5 MW/cm² (b, d).

Curve designations: ● — short- and ▲ — long-lived gratings

Закключение

В работе показана возможность оптимизации условий записи голограмм импульсным лазером в спектральном диапазоне от 450 до 600 нм в кристалле силиката висмута. В ходе экспериментального исследования установлена выраженная спектральная зависимость эффективности формирования коротко- и долгоживущих динамических решеток в фоторефрактивных кристаллах силиката висмута, используемых в системах адаптивной интерферометрии. В результате решения задачи установления времени формирования и релаксации динамических решеток в кристалле силиката висмута определено, что в синей области спектра

(450–500 нм) доминирующими являются короткоживущие решетки с временами жизни на уровне сотни микросекунд. При смещении длины волны записывающего излучения в длинноволновую область спектра их дифракционная эффективность уменьшается, и доминирующими становятся долгоживущие решетки с временем жизни в несколько секунд. Определена оптимальная длина волны для записи динамических решеток в силикате висмута, которая составила 470 нм при толщине кристалла 2 мм. При записи на этой длине волны формируемые динамические голограммы имели наибольшую дифракционную эффективность при фиксированной энергии лазерных импульсов на уровне 1 мДж.

Литература

1. Петров М., Степанов С., Хоменко А. Фоторефрактивные кристаллы в когерентной оптике. СПб.: Наука, 1992. 317 с.
2. Шепелевич В.В. Голография в фоторефрактивных оптически активных кристаллах. Мозырь: УО МГПУ им. И.П. Шамякина, 2012. 287 с.
3. Shandarov S.M., Burimov N.I., Kul'chin Yu.N., Romashko R.V., Tolstik A.L., Shepelevich V.V. Dynamic Denisyuk holograms in cubic photorefractive crystals // *Quantum Electronics*. 2008. V. 38. N 11. P. 1059–1069. <https://doi.org/10.1070/QE2008v038n11ABEH013793>
4. Ромашко Р.В. Адаптивная голографическая интерферометрия: техника, прогресс и приложения // *Вестник Дальневосточного*

References

1. Petrov M., Stepanov S., Khomenko A. *Photorefractive Crystals in Coherent Optics*. St. Petersburg, 1992, 317 p. (in Russian)
2. Shepelevich V.V. *Holography in Photorefractive Optically Active Crystals*. Mazyr, Mazyr State Pedagogical University named after I.P. Shamyakin, 2012, 287 p. (in Russian)
3. Shandarov S.M., Burimov N.I., Kul'chin Yu.N., Romashko R.V., Tolstik A.L., Shepelevich V.V. Dynamic Denisyuk holograms in cubic photorefractive crystals. *Quantum Electronics*, 2008, vol. 38, no. 11, pp. 1059–1069. <https://doi.org/10.1070/QE2008v038n11ABEH013793>
4. Romashko R.V. Adaptive holographic interferometry: technique, progress and applications. *Vestnik of Far Eastern Branch of Russian*

- отделения Российской академии наук. 2021. № 4. С. 40–47. https://doi.org/10.37102/0869-7698_2021_218_04_03
5. Peigné A., Bortolozzo U., Residori S., Molin S., Billault V., Nouchi P., Dolfi D., Huignard J. Adaptive interferometry for high-sensitivity optical fiber sensing // *Journal of Lightwave Technology*. 2016. V. 34. N 19. P. 4603–4609. <https://doi.org/10.1109/JLT.2016.2552495>
 6. Каргин Ю.Ф., Бурков В.И., Марин А.А., Егорышева А.В. Кристаллы $\text{Bi}_{12}\text{Si}_x\text{O}_{20-8}$ со структурой силленита. Синтез, строение, свойства. М.: Изд-во ИОХ РАН, 2004. 312 с.
 7. Koc H., Palaz S., Simsek S., Mamedov A., Ozbay E. Elastic and optical properties of sillenites: First principle calculations // *Ferroelectrics*. 2020. V. 557. N 1. P. 98–104. <https://doi.org/10.1080/00150193.2020.1713354>
 8. Толстик А.Л., Ханон Х.К. Динамика фотоиндуцированного поглощения в кристаллах титаната висмута // *Вестник Белорусского государственного университета. Сер. 1. Физика. Математика. Информатика*. 2012. № 2. С. 3–7.
 9. Isik M., Delice S., Nasser H., Gasanly N.M., Darvishov N.H., Bagiev V.E. Optical characteristics of $\text{Bi}_{12}\text{SiO}_{20}$ single crystals by spectroscopic ellipsometry // *Materials Science in Semiconductor Processing*. 2020. V. 120. P. 105286. <https://doi.org/10.1016/j.mssp.2020.105286>
 10. Kornienko T., Kisteneva M., Shandarov S., Tolstik A. Light-induced effects in sillenite crystals with shallow and deep traps // *Physics Procedia*. 2017. V. 86. P. 105–112. <https://doi.org/10.1016/j.phpro.2017.01.029>
 11. Tolstik A.L., Matusevich A.Yu., Kisteneva M.G., Shandarov S.M., Itkin S.I., Mandel' A.E., Kargin Yu.F., Kul'chin Yu.N., Romashko R.V. Spectral dependence of absorption photoinduced in a $\text{Bi}_{12}\text{TiO}_{20}$ crystal by 532-nm laser pulses // *Quantum Electronics*. 2007. V. 37. N 11. P. 1027–1032. <https://doi.org/10.1070/QE2007v037n11ABEH013371>
 12. Noh T.H., Hwang S.W., Kim J.U., Yu H.K., Seo H., Ahn B., Kim D.W., Cho I. Optical properties and visible light-induced photocatalytic activity of bismuth sillenites ($\text{Bi}_{12}\text{XO}_{20}$, X = Si, Ge, Ti) // *Ceramics International*. 2017. V. 43. N 15. P. 12102–12108. <https://doi.org/10.1016/j.ceramint.2017.06.067>
 13. Shandarov S.M., Kisteneva M.G., Akrestina A.S., Vishnev A.S., Kargin Yu.F., Tolstik A.L. Change in absorption induced in a $\text{Bi}_{12}\text{TiO}_{20}$: Ca crystal by 870-nm infrared radiation // *High Energy Chemistry*. 2008. V. 42. N 7. P. 554–556. <https://doi.org/10.1134/S0018143908070163>
 14. Matusevich A., Tolstik A., Kisteneva M., Shandarov S., Matusevich V., Kiessling A., Kowarschik R. Investigation of photo-induced absorption in a $\text{Bi}_{12}\text{TiO}_{20}$ crystal // *Applied Physics B*. 2008. V. 92. N 2. P. 219–224. <https://doi.org/10.1007/s00340-008-3098-z>
 15. Stankevich A.V., Tolstik A.L., Hanon H.K. Photoinduced absorption in bismuth titanate crystals on nano- and picosecond excitation // *Technical Physics Letters*. 2011. V. 37. N 8. P. 746–749. <https://doi.org/10.1134/S1063785011080268>
 16. Dadenkov I.G., Tolstik A.L., Miksyuk Yu.I., Saechnikov K.A. Photoinduced absorption and pulsed recording of dynamic holograms in bismuth silicate crystals // *Optics and Spectroscopy*. 2020. V. 128. N 9. P. 1401–1406. <https://doi.org/10.1134/S0030400X20090052>
 - Academy of Sciences, 2021, no. 4, pp. 40–47. (in Russian). https://doi.org/10.37102/0869-7698_2021_218_04_03
 5. Peigné A., Bortolozzo U., Residori S., Molin S., Billault V., Nouchi P., Dolfi D., Huignard J. Adaptive interferometry for high-sensitivity optical fiber sensing. *Journal of Lightwave Technology*, 2016, vol. 34, no. 19, pp. 4603–4609. <https://doi.org/10.1109/JLT.2016.2552495>
 6. Kargin Yu.F., Burkov V.I., Marin A.A., Egorysheva A.V. *Bi₁₂Si_xO₂₀₋₈ Crystals with Sillenite Structure. Synthesis, Morphology, Properties*. Moscow, 2004, 312 p. (in Russian)
 7. Koc H., Palaz S., Simsek S., Mamedov A., Ozbay E. Elastic and optical properties of sillenites: First principle calculations. *Ferroelectrics*, 2020, vol. 557, no. 1, pp. 98–104. <https://doi.org/10.1080/00150193.2020.1713354>
 8. Tolstik A.L., Hanon H.K. Dynamics of the photo-induced absorption in crystals of bismuth titanate. *Journal of the Belarusian State University. Physics, Mathematics, Informatics*, 2012, no. 2, pp. 3–7. (in Russian)
 9. Isik M., Delice S., Nasser H., Gasanly N.M., Darvishov N.H., Bagiev V.E. Optical characteristics of $\text{Bi}_{12}\text{SiO}_{20}$ single crystals by spectroscopic ellipsometry. *Materials Science in Semiconductor Processing*, 2020, vol. 120, pp. 105286. <https://doi.org/10.1016/j.mssp.2020.105286>
 10. Kornienko T., Kisteneva M., Shandarov S., Tolstik A. Light-induced effects in sillenite crystals with shallow and deep traps. *Physics Procedia*, 2017, vol. 86, pp. 105–112. <https://doi.org/10.1016/j.phpro.2017.01.029>
 11. Tolstik A.L., Matusevich A.Yu., Kisteneva M.G., Shandarov S.M., Itkin S.I., Mandel' A.E., Kargin Yu.F., Kul'chin Yu.N., Romashko R.V. Spectral dependence of absorption photoinduced in a $\text{Bi}_{12}\text{TiO}_{20}$ crystal by 532-nm laser pulses. *Quantum Electronics*, 2007, vol. 37, no. 11, pp. 1027–1032. <https://doi.org/10.1070/QE2007v037n11ABEH013371>
 12. Noh T.H., Hwang S.W., Kim J.U., Yu H.K., Seo H., Ahn B., Kim D.W., Cho I. Optical properties and visible light-induced photocatalytic activity of bismuth sillenites ($\text{Bi}_{12}\text{XO}_{20}$, X = Si, Ge, Ti). *Ceramics International*, 2017, vol. 43, no. 15, pp. 12102–12108. <https://doi.org/10.1016/j.ceramint.2017.06.067>
 13. Shandarov S.M., Kisteneva M.G., Akrestina A.S., Vishnev A.S., Kargin Yu.F., Tolstik A.L. Change in absorption induced in a $\text{Bi}_{12}\text{TiO}_{20}$: Ca crystal by 870-nm infrared radiation. *High Energy Chemistry*, 2008, vol. 42, no. 7, pp. 554–556. <https://doi.org/10.1134/S0018143908070163>
 14. Matusevich A., Tolstik A., Kisteneva M., Shandarov S., Matusevich V., Kiessling A., Kowarschik R. Investigation of photo-induced absorption in a $\text{Bi}_{12}\text{TiO}_{20}$ crystal. *Applied Physics B*, 2008, vol. 92, no. 2, pp. 219–224. <https://doi.org/10.1007/s00340-008-3098-z>
 15. Stankevich A.V., Tolstik A.L., Hanon H.K. Photoinduced absorption in bismuth titanate crystals on nano- and picosecond excitation. *Technical Physics Letters*, 2011, vol. 37, no. 8, pp. 746–749. <https://doi.org/10.1134/S1063785011080268>
 16. Dadenkov I.G., Tolstik A.L., Miksyuk Yu.I., Saechnikov K.A. Photoinduced absorption and pulsed recording of dynamic holograms in bismuth silicate crystals. *Optics and Spectroscopy*, 2020, vol. 128, no. 9, pp. 1401–1406. <https://doi.org/10.1134/S0030400X20090052>

Авторы

Даденков Иван Геннадьевич — младший научный сотрудник, Белорусский государственный университет, Минск, 220030, Республика Беларусь, <https://orcid.org/0000-0003-0722-4683>, dadenkov.ivan@gmail.com

Толстик Алексей Леонидович — доктор физико-математических наук, профессор, заведующий кафедрой, Белорусский государственный университет, Минск, 220030, Республика Беларусь, <https://orcid.org/0000-0003-4953-4890>, tolstik@bsu.by

Миксюк Юрий Иванович — кандидат физико-математических наук, доцент, доцент, Белорусский государственный педагогический университет имени Максима Танка, Минск, 220050, Республика Беларусь, <https://orcid.org/0000-0002-3969-0454>, miksiuk@mail.ru

Саечников Константин Алексеевич — кандидат физико-математических наук, доцент, доцент, Белорусский государственный педагогический университет имени Максима Танка, Минск, 220050, Республика Беларусь, <https://orcid.org/0000-0001-8591-9312>, skabspu@tut.by

Authors

Ivan G. Dadenkov — Junior Researcher, Belarusian State University, Minsk, 220030, Republic of Belarus, <https://orcid.org/0000-0003-0722-4683>, dadenkov.ivan@gmail.com

Alexei L. Tolstik — D. Sc. (Physics & Mathematics), Head of Department, Belarusian State University, Minsk, 220030, Republic of Belarus, <https://orcid.org/0000-0003-4953-4890>, tolstik@bsu.by

Yuri I. Miksiuk — PhD (Physics & Mathematics), Associate Professor, Associate Professor, Belarusian State Pedagogical University named after Maxim Tank, Minsk, 220050, Republic of Belarus, <https://orcid.org/0000-0002-3969-0454>, miksiuk@mail.ru

Konstantin A. Saechnikov — PhD (Physics & Mathematics), Associate Professor, Associate Professor, Belarusian State Pedagogical University named after Maxim Tank, Minsk, 220050, Republic of Belarus, <https://orcid.org/0000-0001-8591-9312>, skabspu@tut.by

doi: 10.17586/2226-1494-2022-22-6-1031-1036

УДК 681.784.8

Гибридный эндоскоп с телевизионной и многоспектральной обработкой изображений для диагностики рака внутренних органов

Кирилл Вадимович Зайченко¹, Борис Симхович Гуревич²,
Андрей Владимирович Беляев³, Виталия Игоревна Святкина⁴

^{1,2,3,4} Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация

¹ kvz235@mail.ru, <https://orcid.org/0000-0002-2881-4386>

² bgurevich48@gmail.com, <https://orcid.org/0000-0002-7520-7087>

³ Abel2004@inbox.ru, <https://orcid.org/0000-0001-7605-5743>

⁴ svyatkina.vi@edu.spbstu.ru, <https://orcid.org/0000-0002-6456-8776>

Аннотация

Предмет исследования. Эндоскопы широко применяются для диагностики различных заболеваний внутренних органов. Наиболее совершенные эндоскопы содержат встроенные в зонд миниатюрные телевизионные камеры, которые обеспечивают высокое пространственное разрешение. Однако традиционные телевизионные камеры существенно искажают спектральный состав передаваемого изображения, так как в системе RGB воспроизводят только ограниченные по ширине спектральные линии. Вместе с тем полный оптический спектр изображения обследуемого внутреннего органа несет более обширную информацию, которая может оказаться решающей при диагностике онкологических заболеваний. Таким образом, важную роль играет спектральное разрешение изображения. **Метод.** Для эффективного выявления одновременно пространственной и спектральной информации предложено применять многоспектральную обработку. В результате обработки возможно получение серии монохромных субизображений, каждое из которых соответствует выбранному разрешаемому спектральному интервалу. В качестве селективного элемента, позволяющего в реальном времени выделять субизображения, использованы акустооптические перестраиваемые фильтры. При этом увеличение спектрального разрешения сопровождается падением пространственного разрешения. Для решения этой проблемы предложен новый гибридный эндоскоп, в котором реализован метод совместного применения телевизионной камеры и устройства многоспектральной обработки. **Основные результаты.** Показано, что предложенный метод гибридной эндоскопии дает возможность обеспечить высокие разрешения: пространственное с помощью камеры и спектральное (не менее сотни разрешимых интервалов) благодаря многоспектральной обработке. Получаемые оптические изображения позволяют выявлять четкие очертания области новообразования. Результаты многоспектральной обработки диагностируют злокачественную природу новообразований. Представлена структурная схема предложенного эндоскопа, и показаны принципы управления. Определены характеристики телевизионных и многоспектральных изображений. Оценка линейности характеристик акустооптических фильтров показала возможность передачи максимального объема информации за счет линейности преобразования сигнала в звеньях фильтров. **Практическая значимость.** Реализация предложенного метода в конструкции гибридного эндоскопа позволила получать наиболее полную спектральную информацию о новообразованиях и обеспечить возможность диагностики рака внутренних органов на более ранних стадиях развития.

Ключевые слова

гибридный эндоскоп, многоспектральный датчик, телевизионная камера, пространственная информация, спектральная и градиционная информация, рак внутренних органов, ранняя диагностика

Благодарности

Работа поддержана Минобрнауки Российской Федерации, госзадание № 075-00761-22-00, тема № FZZM-2022-0011.

Ссылка для цитирования: Зайченко К.В., Гуревич Б.С., Беляев А.В., Святкина В.И. Гибридный эндоскоп с телевизионной и многоспектральной обработкой изображений для диагностики рака внутренних органов // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1031–1036. doi: 10.17586/2226-1494-2022-22-6-1031-1036

© Зайченко К.В., Гуревич Б.С., Беляев А.В., Святкина В.И., 2022

Hybrid endoscope with television and multispectral image processing for the internal organs cancer early diagnostics

Kirill V. Zaichenko¹, Boris S. Gurevich²✉, Andrey V. Belyaev³, Vitalia I. Svyatkina⁴

^{1,2,3,4} Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation

¹ kvz235@mail.ru, <https://orcid.org/0000-0002-2881-4386>

² bgurevich48@gmail.com✉, <https://orcid.org/0000-0002-7520-7087>

³ Abel2004@inbox.ru, <https://orcid.org/0000-0001-7605-5743>

⁴ svyatkina.vi@edu.spbstu.ru, <https://orcid.org/0000-0002-6456-8776>

Abstract

Endoscopes are widely used to diagnose various internal diseases. The most advanced endoscopes contain miniature television cameras built into the probe which provide high spatial resolution. However, traditional television cameras significantly distort the spectral composition of the transmitted image because in the RGB system they reproduce only limited width spectral lines. However, the full optical spectrum of an image of an internal organ under examination provides a broader range of information that can be crucial in cancer diagnostics. Thus, the spectral resolution of the image plays an important role. Multispectral processing has been suggested for efficient detection of both spatial and spectral information simultaneously. As a result, a series of monochrome sub-images can be obtained, each of which corresponds to a selected resolved spectral interval. Acousto-optical tunable filters are used as a selective element allowing real-time extraction of sub-images. However, an increase in spectral resolution is accompanied by a decrease in spatial resolution. A new hybrid endoscope was suggested to solve this problem. It implements the method of combined application of a TV camera and a multispectral processing device. It is shown that the suggested method of hybrid endoscopy makes it possible to ensure high resolutions: spatial resolution with a camera and spectral resolution (no less than a hundred of resolvable intervals) due to multispectral processing. The resulting optical images reveal a clear outline of the neoplasm area. The results of multispectral processing diagnose the malignant nature of the neoplasm. A schematic diagram of the proposed endoscope is presented, and the principles of control are shown. The characteristics of the television and multispectral images are determined. An estimation of the linearity characteristics of the acousto-optical filters shows the possibility of transmitting the maximum amount of information due to the linearity of the signal transformation in the filter sections. Implementation of the proposed method in the design of a hybrid endoscope allowed obtaining the most complete spectral information about neoplasms and providing an opportunity of diagnostics of cancer of internal organs at earlier stages of development.

Keywords

hybrid endoscope, multispectral probe, television camera, spatial information, spectral and gray scale information, cancer of internal organs, early diagnostics

Acknowledgements

The work has been supported by Education and Science Ministry of Russian Federation, State task No. 075-00761-22-00, project No. FZZM-2022-0011.

For citation: Zaichenko K.V., Gurevich B.S., Belyaev A.V., Svyatkina V.I. Hybrid endoscope with television and multispectral image processing for the internal organs cancer early diagnostics. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1031–1036 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1031-1036

Введение

При диагностике ряда заболеваний, особенно онкологических, применяются эндоскопы для визуального наблюдения фрагментов внутренних органов. В современных эндоскопах используются видеокамеры, которые передают изображение изучаемого объекта на экран монитора. При этом искажается или теряется значительная часть информации о спектре отражения света от внутреннего органа, из-за наличия ограниченный телевизионных систем по градации и спектральному составу света. Эти потери и искажения значительно уменьшают надежность онкологической диагностики.

В работах [1, 2] показано, что общий объем информации, содержащейся в изображении, пропорционален числу спектральных интервалов, разрешаемых системой обработки. Для одновременной передачи пространственной и спектральной информации, при условии, что данные о цветном изображении не могут быть переданы без использования промежуточных устройств, целесообразно применять метод многоспектральной

обработки. Метод включает в себя последовательную передачу множества субизображений, каждое из которых содержит часть данных о полном изображении в узком спектральном интервале [3, 4] и позволяет использовать до сотни и более интервалов [5, 6]. В то же время изображение, переданное с видеокамеры на монитор, содержит лишь три основных цвета — красный, зеленый и синий, в результате могут быть получены только три субизображения. Следовательно, объем спектральной информации, наблюдаемой с экрана монитора, примерно в 30 раз меньше, по сравнению с полученным аналоговым изображением с помощью многоспектральной обработки. К изображениям, образованным с помощью эндоскопов, предъявляются требования к наиболее полному составу информации о наблюдаемом объекте, в том числе о его спектральных характеристиках отражения, что позволяет более надежно определить патологию. Искажение или потеря таких данных может привести к опасным диагностическим ошибкам. В результате сформировалась идея о совмещении телевизионного и многоспектрального

(с передачей данных по оптоволокну) методов передачи изображений в одном устройстве — гибридном эндоскопе. При этом наличие в анализируемом изображении спектральных составляющих, характерных для отражения света от тканей и органов, пораженных раком, сигнализирует о наличии злокачественных новообразований.

Цель работы — разработка метода гибридной эндоскопии, позволяющего минимизировать потери спектральной и пространственной информации путем совмещения процедур передачи изображений по оптическому многоспектральному и телевизионному каналам.

Принцип действия гибридного эндоскопа

На рис. 1 приведена структурная схема реализации принципа предложенного устройства — гибридного эндоскопа.

Рассмотрим принцип работы гибридного эндоскопа. В выбранную область исследования введем зонд, содержащий волоконный световод подсветки 1 и оптоволоконный датчик 2 с регулирующим микроприводом, а также миниатюрную видеокамеру 3 с соответствующим кабелем. С помощью телевизионной камеры на контроллере телевизионного ТВ канала выберем фрагмент органа, и выведем его изображение с помощью датчика 2. Направим с помощью микропривода конец световода 1 с микролинзой в сторону исследуемого участка. Спектральный состав излучения подсветки выберем исходя из принципа, описанного в работах [7–9]. Выполним анализ изображения, сформированного из жгута оптоволокон при поступлении отраженного света на вход датчика 2. Таким образом, два изображения —

переданное оптоволоком и телекамерой — дополняют друг друга. Телевизионное изображение дает более высокое разрешение, а оптоволоконное — менее искаженные спектральную и градиационную информации. По данным телевизионного изображения в пределах выбранного фрагмента через микропривод настроим оптоволоконный датчик на новый участок, и повторим процедуру для получения нового изображения.

При необходимости можно выбрать следующий фрагмент исследуемого органа для анализа и повторить перечисленные процедуры. Таким образом, будет собрана с минимальными потерями визуальная информация об объекте исследования. Более подробно описанные процедуры могут быть проанализированы в соответствии с концепцией многокритериальной оптимизации, сформулированной в ведущей научной школе Российской Федерации № 3455.2012.8, под руководством профессора, доктора технических наук К.В. Зайченко. При этом необходимо присвоить различным видам информации соответствующие иерархические весовые коэффициенты. Такие коэффициенты могут существенно различаться в зависимости от задач, решаемых анализируемым устройством [10]. В частности, в системах, ориентированных на восприятие человеческим глазом, наиболее высокий весовой коэффициент присваивается пространственной информации по горизонтали, несколько меньший — по вертикали, и далее по убыванию — информация по времени, пространственная по глубине, полутоновая информация, цветовая, спектральная, фазовая и поляризационная [11]. Для диагностической же системы, описываемой в настоящей работе, важнейшим видом информации является спектральная, которой присваивается максимальный весовой коэффициент.

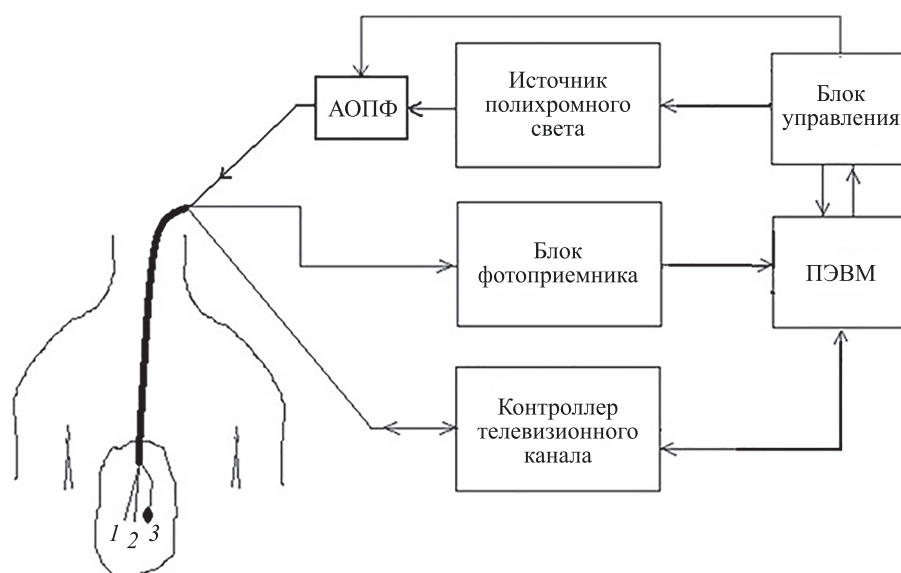


Рис. 1. Структурная схема гибридного эндоскопа с использованием опорного телевизионного изображения.

1 — волоконный световод подсветки; 2 — оптоволоконный датчик передачи монохромного изображения; 3 — видеокамера; АОПФ — акустооптический перестраиваемый фильтр; ПЭВМ — персональный компьютер

Fig. 1. Structural circuit of the hybrid endoscope which uses the reference television image.

Legend: 1 — fiber backlight guide; 2 — fiber probe for monochrome image transmission; 3 — TV camera; АОПФ — acousto-optic tunable filter; ПЭВМ — personal computer

Перестройка длины волны света для освещения исследуемого органа

Одно из основных требований к гибричному эндоскопу — быстрая, желательно программная, перестройка длины волны освещающего пучка света. Это связано с тем, что время нахождения зондирующей головки эндоскопа в полости пациента должно быть минимизировано, поэтому применение механически движущихся деталей для смены длины волны излучения исключается. В качестве источника излучения с переменной, программно управляемой длиной волны выбран осветитель на базе набора светодиодов, построенный в соответствии с принципом, запатентованным в работе [6], который показан на рис. 2.

Источник света состоит из управляемого компьютером блока питания, который подает ток на светодиоды в соответствии с поступающими из компьютера сигналами. Это дает возможность управления интенсивностью и спектральным составом излучения по причине наличия возможности включения светодиодов в произвольной последовательности, изменения подаваемого на любой из светодиодов тока, включения нескольких произвольных светодиодов с заданной интенсивностью. Микрооптическая сборка позволяет сфокусировать излучение с каждого из светодиодов на зеркало, которое отражает это излучение на дифракционную решетку. Зеркало служит для уменьшения габаритов устройства.

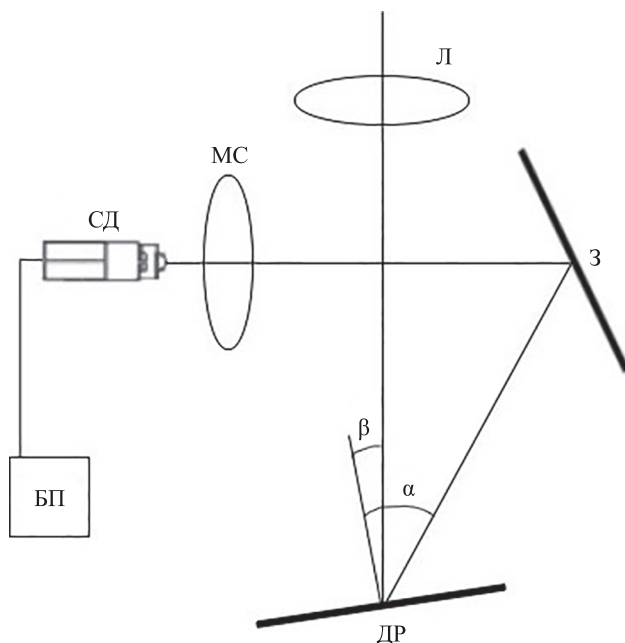


Рис. 2. Схема работы источника света с программным управлением.

БП — блок питания; СД — светодиод; МС — микрооптическая сборка; Л — линза; З — зеркало; ДР — дифракционная решетка

Fig. 2. Operation circuit of the light source with software control.

Legend: БП — feeding unit; СД — light emission diode; МС — micro-optic assembly; Л — lens; З — mirror; ДР — diffraction grating

Излучение с каждого из светодиодов, отраженное от зеркала, попадает на дифракционную решетку таким образом, что для каждого из лучей выполняется условие из уравнения:

$$d(\sin \alpha + \sin \beta) = m\lambda,$$

где d — период дифракционной решетки; α — угол между нормалью к дифракционной решетке и направлением распространения излучения от светоизлучающего элемента; λ — длина волны излучения; β — угол дифракции, измеренный относительно нормали к дифракционной решетке; m — целое число, которое характеризует дифракционный порядок.

В результате первый дифракционный порядок излучения от всех светодиодов уходит в одном и том же направлении. Переключение светодиодов обеспечивает грубую настройку длины волны света. Полоса излучения светодиодов составляет примерно 30–50 нм. Для повышения селективности авторами предложено использовать акустооптические перестраиваемые фильтры (АОПФ) [12, 13], которые управляются изменением частоты внешнего сигнала, подаваемого на пьезоэлектрический преобразователь ячейки Брэгга. Быстродействие этого селективного элемента (АОПФ) ограничивается временем пересечения апертуры пучка света, падающего на ячейку Брэгга и волновым фронтом акустической волны. Это быстродействие определяется так называемой временной апертурой $\tau = D/v$, где D — линейный размер апертуры падающего полихромного пучка света, а v — скорость акустической волны в звукопроводе ячейки Брэгга. Величина τ для наиболее эффективного режима работы ячейки Брэгга из параллеллурита на медленной сдвиговой акустической волне составляет 5–10 мкс. Применение монокристалла параллеллурита в качестве среды звукопровода ячейки Брэгга обусловлено тем, что он ориентирован в определенном кристаллографическом направлении и обеспечивает прохождение сдвиговой акустической волны, для которой эффективность акустооптического взаимодействия (т. е. доля света, направляемая в первый дифракционный порядок) максимальна. Использование акустической мощности порядка 1 Вт в этих условиях позволяет направить в первый дифракционный порядок почти 100 % мощности падающего на ячейку Брэгга оптического излучения.

Теоретические предпосылки использования АОПФ в качестве селективного элемента в многоспектральных диагностических устройствах

АОПФ, как и многие оптические и оптико-электронные приборы, может быть представлен в виде последовательности звеньев, каждое из которых описывает определенное физическое преобразование входного сигнала [14]. В отличие от других видов акустооптических устройств, в схематическое представление АОПФ могут быть включены только два основных звена. В первом из них входное распределение интенсивности света по длине волны преобразуется в распределение

интенсивности света по времени, а второе звено (или группа звеньев) обеспечивает детектирование выходного оптического сигнала. Рассмотрим первое звено.

Отметим, что входной пучок света постоянен во времени. Его интенсивность I может быть представлена как интеграл распределения интенсивности света по длине волны в представляющем интерес диапазоне. При этом выходное распределение интенсивности света для первого звена может иметь вид

$$\frac{dI'}{dt} = \left(\frac{dI}{d\lambda} \right) \left(\frac{d\lambda}{dt} \right), \quad (1)$$

где в правой части уравнения первая производная интенсивности света по длине волны, а вторая производная определяет процедуру выделения длины волны. Тогда запишем

$$k\lambda = F(\omega, t), \quad (2)$$

где k — размерный коэффициент пропорциональности; $F(\omega, t)$ — спектр сигнала, вводимого в ячейку Брэгга АОПФ через пьезоэлектрический преобразователь. Зависимость изменения во времени спектра из уравнения (2) определяется вторым сомножителем правой части уравнения (1). Следовательно,

$$I' = \int_0^T I_\lambda k \left(\frac{d[F(\omega, t)]}{dt} \right) dt, \quad (3)$$

где I_λ — распределение входной интенсивности света по длине волны; T — время накопления, в течение которого на выходе АОПФ сохраняется пучок света с определенным выделением спектральным интервалом. Минимальное значение T близко к временной апертуре прибора τ , т. е. времени, необходимому акустическому волновому фронту в ячейке Брэгга для пересечения апертуры входного светового пучка.

Выражение (3) содержит противоречие. Действительно, если спектр управляющего сигнала, подаваемого на пьезопреобразователь, не изменяется, то интенсивность выходного света равна нулю, что следует

из выражения (3). Однако это может иметь место лишь в случае, когда селективность АОПФ не ограничена, и он выделяет свет с нулевой шириной полосы длин волн. На практике селективность АОПФ составляет величину $\Delta\lambda$, которая определяется характеристиками дифракции Брэгга, и может быть рассчитана из теории связанных волн Когельника [15]. Тогда получим

$$I' = \int_0^T I_\lambda k \frac{F(\omega, t)}{\Delta\lambda} \left(\frac{d\lambda}{dt} \right) dt. \quad (4)$$

Таким образом, можно считать, что выражение (4) представляет собой передаточную функцию первого звена.

Второе звено преобразования сигнала может рассматриваться как сочетание трех субзвеньев с функциями преобразований: входной интенсивности света в мгновенную концентрацию носителей тока; концентрации носителей тока в величину заряда, накопленного на ячейках фотоприемника; распределения заряда по фотоприемнику в электрический ток, снимаемый с фотоприемника [16]. Отметим, что передаточная функция всех трех субзвеньев весьма близка к линейной и позволяет минимизировать потери передаваемой информации.

Заключение

Предложенный гибридный эндоскоп, может обеспечить получение наиболее полного спектра отражения от исследуемых внутренних органов. Это позволит принимать диагностические решения при исследовании онкологических новообразований. При этом источник полихромного света, использующий современные низковольтные светодиоды, обеспечивает минимизацию потерь информации.

Рассмотренный принцип построения гибридного эндоскопа может быть реализован в гастроскопии, бронхоскопии, колоноскопии, лапароскопии и во всех иных возможных областях, где видеoinформация об исследуемых органах позволяет диагностировать наличие патологий.

Литература

1. Зайченко К.В., Гуревич Б.С. Принципы измерений информационных и технических характеристик акустооптических систем многоспектральной обработки изображений // Датчики и системы. 2015. № 2. С. 61–64.
2. Zaichenko K.V., Gurevich S.B., Gurevich B.S. Application of optical freedom degrees principle to acousto-optic devices // *Physics Procedia*. 2015. V. 70. P. 774–778. <https://doi.org/10.1016/j.phpro.2015.08.265>
3. Calpe-Maravilla J., Vila-Frances J., Ribes-Gómez E., Durán-Bosch V., Muñoz-Mari J., Amorós-Lopez J., Gómez-Chova L., Tajahuerce E. 400- to 1000-nm imaging spectrometer based on acousto-optic tunable filters // *Journal of Electronic Imaging*. 2006. V. 15. N 2. P. 023001. <https://doi.org/10.1117/1.2201057>
4. Gupta N. Hyperspectral and polarization imaging application of acousto-optic tunable filters // *Proc. of World Congress on Ultrasonics*. Paris. 2003. P. 345–348.
5. Kutuza B., Pozhar V.E., Pustovoit V.I. AOTF-based imaging spectrometers for research of small-size biological objects //

References

1. Zaychenko K.V., Gurevich B.S. Measurement principles of information and technical characteristics of the multispectral image processing acousto-optic systems. *Sensors & Systems*, 2015, no. 2, pp. 61–64. (in Russian)
2. Zaichenko K.V., Gurevich S.B., Gurevich B.S. Application of optical freedom degrees principle to acousto-optic devices. *Physics Procedia*, 2015, vol. 70, pp. 774–778. <https://doi.org/10.1016/j.phpro.2015.08.265>
3. Calpe-Maravilla J., Vila-Frances J., Ribes-Gómez E., Durán-Bosch V., Muñoz-Mari J., Amorós-Lopez J., Gómez-Chova L., Tajahuerce E. 400- to 1000-nm imaging spectrometer based on acousto-optic tunable filters. *Journal of Electronic Imaging*, 2006, vol. 15, no. 2, pp. 023001. <https://doi.org/10.1117/1.2201057>
4. Gupta N. Hyperspectral and polarization imaging application of acousto-optic tunable filters. *Proc. of World Congress on Ultrasonics*. Paris, 2003, pp. 345–348.
5. Kutuza B., Pozhar V.E., Pustovoit V.I. AOTF-based imaging spectrometers for research of small-size biological objects.

- Proceedings of SPIE. 2003. V. 5143. P. 165–169. <https://doi.org/10.1117/12.500528>
6. Yushkov K.B., Bogomolov D.V., Voloshinov V.B. Acousto-optic imaging by means of wide-angle tunable acousto-optic filter // *Molecular and Quantum Acoustics*. 2007. V. 28. P. 299–304.
 7. Шаповалов В.В., Гуревич Б.С., Колесов И.А., Андреев С.В., Беляев А.В. Источник света с произвольно регулируемым спектральным составом для биомедицинских спектральных анализаторов // *Биомедицинская электроника*. 2009. № 11. С. 16–20.
 8. Semyonov D.V., Nippolainen E., Kamshilin A.A., Belyaev A.V., Andreev S.V., Gurevich B.S. An ultra-fast distance sensor based on dynamic speckles generated by acousto-optic deflection // *Measurement Science and Technology*. 2006. V. 17. N 11. P. 2906–2912. <https://doi.org/10.1088/0957-0233/17/11/007>
 9. Андреев С.В., Беляев А.В., Гуревич Б.С., Земский В.И., Соколов В.Н., Шаповалов В.В. Универсальный источник полихромного излучения. Патент RU 2287736. Бюл. 2006. № 32.
 10. Гуревич С.Б., Гуревич Б.С., Жумалиев К.М. Проблемы информационной оптоэлектроники. СПб.: Наука, 2008. 210 с.
 11. Акаев А.А., Жумалиев К.М., Гуревич С.Б., Гуревич Б.С. Связь оценок качества изображения с информационными характеристиками изображающих систем и памяти // *Оптический журнал*. 1997. Т. 64. № 2. С. 61–66.
 12. Gupta N., Dahmani R. Multispectral and hyperspectral imaging with AOTF for object recognition // *Proceedings of SPIE*. 1999. V. 3584. P. 128–135. <https://doi.org/10.1117/12.339814>
 13. Dong Y., You Z., Gao P. Acousto-optic tunable filter for spectral imaging // *Proceedings of SPIE*. 2002. V. 4919. P. 269–274. <https://doi.org/10.1117/12.465653>
 14. Gurevich B.S., Andreyev S.V., Belyaev A.V., Chelak V.N., Sagymbaeva K.A. Investigation of information transmission processes in acousto-optic spectrophotometer // *Proc. of the World Congress on Ultrasonics*. Paris. 2003. P. 365–368.
 15. Kogelnik H. Coupled wave theory for thick hologram gratings // *Bell System Technical Journal*. 1969. V. 48. N 9. P. 2909–2947. <https://doi.org/10.1002/j.1538-7305.1969.tb01198.x>
 16. Gurevich B.S., Andreyev S.V., Vorobiev V.V., Rodionov A.A., Sokolov V.N. Information losses in acousto-optics, their reasons, and ways to minimize them // *Proceedings of SPIE*. 1998. V. 3464. P. 55–63. <https://doi.org/10.1117/12.323153>
 6. Yushkov K.B., Bogomolov D.V., Voloshinov V.B. Acousto-optic imaging by means of wide-angle tunable acousto-optic filter. *Molecular and Quantum Acoustics*, 2007, vol. 28, pp. 299–304.
 7. Shapovalov V.V., Gurevich B.S., Kolesov I.A., Andreev S.V., Belyaev A.V. Light source with arbitrarily adjustable spectral distribution for biomedical spectrum analyzers. *Journal Biomedical Radioelectronics*, 2009, no. 11, pp. 16–20. (in Russian)
 8. Semyonov D.V., Nippolainen E., Kamshilin A.A., Belyaev A.V., Andreev S.V., Gurevich B.S. An ultra-fast distance sensor based on dynamic speckles generated by acousto-optic deflection. *Measurement Science and Technology*, 2006, vol. 17, no. 11, pp. 2906–2912. <https://doi.org/10.1088/0957-0233/17/11/007>
 9. Andreev S.V., Belyaev A.V., Gurevich B.S., Zemskij V.I., Sokolov V.N., Shapovalov V.V. Universal source of polychromatic optical radiation. *Patent RU 2287736*, 2006. (in Russian)
 10. Gurevich S.B., Gurevich B.S., Zhumaliev K.M. *Information Optoelectronics Problems*. St. Petersburg, Nauka Publ., 2008, 210 p. (in Russian)
 11. Akaev A.A., Zhumaliev K.M., Gurevich S.B., Gurevich B.S. Connecting estimates of image quality to the informational characteristics of imaging systems and memory. *Journal of Optical Technology*, 1997, vol. 64, no. 2, pp. 127–131.
 12. Gupta N., Dahmani R. Multispectral and hyperspectral imaging with AOTF for object recognition. *Proceedings of SPIE*, 1999, vol. 3584, pp. 128–135. <https://doi.org/10.1117/12.339814>
 13. Dong Y., You Z., Gao P. Acousto-optic tunable filter for spectral imaging. *Proceedings of SPIE*, 2002, vol. 4919, pp. 269–274. <https://doi.org/10.1117/12.465653>
 14. Gurevich B.S., Andreyev S.V., Belyaev A.V., Chelak V.N., Sagymbaeva K.A. Investigation of information transmission processes in acousto-optic spectrophotometer. *Proc. of the World Congress on Ultrasonics*. Paris, 2003, pp. 365–368.
 15. Kogelnik H. Coupled wave theory for thick hologram gratings. *Bell System Technical Journal*, 1969, vol. 48, no. 9, pp. 2909–2947. <https://doi.org/10.1002/j.1538-7305.1969.tb01198.x>
 16. Gurevich B.S., Andreyev S.V., Vorobiev V.V., Rodionov A.A., Sokolov V.N. Information losses in acousto-optics, their reasons, and ways to minimize them. *Proceedings of SPIE*, 1998, vol. 3464, pp. 55–63. <https://doi.org/10.1117/12.323153>

Авторы

Зайченко Кирилл Владимирович — доктор технических наук, профессор, заведующий лабораторией, Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация, <https://orcid.org/0000-0002-2881-4386>, kvz235@mail.ru

Гуревич Борис Симхович — доктор технических наук, главный научный сотрудник, Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация, <https://orcid.org/0000-0002-7520-7087>, bgurevich48@gmail.com

Беляев Андрей Владимирович — ведущий электроник, Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация, <https://orcid.org/0000-0001-7605-5743>, Abel2004@inbox.ru

Святкина Виталия Игоревна — техник, Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация, <https://orcid.org/0000-0002-6456-8776>, svyatkina.vi@edu.spbstu.ru

Authors

Kirill V. Zaichenko — D. Sc., Professor, Head of Laboratory, Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation, <https://orcid.org/0000-0002-2881-4386>, kvz235@mail.ru

Boris S. Gurevich — D. Sc., Chief Researcher, Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation, <https://orcid.org/0000-0002-7520-7087>, bgurevich48@gmail.com

Andrey V. Belyaev — Leading Electronic Specialist, Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation, <https://orcid.org/0000-0001-7605-5743>, Abel2004@inbox.ru

Vitalia I. Svyatkina — Technician, Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation, <https://orcid.org/0000-0002-6456-8776>, svyatkina.vi@edu.spbstu.ru

Статья поступила в редакцию 12.07.2022
Одобрена после рецензирования 28.09.2022
Принята к печати 12.11.2022

Received 12.07.2022
Approved after reviewing 28.09.2022
Accepted 12.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1037-1047

УДК 681.78

Моделирование композитного волноводного голографического дисплея

Данила Юрьевич Харитонов¹, Дамир Маратович Ахметов²,
Эдуард Ринатович Муслимов³, Айдар Рустемович Гильфанов⁴,
Надежда Константиновна Павлычева⁵

^{1,2,4} Научно-производственное объединение «Государственный институт прикладной оптики», Казань, 420075, Республика Татарстан

^{3,5} Казанский национальный исследовательский технический университет им. А.Н. Туполева – КАИ, Казань, 420075, Республика Татарстан

³ Нидерландский институт радиоастрономии, Двингело, 7991 PD, Нидерланды

³ Астрофизическая лаборатория Марселя, Марсель, 13388, Франция

¹ danila-haritonov2017@mail.ru, <https://orcid.org/0000-0002-1256-4474>

² akhmetov.damir.97@mail.ru, <https://orcid.org/0000-0001-5827-2648>

³ e0123@mail.ru, <https://orcid.org/0000-0002-3242-9894>

⁴ cataklizm@inbox.ru, <https://orcid.org/0000-0001-8109-5029>

⁵ nkpavlych@rambler.ru, <https://orcid.org/0000-0001-9395-3967>

Аннотация

Предмет исследования. Исследованы оптические схемы дисплеев дополненной реальности волноводного типа. Дисплеи, построенные на основе объемных фазовых голограмм, отличаются малыми размерами, большим выходным зрачком и высоким коэффициентом пропускания в каналах проецируемого изображения и прямого зрения. Однако с увеличением апертуры, поля зрения и рабочего спектрального диапазона увеличивается разброс значений угла падения луча и длины волны излучения при решении задачи дифракции в разных точках поверхности голограммы. Это накладывает ограничения на пространственное разрешение и дифракционную эффективность. Для преодоления данного явления предложено использовать композитную голограмму, в виде объемной фазовой решетки, разделенной на зоны с независимо изменяющимися параметрами наклона полос, формой и толщиной голографического слоя, а также глубиной модуляции. **Метод.** Предложен алгоритм, который позволяет проводить трассировку луча через голограмму, записанную двумя точечными когерентными источниками при помощи вспомогательного асферического зеркала. Первоначальная трассировка луча в схеме записи голограммы выполнена с использованием минимизации функции ошибок методами покоординатного спуска и золотого сечения. На основе полученных результатов с помощью уравнения Уэлфорда вычислены направляющие векторы дифрагированного луча. С использованием результатов трассировки на базе теории связанных волн Когельника определена дифракционная эффективность голограммы. Предложенные алгоритмы реализованы в среде Zemax Optics Studio. **Основные результаты.** Применение представленного композитного голограммного элемента и средств моделирования его работы показаны на примере дисплея, работающего в диапазоне 510–530 нм с полем зрения $7^{\circ}36' \times 5^{\circ}48'$ и диаметром выходного зрачка 8 мм. Предложенные решения позволили повысить дифракционную эффективность в 3,45 раза, а пространственное разрешение на 12,7 %, которое варьируется по полю зрения в пределах $0'44''$ – $1'6''$. **Практическая значимость.** Применение композитных голограмм позволит создавать дисплеи, отличающиеся более высоким пространственным разрешением и яркостью проецируемого изображения, а также равномерностью характеристик по полю зрения.

Ключевые слова

дифракционная эффективность, объемно-фазовая голограмма, волноводный голографический дисплей, дополненная реальность

Благодарности

Работа выполнена при финансовой поддержке гранта РНФ № 21-79-00082. Особую благодарность за проделанную работу хотим высказать Илье Андреевичу Гускову.

Ссылка для цитирования: Харитонов Д.Ю., Ахметов Д.М., Муслимов Э.Р., Гильфанов А.Р., Павлычева Н.К. Моделирование композитного волноводного голографического дисплея // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1037–1047. doi: 10.17586/2226-1494-2022-22-6-1037-1047

© Харитонов Д.Ю., Ахметов Д.М., Муслимов Э.Р., Гильфанов А.Р., Павлычева Н.К., 2022

Modelling of a composite waveguide holographic display

Danila Yu. Kharitonov^{1✉}, Damir M. Akhmetov², Eduard R. Muslimov³,
Aydar R. Gilfanov⁴, Nadezhda K. Pavlycheva⁵

^{1,2,4} OJSC “Scientific and Production Association” State Institute of the Applied Optics, Kazan, 420075, Republic of Tatarstan

^{3,5} Kazan National Research Technical University named after A.N. Tupolev, Kazan, 420111, Republic of Tatarstan

³ NOVA Optical IR Instrumentation Group, Dwingeloo, 7991 PD, Netherlands

³ Laboratoire d’Astrophysique de Marseille, Marseille, 13388, France

¹ danila-kharitonov2017@mail.ru✉, <https://orcid.org/0000-0002-1256-4474>

² akhmetov.damir.97@mail.ru, <https://orcid.org/0000-0001-5827-2648>

³ e0123@mail.ru, <https://orcid.org/0000-0002-3242-9894>

⁴ cataklizm@inbox.ru, <https://orcid.org/0000-0001-8109-5029>

⁵ npavlych@rambler.ru, <https://orcid.org/0000-0001-9395-3967>

Abstract

Optical designs of waveguide-type augmented reality displays are investigated. Displays based on volume phase holograms are notable for their small size, large exit pupil and high transmittance both in the projected image channel and in the direct vision channel. However, with an increase of the aperture, field of view and working spectral range, the spread of the values of the beam angle of incidence and the wavelength increases when solving the diffraction problem at different points on the hologram surface which imposes restrictions on spatial resolution and diffraction efficiency. To overcome this phenomenon, it is proposed to use a composite hologram which represents a volume phase grating divided into zones with independently varying parameters of the fringes tilt, their shape, the holographic layer thickness and the refraction index modulation depth. We propose an algorithm that allows ray tracing through a hologram recorded by two coherent point sources using an auxiliary aspherical mirror. The initial ray tracing in the hologram recording scheme is performed using the error function minimization by the orthogonal descent and golden section methods. Based on the results obtained, the directional vectors of the diffracted beam are calculated using the Welford equation. Using the tracing results, the hologram diffraction efficiency is computed with the Kogelnik’s coupled wave theory. The proposed algorithms are implemented in the Zemax Optics Studio software. The application of the proposed composite hologram element and the tools for operation modeling are shown on an example of display operating in the range of 510–530 nm with the field of view of $7^{\circ}36' \times 5^{\circ}48'$ and the exit pupil diameter of 8 mm. It is shown that the proposed solutions make it possible to increase the diffraction efficiency by 3.45 times. At the same time, the spatial resolution increases by 12.7 % varying across the field of view in the range of $0'44''$ – $1'6''$. The use of composite holograms allows one to create displays with higher spatial resolution and brightness of the projected image as well as uniformity of the characteristics across the field of view.

Keywords

diffraction efficiency, volume phase hologram, waveguide holographic display, augmented reality

Acknowledgements

The research is funded by the Russian Science Foundation grant No. 21-79-00082. We also would like to thank our colleague Ilya A. Guskov for his help.

For citation: Kharitonov D.Yu., Akhmetov D.M., Muslimov E.R., Gilfanov A.R., Pavlycheva N.K. Modelling of a composite waveguide holographic display. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1037–1047 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1037-1047

Введение

Технологии дополненной реальности (Augmented Reality, AR) позволяют совмещать в поле зрения пользователя сформированное компьютером изображение и непосредственно наблюдаемую окружающую сцену. В последние годы они представляют собой быстро развивающуюся область и достигли уровня серийного производства коммерческих устройств. Существует несколько типов оптических систем, используемых в дисплеях дополненной реальности, обладающих особыми преимуществами и недостатками [1, 2].

В частности, дисплей дополненной реальности может быть построен на базе волновода с вводом и выводом излучения при помощи голограммных элементов. Такое решение отличается компактностью, относительной простотой конструкции, высокой яркостью и разрешением проецируемого изображения в ограниченном поле зрения, а также возможностью реализации многократного вывода пучка для увеличе-

ния области согласования с глазом наблюдателя [3–5]. Однако при увеличении поля зрения, апертуры и рабочего спектрального диапазона условия восстановления голограммы значительно изменяются, что приводит к неравномерному распределению яркости и разрешения проецируемого изображения.

В настоящей работе предложено скомпенсировать отмеченные недостатки за счет использования композитной голограммной решетки (КГР), которая представляет собой голограмму, записываемую путем стыковки нескольких зон. При этом в каждой такой зоне изменяются характеристики, такие как дифракционная эффективность, поляризационные, дисперсионные и абберационные свойства. Они варьируются за счет изменения углов падения в схеме записи, аббераций записывающих волновых фронтов и времени экспозиции, которые могут контролироваться локально [6, 7]. При этом расчет и моделирование оптического элемента представляет собой сравнительно сложную вычислительную задачу. Для определения абберационных

свойств КГР необходимо провести трассировку лучей в схеме записи и затем решить уравнение трассировки для каждой точки записанной голограммы. Для расчета дифракционной эффективности (ДЭ) голограммы требуется задать параметры ее структуры, также провести трассировку лучей в схемах записи и работы, и далее рассчитать эффективность дифракции для заданной поляризации. Все указанные вычисления необходимо повторить для каждой зоны КГР со своим набором параметров. В случае, если необходимо провести численную оптимизацию оптической системы, вычисления многократно повторяются в цикле.

Вычисления могут быть проведены существующими средствами моделирования, реализованными в программных пакетах, например Zemax Optics Studio™. Однако в таком случае одна оптическая схема должна быть представлена несколькими моделями — схемами записи и работы, и моделью, учитывающей структуру голограммы для каждой зоны. При этом необходимо дополнительно проверять соответствие моделей друг другу, в частности, справедливость допущения о незначительном влиянии aberrаций высших порядков в схеме записи на распределение ДЭ.

Цель — упрощение процесса моделирования КГР за счет унификации подходов и введения ряда упрощений.

Алгоритм трассировки лучей

Предположим, что при записи КГР использованы вспомогательные оптические элементы для формирования aberrированного волнового фронта. На практике наиболее вероятно использование одного деформируемого зеркала, устанавливаемого в одно из плеч интерферометра записи. Рассмотрим запись и работу голограммной решетки в таком простом случае (рис. 1).

Пусть форма поверхности вспомогательного зеркала 3 представлена уравнением

$$Z_L = \frac{1/R(X_L^2 + Y_L^2)}{1 + \sqrt{1 - (1+k) \frac{X_L^2 + Y_L^2}{R^2}}} + \sum_{p=1}^m a_{2p}(X_L^2 + Y_L^2)^p, \quad (1)$$

где R — радиус кривизны поверхности в вершине; k — коническая постоянная; a, p, m — коэффициент, порядок и число коэффициентов асферичности.

Из уравнения (1) определим нормаль к поверхности в произвольной точке в локальной системе координат зеркала:

$$\mathbf{N} = \frac{\left(\frac{\partial Z_L}{\partial X_L}, \frac{\partial Z_L}{\partial Y_L}, -1 \right)}{\sqrt{\left(\frac{\partial Z_L}{\partial X_L} \right)^2 + \left(\frac{\partial Z_L}{\partial Y_L} \right)^2 + 1}}. \quad (2)$$

Перейдем от локальной (2) к глобальной системе координат, связанной с подложкой 4 записываемой решетки:

$$X_A = -X_L, \quad (3)$$

$$Y_A = -Z_L \sin(i_{22} - i_{21}) + Y_L \cos(i_{22} - i_{21}) + d_{21} \sin(i_{21}), \quad (4)$$

$$Z_A = Z_L \cos(i_{22} - i_{21}) + Y_L \sin(i_{22} - i_{21}) + d_{21} \cos(i_{21}). \quad (5)$$

Координаты вектора нормали преобразуем аналогично уравнениям (3)–(5). Координаты точечных источников записи 1 и 2 имеют вид:

$$X_{r1} = 0, Y_{r1} = d_1 \cos(i_1), Z_{r1} = -d_1 \sin(i_1), \quad (6)$$

$$X_{r2} = 0,$$

$$Y_{r2} = d_{21} \sin(i_{21}) + d_{22} \sin(2i_{22} - i_{21}), \quad (7)$$

$$Y_{r2} = d_{21} \cos(i_{21}) - d_{22} \cos(2i_{22} - i_{21}).$$

Рассчитаем направляющие векторы лучей в схеме, используя уравнения (6), (7):

$$\mathbf{V}_{22} = \frac{(X_A - X_{r2}, Y_A - Y_{r2}, Z_A - Z_{r2})}{\sqrt{(X_A - X_{r2})^2 + (Y_A - Y_{r2})^2 + (Z_A - Z_{r2})^2}}, \quad (8)$$

$$\mathbf{V}_{21} = \frac{(X - X_A, Y - Y_A, Z - Z_A)}{\sqrt{(X - X_A)^2 + (Y - Y_A)^2 + (Z - Z_A)^2}}, \quad (9)$$

$$\mathbf{V}_1 = \frac{(X - X_{r1}, Y - Y_{r1}, Z - Z_{r1})}{\sqrt{(X - X_{r1})^2 + (Y - Y_{r1})^2 + (Z - Z_{r1})^2}}. \quad (10)$$

Предположим, что для вспомогательного зеркала 3 должно выполняться уравнение отражения в векторной форме. Однако при данной постановке задачи остается неизвестной промежуточная точка — точка падения на вспомогательное зеркало. Она может быть найдена численно как ноль функции ошибок, описывающей длину вектора отклонения от условия отражения:

$$F_{err}(X_L, Y_L) = |\mathbf{N} \times \mathbf{V}_{21} - \mathbf{N} \times \mathbf{V}_{22}|. \quad (11)$$

Если локальные координаты точки падения найдены, то перейдем к глобальным координатам (3)–(5) и запишем направляющие векторы (8)–(10). Полученные значения подставим в уравнение Уэлфорда [8] и рассчитаем направление дифрагированного луча в заданной точке падения на голограмму:

$$\mathbf{N}_S \times (\mathbf{V}_i - \mathbf{V}_d) = Q \frac{\lambda}{\lambda_0} \mathbf{N}_S \times (\mathbf{V}_1 - \mathbf{V}_{21}), \quad (12)$$

где Q — порядок дифракции; λ — рабочая длина волны; λ_0 — длина волны записи; $\mathbf{N}_S$ — вектор нормали в точке падения.

Отметим, что при таком подходе возможно значительно сократить число переменных, описывающих схему записи. Вычисления для схемы работы и записи в таком случае могут быть выполнены в одной модели. Тогда задача сводится к численному решению уравнения (11). При использовании уравнения (1) не возникло особых точек, разрывов или множественных локальных минимумов, и даже простые численные методы оказались достаточно эффективными. Например, уравнение (11) решим сочетанием методов золотого сечения и покоординатного спуска [9, 10]. На рис. 2 показаны типичные графики сходимости при трассировке луча с использованием этих методов. Видно, что в пределах 10 шагов можно достичь точности в 1 мкм.

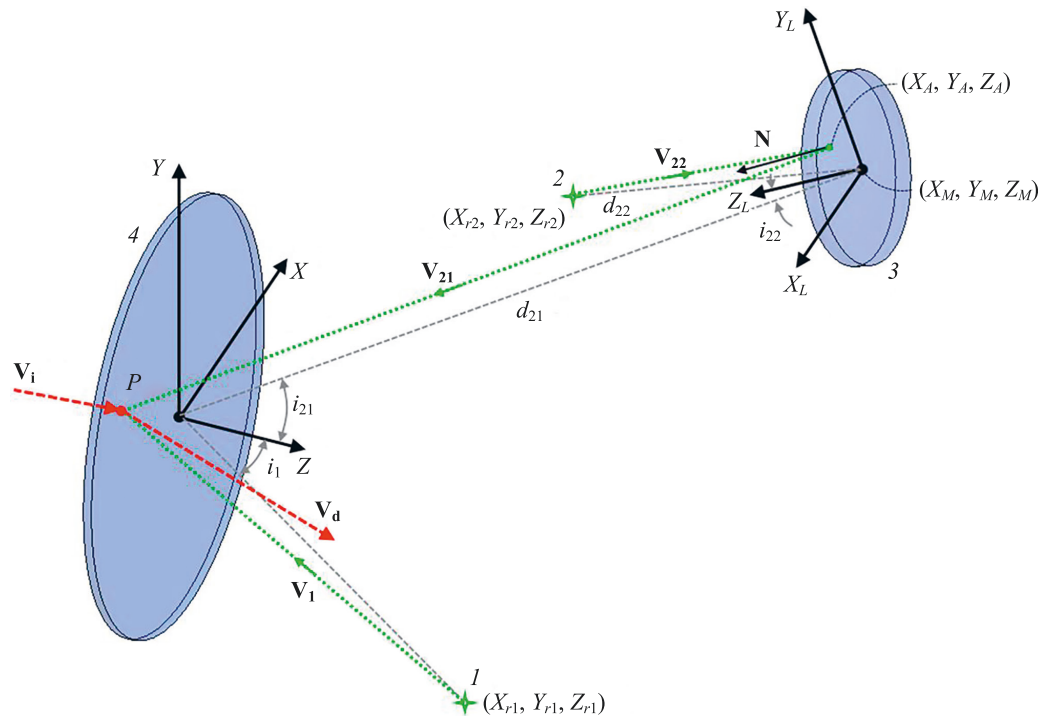


Рис. 1. Схема записи голограммы при помощи вспомогательного зеркала: 1 и 2 — точечные источники записи; 3 — вспомогательное зеркало; 4 — подложка; i_1, i_{21}, i_{22} — углы падения в схеме записи; P — точка на подложке 4; V_i — вектор падающего на подложку луча в схеме работы; V_d — вектор дифрагированного луча в схеме работы; X, Y, Z — декартовы координаты точки падения на поверхности зеркала; индекс L — локальная система координат, связанная с вершиной вспомогательного зеркала; (X_{r2}, Y_{r2}, Z_{r2}) — координаты точечного источника записи 2; (X_{r1}, Y_{r1}, Z_{r1}) — координаты точечного источника записи 1; N — вектор нормали на поверхности вспомогательного зеркала 3; (X_A, Y_A, Z_A) — координаты точки на поверхности вспомогательного зеркала 3; (X_M, Y_M, Z_M) — глобальные координаты начала локальной системы координат

Fig. 1. Scheme of a hologram recording using an auxiliary mirror: 1 and 2 — recording point sources; 3 — auxiliary mirror; 4 — substrate; i_1, i_{21}, i_{22} — angles of incidence in the recording scheme, P — a point on the substrate 4, V_i — the vector of a ray incident on the substrate 4 in the operation scheme, V_d — the diffracted beam vector in the operation scheme, X, Y, Z — cartesian coordinates of the point of incidence on the mirror surface, L index — the local coordinate system associated with the vertex of the auxiliary mirror, (X_{r2}, Y_{r2}, Z_{r2}) — coordinates of the recording point source 2, (X_{r1}, Y_{r1}, Z_{r1}) — coordinates of the recording point source 1, N — the normal vector on the auxiliary mirror surface 3, (X_A, Y_A, Z_A) — coordinates of a point on the surface of the auxiliary mirror 3, (X_M, Y_M, Z_M) — global coordinates of the local coordinate system origin

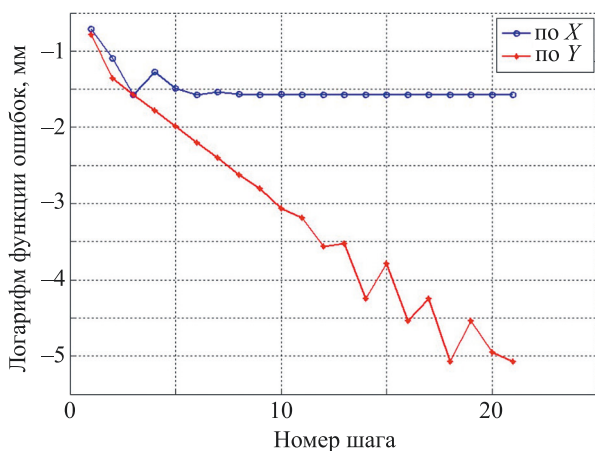


Рис. 2. График численного решения уравнения трассировки луча через вспомогательное зеркало

Fig. 2. Plot of the numerical solution of the raytracing equation through the auxiliary mirror

Алгоритм расчета дифракционной эффективности

В настоящей работе рассмотрены объемно-фазовые голограммные элементы. Рассчитаем ДЭ такого элемента с помощью уравнений теории связанных волн Когельника [11]. Предположим, что толщина структуры голограммы велика по сравнению с ее периодом и длиной волны излучения. При этом эффект конической дифракции (т. е. дифракции в случае, когда падающий луч образует некоторый угол с плоскостью, заданной вектором решетки и ее нормалью [12]) сводится к построению проекции падающего луча на плоскость записи, а дифрагированный луч трассируется только в +1 порядке дифракции. В области точки падения каждого луча на голограмму элементарная решетка (рис. 3) восстановится. При этом такие параметры решетки, как угол наклона полос ϕ и их частота N_{gr} , определим с помощью результатов трассировки в схеме записи (8)–(9):

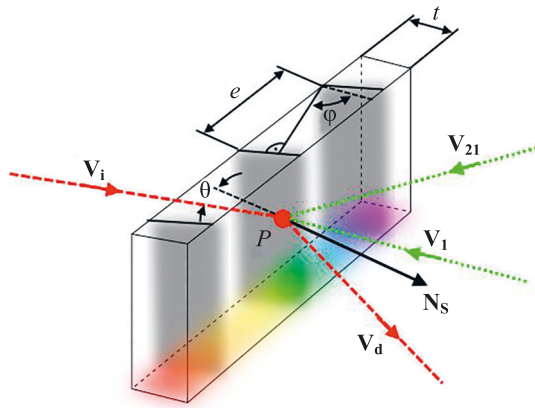


Рис. 3. Определение параметров элементарной решетки для расчета дифракционной эффективности: $e = 1/N_{gr}$ — период элементарной решетки; t — толщина структуры голограммы; φ — угол наклона полосы; θ — угол падения излучения в слое

Fig. 3. Definition of the elementary grating parameters for the diffraction efficiency calculation: $e = 1/N_{gr}$ — fringes period, t — hologram structure thickness, φ — fringes tilt angle, θ — angle of incidence inside the hologram layer

$$\varphi = 90^\circ - \frac{\arcsin\left(\frac{\sin(\alpha_1)}{n}\right) + \arcsin\left(\frac{\sin(\alpha_2)}{n}\right)}{2}, \quad (13)$$

$$N_{gr} = \frac{\sin(\alpha_1) - \sin(\alpha_2)}{\lambda_0}, \quad (14)$$

где α_1 и α_2 — углы падения излучения из воздуха в схеме записи; n — средний показатель преломления светочувствительного материала,

Предположим, что задача конической дифракции в этом случае может быть сведена к построению проекции падающего луча на плоскость записи. Тогда, если остальные предположения теории Когельника верны, можно вычислить дифракционную эффективность для двух состояний поляризации η_{TE} и η_{TM} и далее определить эффективность для неполяризованного излучения η .

$$\eta_{TE(TM)} = \frac{\sin^2 \sqrt{v_{TE(TM)}^2 + \xi^2}}{1 + (\xi/v_{TE(TM)})^2}, \quad (15)$$

$$\eta = \sqrt{\frac{(\eta_{TE}^2 + \eta_{TM}^2)}{2}}, \quad (16)$$

где TM и TE — состояния поляризации; v и ξ — параметры, определяемые как TM и TE .

$$v_{TE} = \frac{\pi \Delta n t}{\lambda \sqrt{C_R C_S}}, \quad v_{TM} = \frac{\pi \Delta n t}{\lambda \sqrt{C_R C_S}} \cos(2(\theta - \varphi)), \quad (17)$$

$$\xi = \frac{\chi t}{2C_S}, \quad \chi = K \cos(\theta - \varphi) - \frac{K^2}{4\pi n} \lambda, \quad (18)$$

$$C_R = \cos(\varphi), \quad C_S = \cos(\varphi) - \frac{K\lambda}{2\pi n} \cos(\theta),$$

где χ — коэффициент отклонения фазы; φ — угол наклона полос, определяемый векторами записи; C_R и

C_S — коэффициенты наклона (индексы R и S соответствуют опорной и объектной волнам; Δn — амплитуда модуляции показателя преломления; θ — угол падения излучения в слое; λ — рабочая длина волны; K — коэффициент поляризации

Расчет ДЭ по (15)–(16) требует задания толщины голограммы и амплитуды модуляции. Остальные параметры определены по формулам (1)–(12). Это позволяет объединить две модели в одну.

Расчеты по формулам (13)–(18) выполнены в библиотеке dll, задающей пользовательский тип поверхности в среде Zemax Optics Studio™. При этом параметры, показанные на рис. 1 и рис. 3 задает пользователь, направляющий вектор дифрагированного луча определен уравнением (12), а пропускание поверхности уравнением (16). Отметим, что описанные алгоритмы реализованы при трассировке каждого луча в последовательном режиме. Это означает, что расчет может быть проведен для голограммы на неплоской подложке, но без учета возможного многократного прохождения луча через поверхность.

Пример расчета оптической системы

Для демонстрации предложенного подхода рассмотрим оптическую систему волноводного дисплея со следующими типичными характеристиками [13–16]: спектральный диапазон 510–530 нм, угловое поле зрения $7^\circ 36' \times 5^\circ 48'$, диаметр выходного зрачка 8 мм, удаление выходного зрачка 60 мм. Общий вид оптической схемы показан на рис. 4. Пучки от излучающей матрицы коллимируются асферической пластиковой линзой с фокусным расстоянием 60 мм и вводятся в плоскопараллельную волноводную пластину через пропускающую голограммную решетку. Решетка имеет пространственную частоту 1920 мм^{-1} и разделена на три зоны с независимо оптимизируемыми параметрами. Волновод выполнен из стекла К8, имеет толщину 1 мм и установлен под углом 5° к падающему пучку.

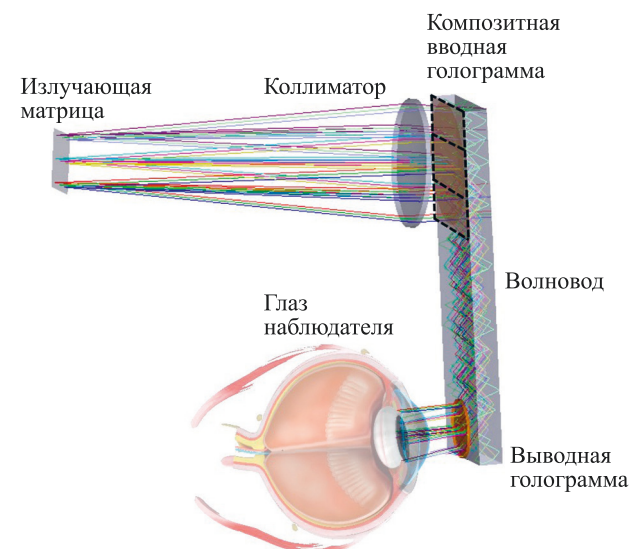


Рис. 4. Общий вид оптической схемы волноводного дисплея
Fig. 4. General view of the waveguide display optical design

Отметим, что наличие небольшого наклона позволяет увеличить зазоры для установки узла излучающей матрицы и снизить пространственную частоту голограммы. Пучок распространяется по волноводу за счет полного внутреннего отражения и выводится через пропускающую голограмму с той же пространственной частотой.

Коллиматор представляет собой одиночную линзу, выполненную из полиметилметакрилата. Первая поверхность является асферической с вершинным радиусом $R_1 = 48,491$ мм, конической постоянной $k = -14,539$ и коэффициентами асферичности $\alpha_4 = 1,143 \cdot 10^{-5}$, $\alpha_6 = -2,110 \cdot 10^{-8}$, вторая поверхность — сферической с радиусом $R_2 = -75,246$ мм. Подобные элементы достаточно типичны для крупносерийных оптических устройств и могут изготавливаться с помощью прецизионного моллирования [17]. Коллиматор позволяет получить коллимированный пучок хорошего качества для монохроматического излучения, но при работе с протяженным полем зрения обладает остаточными aberrациями. На рис. 5, а показано распределение среднеквадратического радиуса точечной диаграммы для коллиматора в плоскости наилучшей установки при расчете в обратном ходе. КГР позволяет ввести дополнительную коррекцию таких остаточных aberrаций. Отметим, что использование однолинзового коллиматора является предельным случаем, позволяющим более наглядно продемонстрировать особенности использования КГР.

Другой эффект, который следует учесть при расчете КГР в подобной схеме, заключается в изменении условий восстановления для разных точек голограммы. На рис. 5, б показано изменение угла падения главного луча на световой зоне голограммы. Очевидно, что такое изменение приведет к локальным изменениям ДЭ и, как следствие, неравномерной яркости проецируемого изображения. Оптимизация параметров структуры голограммы для отдельных зон позволяет уменьшить данный нежелательный эффект.

Для выполнения оптимизации определим следующие условия: выходной зрачок системы совпадает со зрачком глаза, голограмма не совпадает со зрачком коллимационной системы, пучки для разных точек поля зрения разделены в плоскости голограммы. Данные условия позволяют корректировать ДЭ и aberrации для разных точек поля за счет разбиения голограммы на зоны и последующей независимой оптимизации параметров отдельных зон КГР.

На рис. 5, с показаны следы пучков для разных точек зрения на световой зоне вводной голограммы.

Отметим, что известны другие решения, также использующие для повышения характеристик волноводного дисплея изменение параметров структуры голограммного элемента по его площадке. Например, в дисплее DigiLens [18] использована решетка с постоянным периодом и непрерывно изменяющимся наклоном полос, позволяющая компенсировать изменение ДЭ в зависимости от угла падения. В работе [19] предложен вариант с разбиением выводного дифракционного элемента на множество дискретных элементов с независимо изменяющимися значениями периода и толщины структуры. В настоящей работе представлено решение применения небольшого числа дискретных зон, в каждой из которых оптимизированы профиль полос голограммы, определяющий ДЭ и их картина полос, задающая aberrационные свойства. Подход реализован на базе существующей технологии записи физических голограмм и является более гибким и технологичным, особенно в случае изготовления единичного прототипа или малой партии элементов.

Алгоритм оптимизации рассматриваемой оптической схемы аналогичен описанному в работе [6]. Уравнение поверхности зеркала (1) ограничено 4-м порядком. Схема записи вводной КГР на длине волны 532 нм, полученная в результате, показана на рис. 6. Параметры, описывающие каждую из зон, сведены в табл. 1.

Измерение дифракционной эффективности выполнено по стандартным кривым пропускания для трех

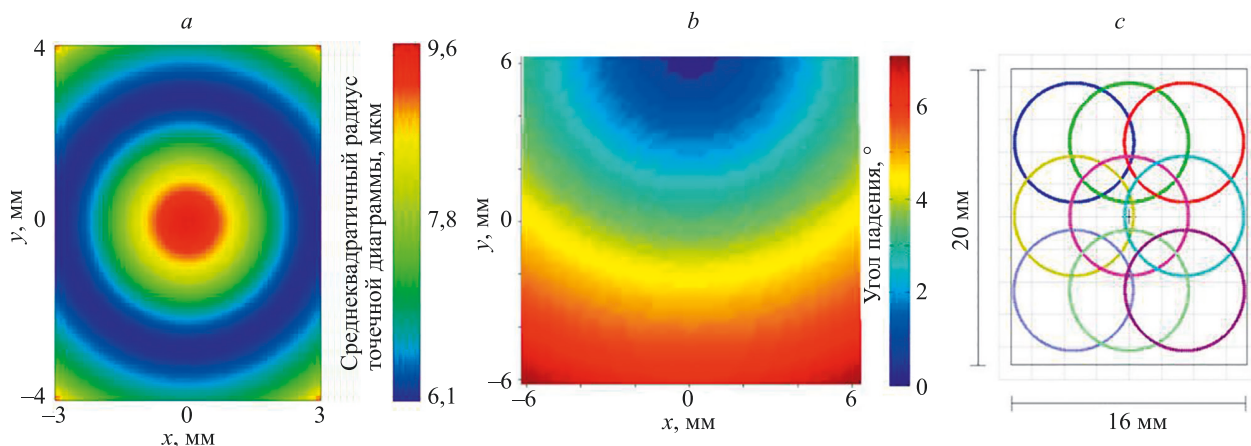


Рис. 5. Эффекты, компенсируемые при помощи композитной голограммной решетки: распределение остаточных aberrаций коллимирующей линзы по полю зрения (а); изменение угла падения главного луча на вводную голограмму по ее световой площадке (б); диаграмма заполнения световой зоны голограммы (с)

Fig. 5. The effects compensated with the composite hologram grating: the distribution of residual aberrations of the collimating lens over the field of view (a); the change in the angle of incidence of the chief ray on the incoupling hologram across its clear aperture (b); filling diagram of the hologram clear aperture (c)

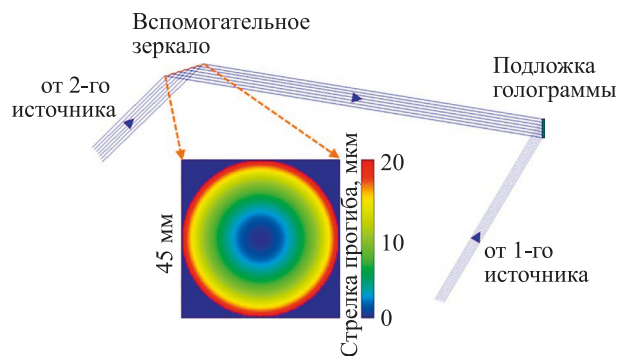


Рис. 6. Схема записи вводной голограммной решетки
Fig. 6. The recording scheme of the incoupling hologram grating

контрольных точек поля зрения (рис. 7). На графиках показано распределение ДЭ в меридиональном сечении. Видно, что значения ДЭ остаются практически неизменными по сагиттальному сечению зрачка. Использование КГР с оптимизированной структурой позволило повысить ДЭ в 1,08–3,45 раз по сравнению с исходным вариантом, основанном на выполнении условия Брэгга, а также повысить ее равномерность.

Выполним оценку влияния изначальных допущений и упрощений. Для этого сравним результаты, полученные с помощью теории Когельника и с использованием численного метода строгого анализа связанных волн

(Rigorous Coupled Wave Analysis, RCWA), реализованного в программе *reticolo* [20]. Результаты расчета ДЭ для главного луча пучка, проходящего через центр вводной голограммы для данных методов представлены на рис. 8. Видно, что использование упрощенной аналитической теории приводит к небольшой ошибке в положении максимума кривой ДЭ. Эффективность порядков дифракции, не учитываемых в расчете, не превосходит 8 %. При введении в расчет угла конической дифракции $\delta = 10^\circ$, превосходящего угловое поле зрения системы, изменение ДЭ не превышает 2,8 %, что можно считать незначительным отклонением.

В данном случае предположение о том, что толщина структуры решетки велика в сравнении с ее периодом и рабочей длиной волны, выполняется не строго, и голограмма работает в промежуточном режиме. Однако, как указано в [21], для подобных случаев использование теории связанных волн остается предпочтительным в сравнении с теорией Рамана–Ната [22]. Как показал выполненный анализ, вносимая данным упрощением ошибка при вычислении ДЭ невелика.

Для оценки качества изображения применены точечные диаграммы (рис. 9). Использование КГР, записываемой с помощью вспомогательного зеркала, позволило уменьшить среднеквадратический угловой размер диаграмм с $0'49''$ – $1'14''$ до $0'44''$ – $1'6''$.

Отметим, что в случае записи вводной голограммы с помощью вспомогательного зеркала изменяются требования к точности изготовления и позиционирования

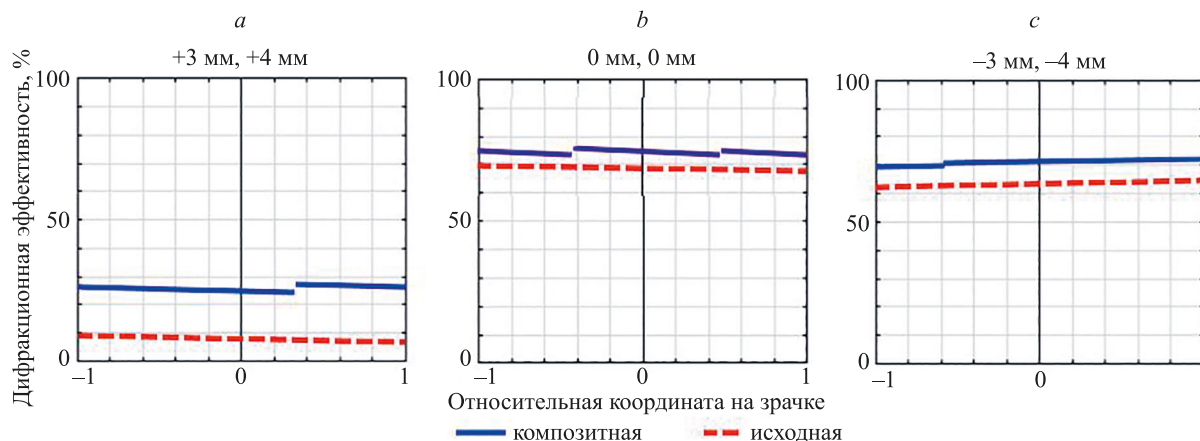


Рис. 7. Пространственное распределение дифракционной эффективности в поле зрения с различными линейными координатами: верхний правый угол (3; 4) мм (a); центр (0; 0) мм (b); нижний левый угол (–3; –4) мм (c)

Fig. 7. Spatial distribution of diffraction efficiency in the field of view with different linear coordinates: upper right corner (3; 4) mm (a), center (0; 0) mm (b), lower left corner (–3; –4) mm (c)

Таблица 1. Параметры композитной голограммной решетки

Table 1. Parameters of composite hologram grating

Параметр	Зона		
	Верхняя	Средняя	Нижняя
Угол падения на вспомогательное зеркало	65°6′	63°12′	63°36′
Стрелка прогиба вспомогательного зеркала, мкм	19,9	18,3	19,1
Толщина голограммы, мкм	3,95	3,74	3,73
Амплитуда модуляции	0,062	0,066	0,069

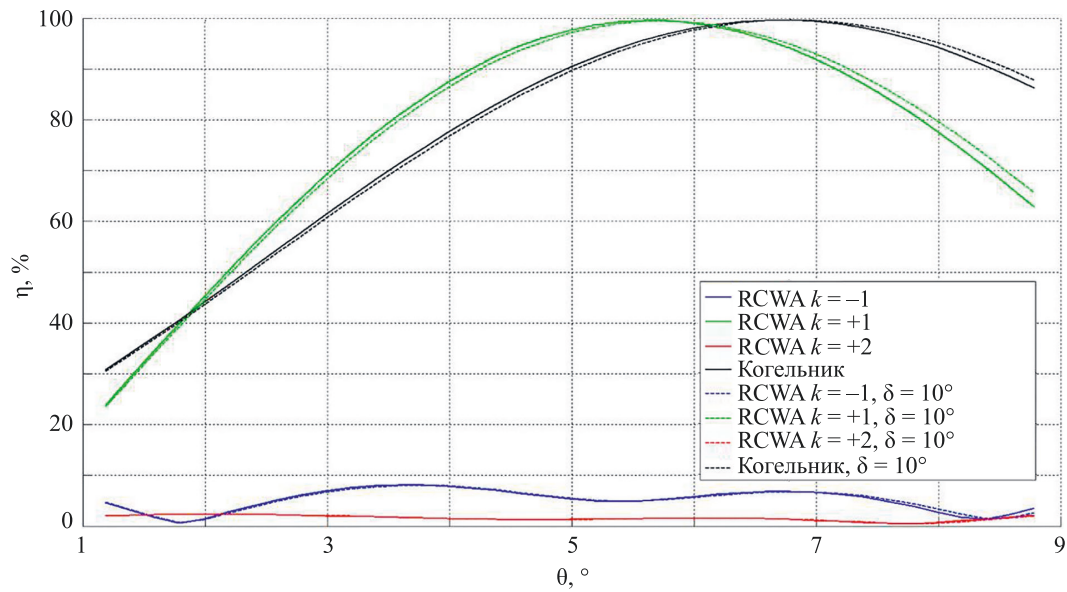


Рис. 8. Сравнение дифракционной эффективности, рассчитанной по аналитическим соотношениям теории Когельника и методом RCWA

Fig. 8. Comparison of the diffraction efficiency calculated by the analytical equations of the Kogelnik theory and the RCWA method

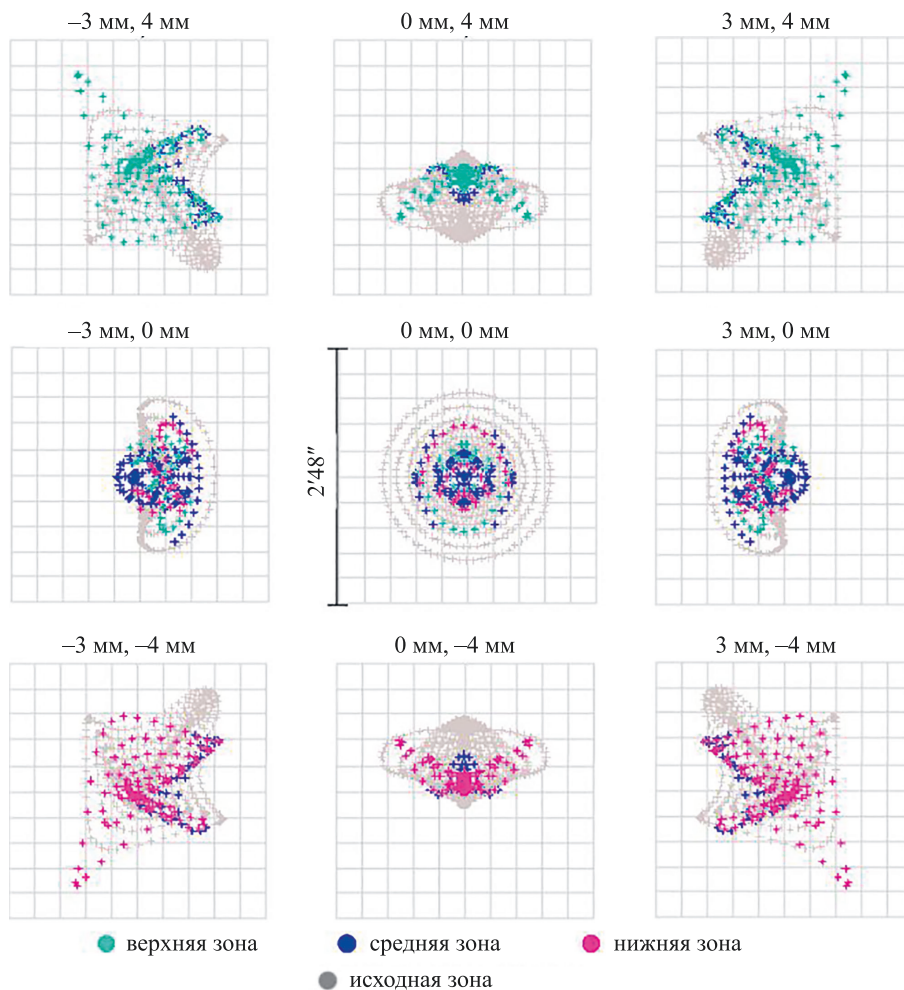


Рис. 9. Точечные диаграммы волноводного дисплея для девяти контрольных точек поля зрения с координатами (x; y) мм в плоскости предметов

Fig. 9. Spot diagrams of the waveguide display for 9 control points across the field of view with coordinates (x; y) mm in the plane of objects

Таблица 2. Изменение чувствительности системы к ошибкам позиционирования

Table 2. Changing the system sensitivity to positioning errors

Номер	Параметр	Исходная схема	Схема с КГР
1	Расстояние матрица-коллиматор, угл. мин/мм	1,164	0,902
2	Расстояние коллиматор-голограмма, угл. мин/мм $\times 10^{-13}$	1,17	0,001
3	Децентрировка матрица-коллиматор по X, угл. мин/мм	0,007	0,022
4	Децентрировка матрица-коллиматор по Y, угл. мин/мм	0,012	0,046
5	Разворот матрица-коллиматор по X, угл. мин/град.	0,030	0,096
6	Разворот матрица-коллиматор по Y, угл. мин/град.	0,022	0,050
7	Децентрировка коллиматор-голограмма по X, угл. мин/мм	$1,14 \cdot 10^{-13}$	$3,02 \cdot 10^{-4}$
8	Децентрировка коллиматор-голограмма по Y, угл. мин/мм	$9,78 \cdot 10^{-14}$	0,015
9	Разворот коллиматор-голограмма по X, угл. мин/град.	$1,12 \cdot 10^{-5}$	$1,08 \cdot 10^{-4}$
10	Разворот коллиматор-голограмма по Y, угл. мин/град.	$1,03 \cdot 10^{-5}$	$9,19 \cdot 10^{-6}$

элементов оптической схемы. Подробный анализ влияния КГР на чувствительность оптических систем, в которых она используется, к ошибкам изготовления и сборки потребует отдельного исследования. Приведем сравнение чувствительности схемы к ошибкам позиционирования элементов, общих для исходной и модифицированной схем — светоизлучающей матрицы и коллимирующей линзы. Расчет проведен с помощью стандартных инструментов Zemax Optics Studio™. В качестве критерия качества изображения использован среднеквадратический радиус точечной диаграммы — среднее значение по полю зрения. Значения чувствительности, полученные для изолированных параметров без использования компенсаторов и выраженные как отношение изменения критерия к изменению конструктивного параметра, приведены в табл. 2.

Величины чувствительности, относящиеся к положению плоской дифракционной решетки с прямыми эквидистантными штрихами в параллельном пучке, т. е. ее линейным смещениям относительно коллиматора (параметры 2, 7 и 8, табл. 2), практически равны нулю. Как и следовало ожидать, продольное и поперечное смещения такой классической решетки никак не сказывается на выбранных критериях качества изображения. Поскольку КГР записывается aberrированными пучками, штрихи решетки не являются строго прямолинейными и эквидистантными, а схема с ней не строго ин-

вариантна к положению решетки в параллельном пучке. Приведенные результаты подтвердили значительный относительный рост чувствительности схемы к положению дисплея и коллиматора. При этом абсолютные значения чувствительности остаются весьма низкими. Следовательно, можно считать, что использование более сложного голограммного элемента в данном случае не приведет к заметному ужесточению требований к изготовлению и сборке.

Закключение

Представлен подход к моделированию оптических систем с композитными голограммными элементами. При наличии ряда допущений к схеме записи голограммного элемента можно выполнить расчет и оптимизацию его aberrаций, а также дифракционной эффективности в единой модели. Реализованы предложенные алгоритмы в программном продукте Zemax Optics Studio™.

На примере компактного монохромного дисплея дополненной реальности с полем зрения $7^{\circ}36' \times 5^{\circ}48'$ и выходным зрачком 8 мм продемонстрирован процесс расчета и анализа работы композитной голограммы. В частности, показаны повышения дифракционной эффективности до 3,45 раз и пространственного разрешения на 12,7 %.

Литература

1. Cakmakci O., Rolland J.P. Head-worn displays: A review // *Journal of Display Technology*. 2016. V. 2. N 3. P. 199–217. <https://doi.org/10.1109/JDT.2006.879846>
2. Kress B.C. Optical waveguide combiners for AR headsets: features and limitations // *Proceedings of SPIE*. 2019. V. 11062. P. 110620J. <https://doi.org/10.1117/12.2527680>
3. Amitai Y., Friesem A.A., Weiss V. Holographic elements with high efficiency and low aberrations for helmet displays // *Applied Optics*. 1989. V. 28. N 16. P. 3405–3416. <https://doi.org/10.1364/AO.28.003405>
4. Yoshida T., Tokuyama K., Takai Y., Tsukuda D., Kaneko T., Suzuki N., Anzai T., Yoshikaie A., Akutsu K., Machida A. A plastic holographic waveguide combiner for light-weight and highly-transparent augmented reality glasses // *Journal of the Society for*

References

1. Cakmakci O., Rolland J.P. Head-worn displays: A review. *Journal of Display Technology*, 2016, vol. 2, no. 3, pp. 199–217. <https://doi.org/10.1109/JDT.2006.879846>
2. Kress B.C. Optical waveguide combiners for AR headsets: features and limitations. *Proceedings of SPIE*, 2019, vol. 11062, pp. 110620J. <https://doi.org/10.1117/12.2527680>
3. Amitai Y., Friesem A.A., Weiss V. Holographic elements with high efficiency and low aberrations for helmet displays. *Applied Optics*, 1989, vol. 28, no. 16, pp. 3405–3416. <https://doi.org/10.1364/AO.28.003405>
4. Yoshida T., Tokuyama K., Takai Y., Tsukuda D., Kaneko T., Suzuki N., Anzai T., Yoshikaie A., Akutsu K., Machida A. A plastic holographic waveguide combiner for light-weight and highly-transparent augmented reality glasses. *Journal of the Society for*

- Information Display. 2018. V. 26. N 5. P. 280–286. <https://doi.org/10.1002/jsid.659>
5. Mukawa H., Akutsu K., Matsumura I., Nakano S., Yoshida T., Kuwahara M., Aiki K. A full-color eyewear display using planar waveguides with reflection volume holograms // *Journal of the Society for Information Display*. 2009. V. 17. N 3. P. 185–193. <https://doi.org/10.1889/JSID17.3.185>
6. Муслимов Э.Р., Павлычева Н.К., Гуськов И.А. Концепция композитных голограммных оптических элементов // *Фотоника*. 2020. Т. 14. № 7. С. 586–599. <https://doi.org/10.22184/1993-7296.FRos.2020.14.7.586.599>
7. Muslimov E., Pavlycheva N., Guskov I., Akhmetov D., Kharitonov D.Yu. Spectrograph with a composite holographic dispersive element // *Proceedings of SPIE*. 2021. V. 11871. P. 187198. <https://doi.org/10.1117/12.2596922>
8. Welford W. A vector raytracing equation for hologram lenses of arbitrary shape // *Optics Communications*. 1975. V. 14. N 3. P. 322–323. [https://doi.org/10.1016/0030-4018\(75\)90327-2](https://doi.org/10.1016/0030-4018(75)90327-2)
9. Гребенникова И.В. Методы оптимизации: учебное пособие. Екатеринбург: УрФУ, 2017. 148 с.
10. Гасников А.В. Современные численные методы оптимизации. Метод универсального градиентного спуска: учебное пособие. М.: МФТИ, 2018. 181 с.
11. Kogelnik H. Coupled wave analysis for thick hologram gratings // *The Bell System Technical Journal*. 1969. V. 48. N 9. P. 2909–2947. <https://doi.org/10.1002/j.1538-7305.1969.tb01198.x>
12. Moharam M.G., Grann E.B., Pommet D.A., Gaylord T.K. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings // *Journal of the Optical Society of America A: Optics and Image Science, and Vision*. 1995. V. 12. N 5. P. 1068–1076. <https://doi.org/10.1364/JOSAA.12.001068>
13. Романова Г.Э., Нгуен Н.Ш. Анализ аберраций клина как компенсационного и функционального элемента в системах дополненной и виртуальной реальности // *Научно-технический вестник информационных технологий, механики и оптики*. 2021. Т. 21. № 6. С. 808–816. <https://doi.org/10.17586/2226-1494-2021-21-6-808-816>
14. Иванюк А.А. Проектирование оптического модуля очков дополненной реальности // *Научно-технический вестник информационных технологий, механики и оптики*. 2020. Т. 20. № 5. С. 642–648. <https://doi.org/10.17586/2226-1494-2020-20-5-642-648>
15. Романова Г.Э., Корешев С.Н., Сидоренко В.С. Расчет и моделирование световодной системы дополненной реальности на базе голографических элементов // *HoloExpo 2019. XVI международная конференция по голографии и прикладным оптическим технологиям: тезисы докладов*, 2019. С. 164–167.
16. Град Я.А., Николаев В.В., Одинокоев С.Б., Соломенко А.Б. Индикатор дополненной реальности на основе световодной пластины с пропускающим ДОЭ // *Голография: наука и практика. XIV международная конференция HOLOEXPO 2017: тезисы докладов*. 2017. С. 133–137.
17. Peixoto C., Valentim P.T., Sousa P.C., Dias D., Araújo C., Pereira D., Machado C.F., Pontes A.J., Santos H., Cruz S. Injection molding of high-precision optical lenses: A review // *Precision Engineering*. 2022. V. 76. P. 29–51. <https://doi.org/10.1016/j.precisioneng.2022.02.002>
18. Waldern J.D., Grant A.J., Popovich M.M. DigiLens switchable Bragg grating waveguide optics for augmented reality applications // *Proceedings of SPIE*. 2018. V. 10676. P. 106760G. <https://doi.org/10.1117/12.2315719>
19. Ангервакс А.Е., Муравьев Н.В., Борисов В.Н., Окунь Р.А., Востриков Г.Н., Попов М.В. Волновод с сегментированными дифракционными оптическими элементами и окологлазный дисплей. Патент RU2752556C1. Бюл. 2021. № 22.
20. Lalanne P., Hugonin J.-P. High-order effective-medium theory of subwavelength gratings in classical mounting: application to volume holograms // *Journal of the Optical Society of America A: Optics and Image Science, and Vision*. 1998. V. 15. N 7. P. 1843–1851. <https://doi.org/10.1364/JOSAA.15.001843>
21. Gerritsen H.J., Thornton D.K., Bolton S.R. Application of Kogelnik's two-wave theory to deep, slanted, highly efficient, relief transmission gratings // *Applied Optics*. 1991. V. 30. N 7. P. 807–814. <https://doi.org/10.1364/AO.30.000807>
22. Moharam M.G., Young L. Criterion for Bragg and Raman-Nath diffraction regimes // *Applied Optics*. 1978. V. 17. N 11. P. 1757–1759. <https://doi.org/10.1364/AO.17.001757>
- Information Display*, 2018, vol. 26, no. 5, pp. 280–286. <https://doi.org/10.1002/jsid.659>
5. Mukawa H., Akutsu K., Matsumura I., Nakano S., Yoshida T., Kuwahara M., Aiki K. A full-color eyewear display using planar waveguides with reflection volume holograms. *Journal of the Society for Information Display*, 2009, vol. 17, no. 3, pp. 185–193. <https://doi.org/10.1889/JSID17.3.185>
6. Muslimov E.R., Pavlycheva N.K., Guskov I.A. Concept of composite holographic optical elements. *Photonics Russia*, 2020, vol. 14, no. 7, pp. 586–599. (in Russian). <https://doi.org/10.22184/1993-7296.FRos.2020.14.7.586.599>
7. Muslimov E., Pavlycheva N., Guskov I., Akhmetov D., Kharitonov D.Yu. Spectrograph with a composite holographic dispersive element. *Proceedings of SPIE*, 2021, vol. 11871, pp. 187198. <https://doi.org/10.1117/12.2596922>
8. Welford W. A vector raytracing equation for hologram lenses of arbitrary shape. *Optics Communications*, 1975, vol. 14, no. 3, pp. 322–323. [https://doi.org/10.1016/0030-4018\(75\)90327-2](https://doi.org/10.1016/0030-4018(75)90327-2)
9. Grebennikova I.V. *Optimization Methods*. Yekaterinburg, UrFU, 2017, 148 c. (in Russian)
10. Gasnikov A.V. *Modern Numerical Methods of Optimization. Universal Gradient Descent Method*. Moscow, MIPT, 2018, 181 p. (in Russian)
11. Kogelnik H. Coupled wave analysis for thick hologram gratings. *The Bell System Technical Journal*, 1969, vol. 48, no. 9, pp. 2909–2947. <https://doi.org/10.1002/j.1538-7305.1969.tb01198.x>
12. Moharam M.G., Grann E.B., Pommet D.A., Gaylord T.K. Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, 1995, vol. 12, no. 5, pp. 1068–1076. <https://doi.org/10.1364/JOSAA.12.001068>
13. Romanova G.E., Nguyen N.S. Aberration analysis of a wedge as a compensator element in augmented and virtual reality systems. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 6, pp. 808–816. (in Russian). <https://doi.org/10.17586/2226-1494-2021-21-6-808-816>
14. Ivaniuk A.A. Optical module design for augmented reality glasses. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 5, pp. 642–648. (in Russian). <https://doi.org/10.17586/2226-1494-2020-20-5-642-648>
15. Romanova G.E., Koreshev S.N., Sidorenko V.S. Calculation and modeling of the augmented reality optical-waveguide structure based on holographic elements. *HoloExpo 2019: XVI International Conference on Holography and Applied Optical Technology. Proceedings*, 2019, pp. 164–167. (in Russian)
16. Grad Ya.A., Nikolaev V.V., Odinkov S.B., Solomashenko A.B. Augmented reality indicator based on a lightguide plate with transmissive DOEs. *Holography. Science and Practice XIV International Conference HOLOEXPO 2017. Proceedings*, 2017, pp. 133–137. (in Russian)
17. Peixoto C., Valentim P.T., Sousa P.C., Dias D., Araújo C., Pereira D., Machado C.F., Pontes A.J., Santos H., Cruz S. Injection molding of high-precision optical lenses: A review. *Precision Engineering*, 2022, vol. 76, pp. 29–51. <https://doi.org/10.1016/j.precisioneng.2022.02.002>
18. Waldern J.D., Grant A.J., Popovich M.M. DigiLens switchable Bragg grating waveguide optics for augmented reality applications. *Proceedings of SPIE*, 2018, vol. 10676, pp. 106760G. <https://doi.org/10.1117/12.2315719>
19. Angervaks A.E., Muravev N.V., Borisov V.N., Okun R.A., Vostrikov G.N., Popov M.V. Waveguide with segmented diffraction optical elements and near-eye display. *Patent RU2752556C1*, 2021. (in Russian)
20. Lalanne P., Hugonin J.-P. High-order effective-medium theory of subwavelength gratings in classical mounting: application to volume holograms. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, 1998, vol. 15, no. 7, pp. 1843–1851. <https://doi.org/10.1364/JOSAA.15.001843>
21. Gerritsen H.J., Thornton D.K., Bolton S.R. Application of Kogelnik's two-wave theory to deep, slanted, highly efficient, relief transmission gratings. *Applied Optics*, 1991, vol. 30, no. 7, pp. 807–814. <https://doi.org/10.1364/AO.30.000807>
22. Moharam M.G., Young L. Criterion for Bragg and Raman-Nath diffraction regimes. *Applied Optics*, 1978, vol. 17, no. 11, pp. 1757–1759. <https://doi.org/10.1364/AO.17.001757>

Авторы

Харитонов Данила Юрьевич — инженер, Научно-производственное объединение «Государственный институт прикладной оптики», Казань, 420075, Республика Татарстан, [sc 57338964100](https://orcid.org/0000-0002-1256-4474), <https://orcid.org/0000-0002-1256-4474>, danila-haritonov2017@mail.ru

Ахметов Дамир Маратович — инженер-конструктор, Научно-производственное объединение «Государственный институт прикладной оптики», Казань, 420075, Республика Татарстан, [sc 57339030100](https://orcid.org/0000-0001-5827-2648), <https://orcid.org/0000-0001-5827-2648>, akhmetov.damir.97@mail.ru

Муслимов Эдуард Ринатович — доктор технических наук, профессор, Казанский национальный исследовательский технический университет им. А.Н. Туполева — КАИ, Казань, 420075, Республика Татарстан; инженер-оптик, Нидерландский институт радиоастрономии, Двингело, 7991 PD, Нидерланды; ассоциированный исследователь, Астрофизическая лаборатория Марселя, Марсель, 13388, Франция, [sc 55785536800](https://orcid.org/0000-0002-3242-9894), <https://orcid.org/0000-0002-3242-9894>, e0123@mail.ru

Гильфанов Айдар Рустемович — инженер-технолог, Научно-производственное объединение «Государственный институт прикладной оптики», Казань, 420075, Республика Татарстан, [sc 6701787491](https://orcid.org/0000-0001-8109-5029), <https://orcid.org/0000-0001-8109-5029>, cataklizm@inbox.ru

Павлычева Надежда Константиновна — доктор технических наук, профессор, профессор, Казанский национальный исследовательский технический университет им. А.Н. Туполева — КАИ, Казань, 420075, Республика Татарстан, [sc 6701787491](https://orcid.org/0000-0001-9395-3967), <https://orcid.org/0000-0001-9395-3967>, nkpavlych@rambler.ru

Authors

Danila Yu. Kharitonov — Engineer, OJSC “Scientific and Production Association” State Institute of the Applied Optics, Kazan, 420075, Republic of Tatarstan, [sc 57338964100](https://orcid.org/0000-0002-1256-4474), <https://orcid.org/0000-0002-1256-4474>, danila-haritonov2017@mail.ru

Damir M. Akhmetov — Design Engineer, OJSC “Scientific and Production Association” State Institute of the Applied Optics, Kazan, 420075, Republic of Tatarstan, [sc 57339030100](https://orcid.org/0000-0001-5827-2648), <https://orcid.org/0000-0001-5827-2648>, akhmetov.damir.97@mail.ru

Eduard R. Muslimov — D. Sc., Professor, Kazan National Research Technical University named after A.N. Tupolev, Kazan, 420111, Republic of Tatarstan; Optical Engineer, NOVA Optical IR Instrumentation Group, Dwingeloo, 7991 PD, Dwingeloo, The Netherlands; Associated Researcher, Laboratoire d’Astrophysique de Marseille, Marseille, 13388, France, [sc 55785536800](https://orcid.org/0000-0002-3242-9894), <https://orcid.org/0000-0002-3242-9894>, e0123@mail.ru

Aydar R. Gilfanov — Engineer-Technologist, OJSC “Scientific and Production Association” State Institute of the Applied Optics, Kazan, 420075, Republic of Tatarstan, [sc 6701787491](https://orcid.org/0000-0001-8109-5029), <https://orcid.org/0000-0001-8109-5029>, cataklizm@inbox.ru

Nadezhda K. Pavlycheva — D. Sc., Full Professor, Kazan National Research Technical University named after A.N. Tupolev, Kazan, 420111, Republic of Tatarstan, [sc 6701787491](https://orcid.org/0000-0001-9395-3967), <https://orcid.org/0000-0001-9395-3967>, nkpavlych@rambler.ru

Статья поступила в редакцию 13.07.2022
Одобрена после рецензирования 10.10.2022
Принята к печати 30.11.2022

Received 13.07.2022
Approved after reviewing 10.10.2022
Accepted 30.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

ОПТИЧЕСКИЕ СИСТЕМЫ И ТЕХНОЛОГИИ OPTICAL ENGINEERING

doi: 10.17586/2226-1494-2022-22-6-1048-1054

УДК 543.421/.424

Применение методов инфракрасной спектроскопии в исследовании составов для проклейки бумаги

Халима Абишевна Бабаханова^{1✉}, Акмал Абдулло ўгли Садриддинов²,
Мансур Мехриддинович Абдуназаров³, Мадина Авазовна Бабаханова⁴,
Ирина Григорьевна Громыко⁵

^{1,2,3} Ташкентский институт текстильной и легкой промышленности, Ташкент, 100100, Республика Узбекистан

⁴ Ташкентский государственный технический университет ГУП «Фан ва тараккият», Ташкент, 100174, Республика Узбекистан

⁵ Белорусский государственный технологический университет, Минск, 220006, Республика Беларусь

¹ halima300@inbox.ru✉, <https://orcid.org/0000-0002-6956-2824>

² akmalsadridin0777@gmail.com, <https://orcid.org/0000-0002-0742-2439>

³ Abdunazarov.1977@mail.ru, <https://orcid.org/0000-0002-0099-8312>

⁴ madina89@mail.ru, <https://orcid.org/0000-0001-6832-3373>

⁵ gromyko@belstu.by, <https://orcid.org/0000-0002-0896-7262>

Аннотация

Предмет исследования. В работе предложен метод инфракрасной спектроскопии для объективного контроля клеящего состава при производстве бумаги. Для проклейки бумаги с целью повышения механической прочности и снижения водопоглощения используются гидрофобизирующие материалы. Один из наиболее эффективных и используемых гидрофобизирующих материалов — канифоль из сосновой живицы. В работе исследованы свойства нового клея на основе канифоли из живицы черешневого дерева для использования его при проклейке в бумажной массе, с целью расширения номенклатуры проклеиваемых веществ и снижения доли импортной составляющей в структуре местного бумажного производства. **Метод.** Для идентификации клея на канифольной основе из живицы черешневого дерева использован метод инфракрасной спектроскопии. Встроенное устройство с алмазным кристаллом универсального и компактного ИК-Фурье спектрометра Nicolet iS50 от Thermo Scientific позволило зарегистрировать спектры в среднем и дальнем инфракрасных диапазонах вплоть до 100 см⁻¹. Эффективность проклейки образцов бумаги исследована по стандартным методикам и оценена по следующим свойствам: плотность, прочность на разрыв при растяжении на разрывной машине, поверхностная впитываемость при одностороннем смачивании. **Основные результаты.** По инфракрасным спектрам на основании близости химического состава выявлена идентичность проклеиваемого вещества из живицы черешневого дерева клею из сосновой живицы, используемого для сравнения. Сравнительный анализ механических и поверхностных свойств образцов бумаги подтвердил возможность использования исследуемого проклеиваемого вещества, так как получены значения свойств, близкие к клею, примененному для сравнения. При этом целесообразным является добавление в предлагаемое вещество сухого целлюлозного сырья в пределах от 1 до 1,5 г на 100 г. **Практическая значимость.** Применение метода инфракрасной спектроскопии для экспресс-анализа состава проклеиваемого вещества является перспективным, так как появляется возможность управления технологическим процессом и создание проклеенных видов бумажной продукции с заданными свойствами, обеспечивающими качество печати без потерь мелких деталей изображения.

Ключевые слова

проклейка бумаги, живица черешневого дерева, канифоль из сосновой живицы, ИК спектроскопия, механические и поверхностные свойства бумаги

Ссылка для цитирования: Бабаханова Х.А., Садриддинов А.А., Абдуназаров М.М., Бабаханова М.А., Громыко И.Г. Применение методов инфракрасной спектроскопии в исследовании составов для проклейки бумаги // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1048–1054. doi: 10.17586/2226-1494-2022-22-6-1048-1054

Application of infrared spectroscopy methods in studying compositions for paper sizing

Khalima A. Babakhanova^{1✉}, Akmal A. Sadridinov², Mansur M. Abdunazarov³,
Madina A. Babakhanova⁴, Irina G. Gromyko⁵

^{1,2,3} Tashkent Institute of Textile and Light Industry, Tashkent, 100100, Republic of Uzbekistan

⁴ Tashkent State Technical University SUE "Fan va taraqqiyot", Tashkent, 100174, Republic of Uzbekistan

⁵ Belarusian State Technological University, Minsk, 220006, Republic of Belarus

¹ halima300@inbox.ru✉, <https://orcid.org/0000-0002-6956-2824>

² akmaalsadridin0777@gmail.com, <https://orcid.org/0000-0002-0742-2439>

³ Abdunazarov.1977@mail.ru, <https://orcid.org/0000-0002-0099-8312>

⁴ madina89@mail.ru, <https://orcid.org/0000-0001-6832-3373>

⁵ gromyko@belstu.by, <https://orcid.org/0000-0002-0896-7262>

Abstract

The paper proposes a method of infrared spectroscopy for objective control of the adhesive composition in the production of paper. For sizing paper, in order to increase mechanical strength and reduce water absorption, water-repellent materials are used. One of the most effective and used water-repellent material is ordinary pine resin rosin which is in short supply for the Republic of Uzbekistan. The paper investigates the properties of a new adhesive based on rosin from cherry resin for use in mass sizing in order to expand the range of sizing agents and reduce the share of imported components in the structure of local paper production. Infrared spectroscopy was used to identify the rosin-based glue from the resin of the cherry tree. The built-in device with a diamond crystal of the universal and compact Thermo Scientific Nicolet iS50 IR Fourier spectrometer made it possible to register spectra in the mid and far IR ranges up to 100 cm⁻¹. The efficiency of sizing paper samples was studied by standard methods and evaluated by the following properties: density, tensile strength at stretching on a tensile machine, surface absorbency with one-sided wetting. Infrared spectra revealed the identity of the sizing agent from the resin of the cherry tree to the glue from the pine resin, taken for comparison, because their chemical composition is the same. A comparative analysis of the mechanical and surface properties of paper samples confirmed the possibility of using the studied sizing agent, since the obtained values are close to each other, while it is advisable to add in the range from 1 to 1.5 g per 100 g of dry cellulose raw materials. The use of infrared spectroscopy for express analysis of the composition of the sizing agent is promising, since it becomes possible to control the technological process and create glued types of paper products with desired properties that ensure print quality without loss of fine image details.

Keywords

paper sizing, cherry tree resin, pine resin rosin, IR spectroscopy, mechanical and surface properties of paper

For citation: Babakhanova Kh.A., Sadridinov A.A., Abdunazarov M.M., Babakhanova M.A., Gromyko I.G. Application of infrared spectroscopy methods in studying compositions for paper sizing. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1048–1054 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1048-1054

Введение

Для динамического развития целлюлозно-бумажной промышленности из-за дефицита и дороговизны используемых основных и вспомогательных компонентов бумажной массы остро встает вопрос создания новых композиций и материалов из вторичных ресурсов. Проблеме переработки вторичных ресурсов для дальнейшего использования в качестве компонентов бумажной массы посвящены многочисленные работы ученых и специалистов этой отрасли. Заметим, что внедрение новых материалов в целлюлозно-бумажной промышленности требует расширения объективных методов контроля. Один из наиболее информативных методов идентификации материалов — метод инфракрасной (ИК) спектроскопии.

В работах [1–3] для снижения потребления дефицитного древесного сырья в качестве основного использованы вторичные волокна, получаемые в результате процесса переработки макулатуры или из отходов производства. Разработаны технологии производства бумаги из ежегодно возобновляющихся и имеющихся в большом количестве растительных материалов — сельскохозяйственных отходов [4–8], отходов текстильной [9–11] и фармацевтической отраслей [12, 13].

В работах [14–16] для подтверждения возможности использования целлюлозной массы из внутреннего слоя коры ветвей тутового дерева исследованы физико-механические, поверхностные и оптические свойства. Денситометрическое и микроскопическое исследования печатно-технических свойств бумаги выявили возможность их использования при печати маловязкими печатными красками [17, 18].

Для печати вязкими красками поверхность бумаги должна быть стойкой к воде и органическим соединениям, поэтому применяют гидрофобизирующие проклеивающие вещества [19, 20].

В зависимости от вида целлюлозы, качества производственной воды и других факторов используются различные проклеивающие вещества. Отсюда следует, что не существует универсального клея, пригодного для различных условий работы бумажной фабрики. Наиболее классическим проклеивающим материалом является белый клей (содержание свободной смолы 15–25 %) из живичной канифоли, который пригоден для широкого ассортимента бумаг. Канифольный клей осаждается на волокнах бумажной массы и в процессе сушки бумаги плавится, растекается по поверхности волокон, образуя сплошную пленку и, тем самым, обеспечивает чернило- и водонепроницаемость бумаги.

Цель настоящей работы — изучение возможности применения ИК спектроскопии для идентификации клея на канифольной основе из живицы черешневого дерева, используемого при проклейке в бумажной массе вместо канифоли из сосновой живицы, который является дефицитом для Республики Узбекистан.

Для достижения цели необходимо решить следующие задачи:

- выполнить идентификацию клея, полученного из живицы черешневого дерева, а также провести сравнительный анализ его спектра со спектром используемого образца методом ИК спектроскопии;
- для оценки эффективности действия проклеивающего вещества склеить образцы бумаги из целлюлозной массы веток тутового дерева канифольным клеем из живицы черешневого дерева и клеем из сосновой живицы и выполнить их сравнение;
- исследовать степень влияния проклеивающего вещества на механические и поверхностные свойства бумаги;
- изучить свойства бумаги и предложить рекомендации по применению.

Объекты и методы исследования

Для идентификации клея целесообразно использовать ИК спектроскопию с фурье-преобразованием, превосходящую оптические приборы по информативности и обеспечивающую высокое разрешение.

Фурье-спектрометр построен по принципу интерферометра Майкельсона, в котором исследуемое излучение служит источником освещения, а одно из зеркал имеет определенную постоянную скорость движения (рис. 1, *b*). Полученная кривая, описывающая зависимость отсчета приемника от разности хода лучей, после фурье-анализа имеет вид распределения интенсивности излучения по длинам волн (частотам) [21].

В данной работе использован универсальный и компактный ИК-Фурье спектрометр Nicolet iS50¹ от Thermo Scientific лаборатории физико-химических методов исследований² (рис. 1). По функциональным возможностям, техническим параметрам и простоте эксплуатации спектрометру Nicolet iS50 аналогов нет [22].

Встроенное устройство с алмазным кристаллом, имеющее собственный детектор, позволяет регистрировать спектры в среднем и дальнем ИК диапазонах вплоть до 100 см⁻¹. Способность прибора работать в широком спектральном диапазоне (от дальнего ИК до видимой области) легко идентифицирует неорганические компоненты в полимерах. При совместном применении встроенного устройства в режиме нарушенного полного внутреннего отражения (НПВО) и устройства автоматической смены светодетектора спектры в среднем и дальнем ИК диапазонах могут быть автомати-

чески зарегистрированы и «сшиты» с использованием программы на Visual Basic для получения единого спектра в диапазоне от 4000 до 100 см⁻¹ [23].

Метод ИК спектроскопии применен для идентификации клея из живицы черешневого дерева, предназначенного для проклейки волокнистой массы из целлюлозы веток тутового дерева, с целью расширения номенклатуры проклеивающих веществ и снижения доли импортной составляющей в структуре местного бумажного производства.

В качестве объектов исследования использованы клей из сырой смолы (живицы) черешневого дерева и образцы бумаги из целлюлозной массы внутренней коры веток тутового дерева.

Полученный клей из сырой смолы (живицы) черешневого дерева при температуре 85–95 °С был выпарен для очистки от скипидара и других летучих веществ, содержание которых в сырой смоле составляет 25 %. В полученном без использования химических средств натуральном продукте содержание воды составило не более 0,5 % при остаточной кислотности не более 0,3 %.

Для оценки эффективности действия проклеивающего вещества на листоотливном аппарате получены бумажные отливки массой 80 ± 2 г/м², проклеенные в массу канифольным клеем из живицы черешневого дерева и клеем из сосновой живицы. Механические и поверхностные свойства бумаги исследованы по методикам на основании: ГОСТ 27015-86³ — определена плотность; ГОСТ 13525.1-79⁴ — прочность на разрыв при растяжении на разрывной машине; ГОСТ 12605-97 (ISO 535-91)⁵ — поверхностная впитываемость при одностороннем смачивании.

Результаты исследования и их обсуждение

В результате анализа получены спектры образцов, которые представлены на рис. 2. ИК спектры скорректированы с использованием алгоритма коррекции приставки НВПО, входящего в программный пакет OMNIC. Сравнение ИК спектров проклеивающего вещества из живицы черешневого дерева (рис. 2, *a*) и клея из сосновой живицы, взятого для сравнения (рис. 2, *b*) показывает, что химический состав их близок.

В сравниваемых спектрах присутствуют характеристические полосы поглощения С=О-связи циклической ангидридной группы с максимумом при 1636 и 2123 см⁻¹. Кроме того, в обоих спектрах имеются полосы 2848 и 3343 см⁻¹, отвечающие за симметричные и асимметричные валентные колебания С–Н-связей ме-

¹ [Электронный ресурс]. Режим доступа: <http://www.bvl.by/oborudovanie/molekulyarnyi-analiz/ik-fure-spektrometry-1/ik-fure-spektrometr-nicolet-is50> (дата обращения: 10.02.2022).

² [Электронный ресурс]. Режим доступа: <https://biochem.uz> (дата обращения: 10.02.2022).

³ ГОСТ 27015-86 Бумага и картон. Методы определения толщины, плотности и удельного объема. Введен 01.01.88. М.: Издательство стандартов, 1986. 4 с.

⁴ ГОСТ 13525.1-79 Полуфабрикаты волокнистые. Бумага и картон. Методы определения прочности на разрыв и удлинения при растяжении. Введен 01.07.80. М.: Издательство стандартов, 2007. 5 с.

⁵ ГОСТ 12605-97 (ИСО 535-91) Бумага и картон. Метод определения поверхностной впитываемости воды при одностороннем смачивании (Метод Кобба). Введен 01.07.2001. М.: Издательство стандартов, 2001. 8 с.

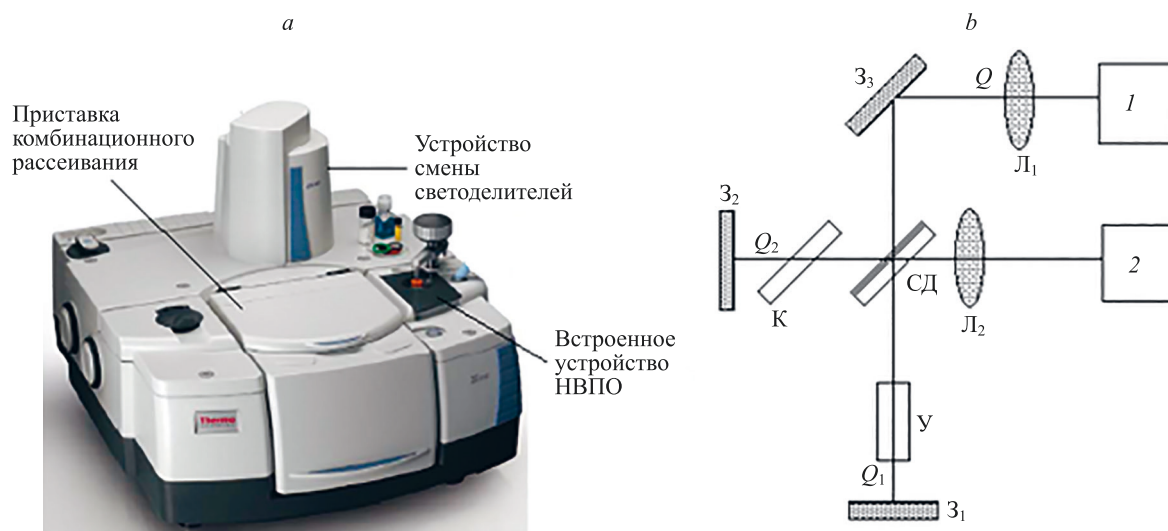


Рис. 1. ИК-Фурье спектрометр Nicolet iS50: общий вид (а) и оптическая система (б).

1 — источник света; 2 — фотоприемник; 3₁, 3₂, 3₃ — зеркала; СД — светоделитель; К — компенсатор; У — устройство, измеряющее разность хода; Л₁, Л₂ — линзы; Q, Q₁, Q₂ — интенсивность волны в плече интерферометра

Fig. 1. FTIR spectrometer Nicolet iS50: general view (a) and optical system (b).

1 — light source; 2 — photodetector; 3₁, 3₂, 3₃ — mirrors; СД — beam splitter; К — compensator; У — a device that measures the path difference; Л₁, Л₂ — lenses; Q, Q₁, Q₂ — wave intensity in the interferometer arm

тиленовых и метильных групп фенантренового скелета смоляных кислот канифоли; 1636 см⁻¹ — характеристическая полоса валентных колебаний С=О-группы в карбоксильной группе СООН для ненасыщенных кислот; 1463, 1400, 1382, 1360 см⁻¹ — группа полос поглощения, характеризующая деформационные (плоские ножничные) колебания С–Н-связи метиленовых и метильных групп фенантренового скелета смоляных кислот канифоли; 1261 см⁻¹ — валентные колебания С–О-связи в карбоксильной группе. В связи с тем, что в клее присутствует мочевиная, для которой характерно в области 3345 см⁻¹ поглощение N–Н-группы, то данная полоса также была найдена в исследуемых спектрах. Как следует из анализа результатов обработки спектров по поисковой базе, качественный и частично количественный состав совпадают в обоих образцах. На основании рассмотренных экспериментальных данных, которые включают ИК спектры клеев из живицы черешневого дерева (образцы AAS-atr-TITLP-SadriiddinovA-1) и из живицы сосны (образцы AAS-atr-TITLP-SadriiddinovA-2) можно утверждать, что их состав идентичен.

Для оценки эффективности действия проклеивающего вещества определена прочность на разрыв при растяжении на разрывной машине по ГОСТ 13525.1-79¹. В соответствии с методом, образец был закреплен в зажимах разрывной машины, не касаясь его испытуемой части, с силой натяжения не более 0,3 Н. В результате определено усилие, вызывающее разрушение образца и его удлинение до момента разрыва.

¹ ГОСТ 13525.1-79 Полуфабрикаты волокнистые. Бумага и картон. Методы определения прочности на разрыв и удлинения при растяжении. Введен 01.07.80. М.: Издательство стандартов, 2007. 5 с.

Разрывная длина L вычислена по формуле:

$$L = \frac{l_0 F}{m},$$

где F — разрушающее усилие; l_0 — номинальное расстояние между зажимами; m — масса 1 м² образца.

Значение предела прочности при растяжении σ рассчитано по формуле:

$$\sigma = \frac{F}{bh},$$

где h — толщина образца бумаги; b — ширина образца бумаги.

Относительная погрешность определения предела прочности при растяжении не превышает $\pm 10\%$ при доверительной вероятности 0,95.

Сравнительный анализ механических свойств образцов бумаги подтвердил возможность использования исследуемого проклеивающего вещества, так как получены значения, близкие друг к другу, при этом целесообразно добавление проклеивающего вещества в пределах от 1 до 1,5 г на 100 г сухого целлюлозного сырья, так как увеличение добавки до 2 г способствует уменьшению межволоконных сил связи на 21 %.

Для выявления степени влияния проклеивающего вещества на гидрофобность бумаги определена поверхностная впитываемость при одностороннем смачивании (ГОСТ 12605-97 (ISO 535-91)²) по методу

² ГОСТ 12605-97 (ISO 535-91) Бумага и картон. Метод определения поверхностной впитываемости воды при одностороннем смачивании (Метод Кобба). Введен 01.07.2001. М.: Издательство стандартов, 2001. 8 с.

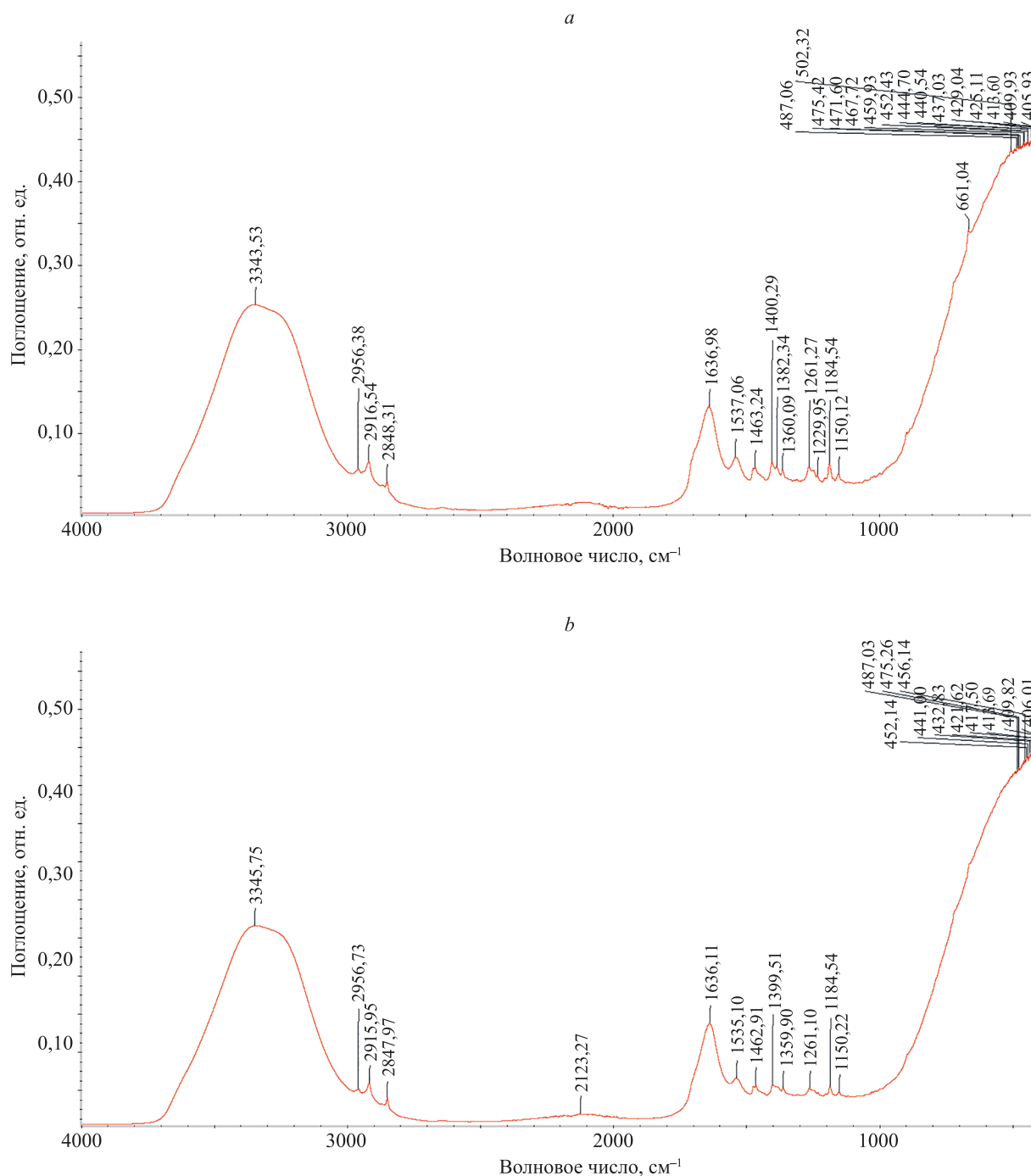


Рис. 2. Инфракрасные спектры проклеивающего вещества из живицы черешневого дерева (а) и клея из сосновой живицы, используемого для сравнения (b)

Fig. 2. Infrared spectra sizing agent from cherry tree resin (a) and glue from pine resin taken for comparison (b)

Кобб(30), путем определения массы воды, поглощенной поверхностью бумаги при смачивании одной стороны испытуемого образца

$$\text{Кобб}_{(30)} = 100(m_2 - m_1),$$

где m_1 и m_2 — массы образца до и после испытаний.

Результаты исследований зависимости свойств бумаги от вида проклеивающего вещества и содержания наполнителя приведены в таблице.

В ходе анализа поверхностных свойств образцов бумаги (таблица) выявлена эффективность проклейки от 22,5 до 30 % с увеличением добавления проклеивающего клея, что можно объяснить уплотнением структуры, которая препятствует проникновению водных растворов.

Таблица. Механические и поверхностные свойства бумаги из целлюлозной массы внутренней коры веток тутового дерева
Table. Mechanical and surface properties of paper from the cellulose pulp of the inner bark of mulberry branches

Номер образца	Масса вводимого связующего на 100 г сухого целлюлозного сырья, г		Свойства бумаги					
			механические				поверхностные	
	из живицы черешневого дерева	из живицы сосны	плотность, г/см ³	разрушающее усилие, Н	прочность на изгиб, число двойных перегибов	предел прочности, МПа	поверхностная впитываемость, г/м ²	влажность, %
1	1,0	—	0,65	69,0	1395	55,7	40	9,3
2	1,5	—	0,67	65,1	1203	52,2	31	8,8
3	2,0	—	0,68	55,1	1180	44,3	28	8,2
4	—	1,0	0,65	71,0	1394	56,0	41	9,2
5	—	1,5	0,66	66,2	1205	51,8	30	8,8
6	—	2,0	0,68	54,0	1164	43,8	25	8,1

Заключение

Сравнение инфракрасных спектров проклеивающего вещества из живицы черешневого дерева и клея из сосновой живицы, используемого для сравнения, показало близкие значения химического состава. Сравнительный анализ механических свойств также подтвердил возможность использования исследуемого проклеивающего вещества наравне с клеем из сосновой живицы. При этом целесообразным является добавление клея из живицы черешневого дерева в пределах от 1 до 1,5 г на 100 г сухого целлюлозного сырья, так как увеличение добавки до 2 г способствует уменьшению

межволоконных сил связи на 21 %. Выявлена эффективность проклейки от 22,5 до 30 % с увеличением добавления проклеивающего клея, что можно объяснить уплотнением структуры, которая препятствует проникновению водных растворов. Полученные результаты свидетельствуют о перспективности используемого метода инфракрасной спектроскопии для экспресс-анализа состава проклеивающего вещества, так как появляется возможность управления технологическим процессом и создания проклеенных видов бумажной продукции с заданными свойствами, обеспечивающими качество печати без потерь мелких деталей изображения.

Литература

- Махотина Л.Г. Современные тенденции в технологии бумаги для печати // Целлюлоза. Бумага. Картон. 2008. № 3. С. 52–55.
- CEPI Sustainability Report. Brussels, CEPI, 2005. 36 p.
- Пузырев С.С. Ресурсосберегающая технология переработки макулатуры // ЛесПромИнформ. 2006. № 3(34). С. 104–109.
- Барбаш В.А., Трембус И.В. Органосольвентный способ получения волокнистых полуфабрикатов из отходов сельского хозяйства // Экотехнологии и ресурсосбережение. 2002. № 6. С. 29–32.
- Рю Х., Ким Ч.К., Вон Ж.-М. Способ получения целлюлозной массы из стеблей кукурузы. Патент RU 2249636C2. Бюл. 2005. № 10.
- Siegl S. Производство целлюлозы из соломы, тростника и багассы // Natural pulping update and progress. 2002. С. 237–249.
- Huang G., Chen Z., Zhang C. Aqueous ammonia-caustic potash pulping of rice straw // Chemistry and Industry of Forest Products. 2002. V. 22. N 4. P. 31–36.
- Lee K.-H., Won J.-M. Process for producing pulp utilizing bamboo and pulp and papers produced using the same. Patent WO2007/004757A1. 2007.
- Бабаханова Х.А., Алимова Х.А. К оценке эффекта применения в бумажных изделиях отходов натурального шелка и кенафа // Шелк. 2000. № 2. С. 20–21.
- Мухамадрасулов Ш.Х. Повышение эффективности технологии производства натурального шелка // Физика волокнистых материалов: структура, свойства, наукоемкие технологии и материалы (SMARTKX). 2015. № 1. С. 216–220.
- Babakanova H.A. Papier mit Faserabfaellen aus der Textilindustrie // Textile Problems. 2003. N 4. P. 34–35. (in German)
- Набиев Д.С., Набиева И.А., Бабаханова Х.А., Шахидова Ф.Н. Способ получения целлюлозы. Патент РУз IAP04879. 2014.
- Бабаханова Х.А., Галимова З.К., Абдуназаров М.М. Использование фармоотходов в бумажной отрасли Республики

References

- Mahotina L.G. The modern trends in the technology of the printing paper production. *Celluloza. Bumaga. Karton*, 2008, no. 3, pp. 52–55. (in Russian)
- CEPI Sustainability Report. Brussels, CEPI, 2005, 36 p.
- Puzyrev S.S. Resource saving wastepaper recycling technology. *LesPromInform*, 2006, no. 3(34), pp. 104–109. (in Russian)
- Barbash V.A., Trembus I.V. Organosolv method of obtaining fibrous semi-finished products from agricultural waste. *Jekotekhnologii i resursosberezhenie*, 2002, no. 6, pp. 29–32. (in Russian)
- Rju Kh., Kim Ch.K., Von Z.-M. Method of preparing pulp from corn stems. *Patent RU 2249636C2*, 2005. (in Russian)
- Siegl S. Pulp production from straw, reed and bagasse. *Natural pulping update and progress*, 2002, pp. 237–249.
- Huang G., Chen Z., Zhang C. Aqueous ammonia-caustic potash pulping of rice straw. *Chemistry and Industry of Forest Products*, 2002, vol. 22, no. 4, pp. 31–36.
- Lee K.-H., Won J.-M. Process for producing pulp utilizing bamboo and pulp and papers produced using the same. *Patent WO2007/004757A1*, 2007.
- Babakanova Kh.A., Alimova Kh.A. To the assessment of the effect of using natural silk and kenaf waste in paper products. *Shelk*, 2000, no. 2, pp. 20–21. (in Russian)
- Muxamadrasulov Sh.X. Improving the efficiency of Manufacturing Natural Silk. *Fizika voloknistykh materialov: struktura, svoystva. naukoemykiye tekhnologii i materialy (SMARTKX)*, 2015, no. 1, pp. 216–220. (in Russian)
- Babakanova H.A. Papier mit Faserabfaellen aus der Textilindustrie. *Textile Problems*, 2003, no. 4, pp. 34–35. (in German)
- Nabiev O.S., Nabieva I.A., Babakanova H.A., Shahidova F.N. The Method of Cellulose Production. *Patent UZ IAP04879*, 2014. (in Russian)

- Узбекистан // Химия растительного сырья. 2020. № 3. С. 285–290. <https://doi.org/10.14258/Jcprm.2020037298>
14. Бабаханова Х.А., Галимова З.К., Абдуназаров М.М., Исmoilov И.И. Свойства бумаги, в составе которой целлюлозная масса из веток тутта // Высшая школа: научные исследования: Материалы межвузовского научного конгресса, Москва, 30 апреля 2020 г. Уфа: Инфинити, 2020. С. 118–122.
 15. Бабаханова Х.А., Галимова З.К., Абдуназаров М.М., Исmoilov И.И. Структура бумаги с добавлением целлюлозной массы из коры веток тутового дерева // Химия растительного сырья. 2020. № 4. С. 261–266. <https://doi.org/10.14258/jcprm.2020047761>
 16. Бабаханова Х.А., Галимова З.К., Абдуназаров М.М., Исmoilov И.И. Целлюлозная масса из коры веток тутовника для бумажной отрасли // Известия вузов. Лесной журнал. 2020. № 5. С. 193–200. <https://doi.org/10.37482/0536-1036-2020-5-193-200>
 17. Бабаханова Х.А., Галимова З.К., Абдуназаров М.М., Исmoilov И.И. Исследование красковосприимчивости бумаги с добавлением целлюлозной массы из коры веток тутового дерева // Лесной вестник / Forestry Bulletin. 2021. Т. 25. № 5. С. 97–105. <https://doi.org/10.18698/2542-1468-2021-5-97-105>
 18. Бабаханова Х.А., Садриддинов А.А., Галимова З.К., Исmoilov И.И. Исследование микрогеометрии поверхности бумаги из целлюлозы, полученной из коры веток тутового дерева // Лесной вестник / Forestry Bulletin. 2022. Т. 26. № 1. С. 84–90. <https://doi.org/10.18698/2542-1468-2022-1-84-90>
 19. Муллина Э.Р., Мишурина О.А., Чупрова Л.В., Ершова О.В. Влияние химической природы проклеивающих компонентов на гидрофильные и гидрофобные свойства целлюлозных материалов // Современные проблемы науки и образования. 2014. № 6. С. 250.
 20. Остапенко А.А., Мороз В.Н., Барбаш В.А., Кожевников С.Ю., Дубовый В.К., Ковернинский И.Н. Повышение качества бумаги из макулатуры химическими функциональными веществами // Химия растительного сырья. 2012. № 1. С. 187–190.
 21. Ефимова А.И., Зайцев В.Б., Болдырев Н.Ю., Кошкарёв П.К. Инфракрасная Фурье-спектроскопия: учебное пособие. М., 2008. 123 с.
 22. Стюарт И.Г. Введение в фурье-оптику. М.: Мир, 1985. 182 с.
 23. Jansen J. Определение причин возникновения брака полимерных изделий // Пластические массы. 2018. № 9–10. С. 26–29.
 13. Babakhanova Kh.A., Galimova Z.K., Abdunazarov M.M. Use of pharmaceutical waste in paper industry of the Republic of Uzbekistan. *Khimiya Rastitel'nogo Syr'ya*, 2020, no. 3, pp. 285–290. <https://doi.org/10.14258/Jcprm.2020037298>. (in Russian)
 14. Babakhanova Kh.A., Galimova Z.K., Abdunazarov M.M., Ismoilov I.I. Properties of paper containing pulp from mulberry branches. *Inter-University Scientific Congress. Higher school: Scientific research*. Moscow. 30.04.2020. Ufa, Infiniti, 2020, pp. 118–122 c. (in Russian)
 15. Babakhanova Kh.A., Galimova Z.K., Abdunazarov M.M., Ismoilov I.I. Texture of paper with the addition of cellulose pulp from the bark of mulberry branches. *Khimiya Rastitel'nogo Syr'ya*, 2020, no. 4, pp. 261–266. <https://doi.org/10.14258/jcprm.2020047761>. (in Russian)
 16. Babakhanova Kh.A., Galimova Z.K., Abdunazarov M.M., Ismailov I.I. Cellulose pulp from mulberry branch bark for the paper industry. *Lesnoy Zhurnal (Russian Forestry Journal)*, 2020, no. 5, pp. 193–200. <https://doi.org/10.37482/0536-1036-2020-5-193-200>. (in Russian)
 17. Babakhanova Kh.A., Galimova Z.K., Abdunazarov M.M., Ismailov I.I. Paper trapping research after adding mulberry tree branches bark cellulose pulp. *Lesnoy vestnik / Forestry Bulletin*, 2021, vol. 25, no. 5, pp. 97–105. <https://doi.org/10.18698/2542-1468-2021-5-97-105>. (in Russian)
 18. Babakhanova KH.A., Sadriddinov A.A., Galimova Z.K., Ismailov I.I. Microgeometry of cellulose paper top made of mulberry tree bark branches. *Lesnoy vestnik / Forestry Bulletin*, 2022, no. 1, pp. 84–90. <https://doi.org/10.18698/2542-1468-2022-1-84-90>. (in Russian)
 19. Mullina E.R., Mishurina O.A., Chuprova L.V., Ershova O.V. Influences of sizing compound chemistry on hydrophilic and hydrophobic properties of cellulosic materials. *Modern problems of science and education*, 2014, no. 6, pp. 250. (in Russian)
 20. Ostapenko A.A., Moroz V.N., Barbash V.A., Kozhevnikov S.Iu., Duboviy V.K., Koverninskii I.N. Improving the quality of paper from waste paper from paper with chemical functional substances. *Khimiya Rastitel'nogo Syr'ya*, 2012, no. 1, pp. 187–190. (in Russian)
 21. Efimova A.I., Zaitcev V.B., Boldyrev N.Iu., Koshkarev P.K. *Infrared Fourier Spectrometry*. Tutorial. Moscow, 2008, 123 p. (in Russian)
 22. Steward E.G. *Fourier Optics: An Introduction*. NY, E. Horwood Publ., 1983, 185 p.
 23. Jansen J. Root cause analysis of a failed plastic cover. *Plasticheskie massy*, 2018, no. 9-10, pp. 26-29. (in Russian)

Авторы

Бабаханова Халима Абишевна — доктор технических наук, профессор, профессор, Ташкентский институт текстильной и легкой промышленности, Ташкент, 100100, Республика Узбекистан, 57201773008, <https://orcid.org/0000-0002-6956-2824>, halima300@inbox.ru

Садриддинов Акмал Абдулло ўгли — докторант, Ташкентский институт текстильной и легкой промышленности, Ташкент, 100100, Республика Узбекистан, 57365737500, <https://orcid.org/0000-0002-0742-2439>, akmalsadriddin0777@gmail.com

Абдуназаров Мансур Мехриддинович — старший преподаватель, Ташкентский институт текстильной и легкой промышленности, Ташкент, 100100, Республика Узбекистан, <https://orcid.org/0000-0002-0099-8312>, Abdunazarov.1977@mail.ru

Бабаханова Мадина Авазовна — PhD, докторант, Государственное унитарное предприятие «Фан ва тараққиёт» Ташкентского государственного технического университета, Ташкент, 100174, Республика Узбекистан, 57366173600, <https://orcid.org/0000-0001-6832-3373>, madina89@mail.ru

Громыко Ирина Григорьевна — кандидат технических наук, доцент, заведующая кафедрой, Белорусский государственный технологический университет, Минск, 220006, Республика Беларусь, <https://orcid.org/0000-0002-0896-7262>, gromyko@belstu.by

Authors

Khalima A. Babakhanova — D. Sc., Full Professor, Tashkent Institute of Textile and Light Industry, Tashkent, 100100, Republic of Uzbekistan, 57201773008, <https://orcid.org/0000-0002-6956-2824>, halima300@inbox.ru

Akmal A. Sadriddinov — Doctoral Student, Tashkent Institute of Textile and Light Industry, Tashkent, 100100, Republic of Uzbekistan, 57365737500, <https://orcid.org/0000-0002-0742-2439>, akmalsadriddin0777@gmail.com

Mansur M. Abdunazarov — Senior Lecturer, Tashkent Institute of Textile and Light Industry, Tashkent, 100100, Republic of Uzbekistan, <https://orcid.org/0000-0002-0099-8312>, Abdunazarov.1977@mail.ru

Madina A. Babakhanova — PhD, Doctoral Student, SUE «Fan va taraqqiyot» of the Tashkent State Technical University, Tashkent, 100174, Republic of Uzbekistan, 57366173600, <https://orcid.org/0000-0001-6832-3373>, madina89@mail.ru

Irina G. Gromyko — PhD, Associate Professor, Head of Department, Belarusian State Technological University, Minsk, 220006, Republic of Belarus, <https://orcid.org/0000-0002-0896-7262>, gromyko@belstu.by

Статья поступила в редакцию 10.06.2022

Одобрена после рецензирования 29.09.2022

Принята к печати 12.11.2022

Received 10.06.2022

Approved after reviewing 29.09.2022

Accepted 12.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1055-1062

УДК 681.772.7

Методика оптимизации распределения плотности пикселей по зоне наблюдения

Владимир Владимирович Волхонский¹✉, Владислав Александрович Ковалевский²

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ volkhonski@mail.ru✉, <https://orcid.org/0000-0001-9628-2046>

² Kovalvlad9@gmail.com, <https://orcid.org/0000-0001-7042-7905>

Аннотация

Предмет исследования. Рассмотрена задача оптимизации параметров матрицы видеокамеры для снижения избыточности информативности видеосигнала. Актуальность темы обусловлена существующей избыточностью формируемых матрицами видеосигналов в разных частях зоны наблюдения и, как следствие, избыточной нагрузкой устройств передачи, хранения и обработки сигналов. Данная задача решена путем достижения равномерности распределения плотности пикселей по зоне наблюдения. Методика основана на решении задачи программным способом в отличие от существующих решений, использующих аппаратные методы. **Метод.** Предложен метод представления зоны обзора как совокупности подмножеств пространства задач наблюдения с последующей ее фрагментацией. Задачи наблюдения определяют требуемую минимальную плотность пикселей в разных фрагментах. **Основные результаты.** Осуществлена сегментация матрицы в соответствии с размерами фрагментов зоны наблюдения и заданным допустимым диапазоном изменения распределения плотности пикселей в зоне обзора телекамеры. Предложены обязательные и дополнительные критерии оптимизации распределения. Сформулирована методика оптимизации. В соответствии с коэффициентом избыточности распределения плотности пикселей в сегментах произведено объединение пикселей в группы, различные в разных сегментах. Приведены примеры решения задач оптимизации по разным критериям. **Практическая значимость.** Предложенный подход позволил минимизировать избыточность плотности пикселей и, как следствие, снизить нагрузку каналов связи, объем памяти устройств хранения видеoinформации и требования к производительности устройств обработки видеосигналов. Решена задача формирования непрерывного изображения зоны наблюдения. Результаты работы могут найти применение при обработке видеосигналов и разработке новых телекамер систем наблюдения.

Ключевые слова

распределение плотности пикселей, плотность пикселей, задача наблюдения, матрица камеры, зона обзора

Ссылка для цитирования: Волхонский В.В., Ковалевский В.А. Методика оптимизации распределения плотности пикселей по зоне наблюдения // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1055–1062. doi: 10.17586/2226-1494-2022-22-6-1055-1062

Distribution optimization method of pixel density by surveillance area

Vladimir V. Volkhonskiy¹✉, Vladislav A. Kovalevskiy²

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ volkhonski@mail.ru✉, <https://orcid.org/0000-0001-9628-2046>

² Kovalvlad9@gmail.com, <https://orcid.org/0000-0001-7042-7905>

Abstract

The task of optimizing the parameters of the video camera matrix to reduce the redundancy of the generated video signal is considered. The relevance of the topic is due to the existing redundancy of video signals generated by matrices in different parts of the observation zone and, as a result, excessive loading of signal transmission, storage and processing devices. The problem is solved by achieving a uniform distribution of pixel density over the observation area. The

methodology is based on solving the problem programmatically, unlike existing solutions using hardware methods. It is based on the representation of the view area as a set of subsets of the observation task space with its subsequent fragmentation in accordance with the observation tasks being solved which determine the required minimum pixel density in different fragments. To solve the problem, the matrix is segmented in according to the size of the fragments of the observation area and the specified allowable range of changes in the distribution of pixel density in the viewing area of the camera. Mandatory and additional distribution optimization criteria are proposed. The optimization technique is formulated. In according to the redundancy coefficient of the pixel density distribution in different segments, pixels are combined into groups that are different in different segments. Examples of solving optimization problems according to different criteria are given. The proposed approach makes it possible to minimize the pixel density redundancy and thereby reduce the load on communication channels, the amount of memory in video information storage devices, and the performance requirements for video signal processing devices. In this case, the problem of forming a continuous image of the observation zone is also solved. Results could be used for video signal processing and design of new cameras for video surveillance systems.

Keywords

pixel density distribution, pixel density, observation task, camera matrix, view area

For citation: Volkhonskiy V.V., Kovalevskiy V.A. Distribution optimization method of pixel density by surveillance area. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1055–1062 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1055-1062

Введение

Для решения задач телевизионного наблюдения необходимо сформировать видеоизображение требуемого качества. Оценка качества видеоизображения производится с использованием различных характеристик, но система телевизионного наблюдения рассматривается с точки зрения получения максимально информативного изображения контролируемой зоны. Информативность видеоизображения зависит от многих характеристик и параметров оборудования системы наблюдения, а также способов и условий его использования. Отметим, что наиболее важный критерий оценки информативности — степень детализации изображения, характеризующая возможность различать мелкие детали. Степень детализации изображения, в свою очередь, определяется плотностью пикселей (ПП) в различных частях зоны наблюдения. Необходимое значение ПП задается выбранной задачей наблюдения (мониторинг, обнаружение, обзор, распознавание или идентификация) в наиболее дальней точке зоны обзора [1]. Как результат, в силу особенностей формирования видеоизображения, ПП в ближних частях зоны обзора оказывается выше необходимой — вблизи телевизионной камеры она может многократно превышать минимально необходимый уровень. Это приводит к тому, что значительная часть изображения содержит избыточную информацию и, как следствие, неоправданно увеличиваются: загрузка канала связи, объем памяти устройств хранения видеосигнала и требования к ресурсам средств обработки видеосигнала. В результате возрастает итоговая стоимость оборудования системы и снижается быстродействие.

Таким образом, существует проблема достижения: необходимой минимальной ПП в дальней части зоны видеонаблюдения и обеспечения равномерности распределения ПП по зоне наблюдения. Известны ряд методов и способов их реализации, которые могут решить проблему получения видеоизображения требуемого качества [2–11]. Однако они приводят либо к заметному усложнению технических решений, а следовательно, и к увеличению стоимости, либо к формированию набор отдельных видеоизображений, различных участков

заданной зоны наблюдения с перекрытиями или пропусками в той или иной части зоны [1, 12]. Такие усложнения решений заметно затрудняют работу операторов, особенно во внештатных ситуациях и снижают эффективность системы телевизионного наблюдения. По этой причине задача формирования единого видеоизображения зоны наблюдения с требуемыми параметрами, в первую очередь с заданной ПП, в идеальном случае равномерно распределенной по всей зоне наблюдения, является актуальной.

Для разработки методики, обеспечивающей равномерное изменение значения ПП для решения заданной задачи наблюдения, требуется выполнить: анализ особенностей формирования изображения в различных частях зоны обзора, сравнение известных решений этой проблемы и формулировку критериев оценки равномерности ПП. В результате необходимо разработать методику решения поставленной задачи. Также важно иметь возможность программной автоматизации процесса адаптации параметров телевизионной камеры к конкретным условиям для получения единого изображения зоны наблюдения без пропусков и перекрытий зон обзора. При этом очевидно, что адаптивно менять параметры камеры с соответствующей компенсацией настроек возможно только с использованием электронных регулировок ее параметров. В настоящей работе предложено решение перечисленных вопросов.

Зона эффективного обзора

С точки зрения решения задач наблюдения, целесообразно учитывать характер распределения ПП по зоне обзора и влияние на это распределение различных особенностей формы зоны обзора камеры, положения в ней субъекта или объекта наблюдения (СОИ) и параметров установки камеры (места и высоты установки, углов наклона, поворота и обзора).

В общем случае основные требования к формируемому видеоизображению всей зоны наблюдения должны быть следующими:

- достижение минимальной ПП, требуемой для решения заданной задачи наблюдения, во всей зоне обзора;

- обеспечение равномерного распределения ПП по всей зоне наблюдения при заданных критериях и ограничениях;
- формирование единого видеоизображения, полностью перекрывающего всю зону наблюдения без не просматриваемых участков и перекрытия.

Для оценки характера распределения ПП необходимо проанализировать особенности формирования зоны обзора камеры. В общем случае зона обзора телевизионной камеры представляет собой косоугольную пирамиду (рис. 1, пунктирные линии).

Предположим, что зона наблюдения — область возможного нахождения СОН. В свою очередь, зона наблюдения должна полностью перекрываться зоной обзора телевизионной камеры, поэтому пространство наблюдения — вся область зоны обзора. Пространство наблюдения, учитывая реальное распределение ПП в этом пространстве, можно представить совокупностью подпространств разных задач наблюдения. Размеры, как пространства наблюдения, так и подпространств задач наблюдения, будут зависеть от решаемой задачи наблюдения, параметров как самой телевизионной камеры, так и ее установки, и параметров СОН.

Форма зоны обзора, показанная на рис. 1, является теоретической, так как может формироваться видеоизображение любого СОН, попадающего в зону. Однако для решения задач наблюдения необходимы определенные значения ПП. Из-за этого по дальности зона обзора будет ограничиваться возможностью достижения этих значений. Также, кроме попадания СОН в зону обзора, в ближней зоне будут ограничения, связанные с глубиной резкости и искажениями [1], которые в настоящей работе не учитываются.

С учетом упомянутых свойств зоны обзора можно говорить о пространстве Z возможного нахождения СОН, в котором формируется его изображение с

той или иной информативностью. Последняя будет определяться ПП в месте нахождения СОН. В силу дискретности видеосигнала, формируемого матрицей камеры, изображение будет состоять из элементарных элементов, размер которых сравним с размером СОН и зависит от расстояния до камеры, угла обзора и разрешения матрицы. В результате можно говорить о совокупности элементов z_{ij} множества Z пространства наблюдения. Положение каждого элемента множества $z_{ij} = f(l, \varphi_i, \psi_j)$, где $i = 1, \dots, I; j = 1, \dots, J; I$ и J — число строк и столбцов матрицы камеры в пространстве наблюдения будут функцией дальности l , а также углов φ_i и ψ_j направления на ij -й элемент СОН относительно оси камеры.

Введем понятие *зоны эффективного обзора* (ЗЭО), как части зоны обзора, в которой выполняются требования по заданной ПП. Физически по дальности ЗЭО будет ограничиваться поверхностью равной ПП (рис. 1), находящейся на максимально возможном расстоянии $L_{\max}$ нахождения СОН. В каждом ij -ом элементе этой сферической поверхности ПП одинакова. Заметим, что выбор расстояния $L_{\max}$ зависит от формы и положения самого СОН относительно камеры. Так, для человека это могут быть расстояния до верхней L^B (голова стоящего) или нижней L^H (ноги или голова лежащего) частей тела, например для идентификации стоящего или лежащего человека.

С учетом необходимости достижения требуемой ПП D_p^0 во всей зоне обзора множество Z пространства наблюдения можно представить в виде совокупности $Z \subseteq [Z_0, Z_1, Z_2, \dots, Z_K, Z_{K+1}]$ подмножеств Z_k разных задач наблюдения. При этом множество Z включает в себя подмножества $Z \subseteq Z_k$, где $k = 0, 1, \dots, K + 1$. Подмножества определяются пространственными областями зоны обзора, в которых решаются различные задачи наблюдения. Им можно поставить в соответствие значения индекса k , например для типовых задач наблюдения субъекта [1], следующим образом: 1 — идентификация, 2 — распознавание, 3 — обзор, 4 — обнаружение, 5 — мониторинг, а также 0 — как часть зоны обзора неуверенных попадания и идентификации СОН в эту зону и 6 — неуверенного попадания и мониторинга. В частных случаях других объектов наблюдения и, следовательно, иных формулировок задач наблюдения или требований к ПП эти индексы могут быть другими [1].

Во многих практических случаях, в частности, когда высота установки камеры значительно меньше протяженности зоны обзора, допустимо рассматривать двумерное изображение зоны обзора как проекции на горизонтальную плоскость (рис. 2). При этом границы подмножеств в случае малых углов обзора допустимо считать прямыми¹. Можно говорить о двух основных моделях СОН. Это распределенный СОН — соизмеримый по размерам с размерами ЗЭО, и квазиточечный, размеры которого намного меньше размеров ЗЭО. Тогда двумерное представление будет применимо для случая

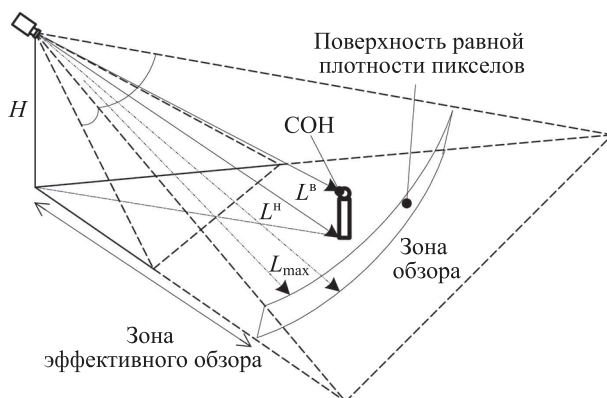


Рис. 1. Зона обзора камеры.

$L_{\max}$ — максимально возможное расстояние; L^B и L^H — расстояния до верхней и нижней частей тела человека; H — высота установки камеры. Пунктирные линии — зона обзора телевизионной камеры

Fig. 1. Camera field of view.

$L_{\max}$ is the maximum possible distance; L^B and L^H are distances to the upper and lower parts of the human body; H is camera installation height. Dashed lines — TV camera field of view

¹ Panomera cameras [Электронный ресурс]. Режим доступа: <https://www.dallmeier.com/technology/panomera-cameras/> (дата обращения: 29.09.2022).

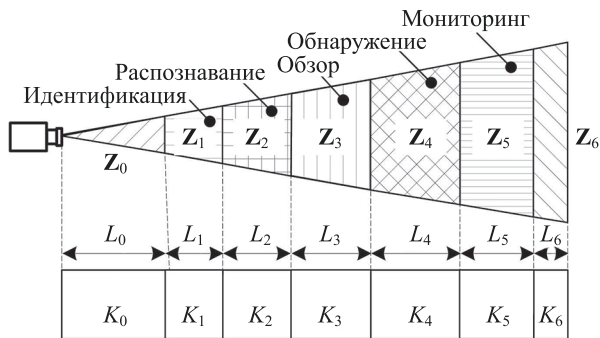


Рис. 2. Двумерное изображение протяженной зоны обзора.

Индексы обозначают: 0 — часть зоны обзора неуверенных попадения и идентификации субъекта или объекта наблюдения; 6 — часть зоны обзора неуверенных попадения и мониторинга; 1–5 — подмножества зоны эффективного обзора

Fig. 2. 2D image of an extended field of view.

Indices indicate: 0 is part of the viewing area of uncertain hits and identification of the subject or object of observation; 6 is part of the area of view of uncertain hits and monitoring; 1–5 are subsets of the effective view area

квазиточечного объекта наблюдения, либо критически важного для решения задачи наблюдения элемента СОН, например глаза. Но в более общем случае необходимо учитывать размер и положение СОН, включая его возможные наклон и поворот относительно камеры при разных размерах формы и размера в разных плоскостях или ракурсах.

Способы достижения равномерной плотности пикселей

Известны несколько технических решений, позволяющих в определенной степени решить задачу достижения равномерной ПП. Наиболее эффективные решения:

- 1) перекрытие зоны наблюдения зонами обзора отдельных камер, установленных в одном или нескольких местах в зависимости от требуемого ракурса;
- 2) установка мультимодульной камеры с независимой ориентацией модулей, обеспечивающих полное перекрытие зоны обзора (рис. 3).

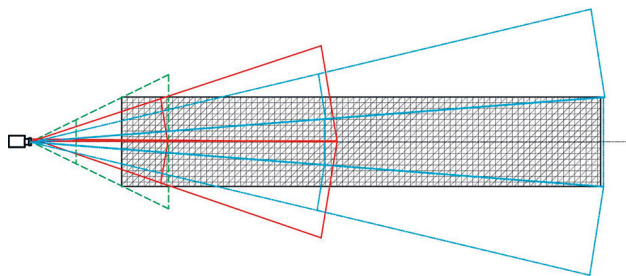


Рис. 3. Формирование зоны обзора мультимодульными камерами.

Различные части общей зоны обзора соответствуют зонам обзора, формируемым каждым модулем «камера-объектив»

Fig. 3. Formation of the view area by multimodule cameras. Different parts of the total view area corresponding to the area of view formed by each module “camera-lens”

Таким образом, необходимо использование нескольких камер или устройства, состоящего из нескольких модулей — «камера-объектив», каждый из которых контролирует часть общей зоны обзора с последующей стыковкой отдельных изображений в единое целое для использования в процессе наблюдения [12]. Это позволяет достичь заметного улучшения равномерности распределения ПП. Однако достигается значительным усложнением аппаратной части конструкции и программного обеспечения, и, как следствие, больших вычислительных ресурсов. В результате имеет место заметное удорожание устройства.

Решения 1 и 2 требуют выполнение ряда действий, связанных с регулировками углов наклона, поворота и обзора, и высоты установки отдельных камер или модулей (рис. 3). Кроме того, возникает необходимость повышенной точности регулировок для стыковки изображений отдельных камер в одно общее изображение. Автоматизация процесса настройки практически не реализована.

Идея, позволяющая устранить перечисленные недостатки, лежит в основе мультимодульной телевизионной камеры¹, обеспечивающей необходимую ПП и ее равномерность вдоль зоны обзора [13–15]. Кроме того, в таких камерах возможно формирование зон эффективного обзора с точно «сшитым» изображением. Это решение аналогично решению 2 и может быть реализовано на этапе изготовления устройства. При этом в одном корпусе устанавливается несколько модулей камер с заранее подобранными заводскими установками, позволяющими состыковать точно формируемые ими изображения и уменьшить разброс ПП по зоне видеоконтроля. Решение 2 достаточно сложное хотя и эффективное. Главное преимущество — точно «сшитое» изображение при отсутствии необходимости сложных регулировок положения отдельных камер и подбора их параметров, а также положения (высоты установки, углов поворота и наклона). Отметим, что в рассмотренных случаях отсутствует возможность автоматизированного выбора параметров камеры для адаптации к конкретным условиям наблюдения.

В результате можно говорить о целесообразности разработки подхода, позволяющего оптимизировать распределение ПП для заданных условий установки.

Критерии оптимизации распределения

Для решения задачи оптимизации распределения ПП необходимо сформулировать критерии оптимизации. Главными и обязательными критериями являются:

- 1) обеспечение минимально требуемой D_p^0 ПП по всей зоне наблюдения $D_p^{\min} \geq D_p^0$, где D_p и D_p^0 — фактическое и минимально требуемое значения ПП;
- 2) наличие заданного допустимого диапазона изменения значений ПП $\Delta D = D_p^{\max} - D_p^{\min}$ (либо заданное максимально допустимое $D_p^{\max}$ значение ПП аналогично критерию 1).

¹ Panomera cameras [Электронный ресурс]. Режим доступа: <https://www.dallmeier.com/technology/panomera-cameras/> (дата обращения: 29.09.2022).

Также могут применяться дополнительные критерии, в том числе разные для различных фрагментов зоны обзора, выбор которых зависит от конкретной задачи, например:

- 1) минимум максимального значения $\min\{\max D_p^{\max}\}$;
- 2) минимальный диапазон изменения значений ПП $\min\{D_p^{\max} - D_p^{\min}\}$;
- 3) равнозначные максимумы значений ПП $D_p^{n\max} = D_p^{m\max}$;
- 4) равнозначные минимумы значений ПП $D_p^{n\min} = D_p^{m\min}$.

Индексы « n » и « m » соответствуют разным фрагментам зоны обзора.

Отметим, что критерии могут формулироваться для всей или отдельной части ЗЭО; любое значение ПП в любой точке ЗЭО должно быть больше или равно заданному значению ПП $D_p \geq D_p^0$, необходимому для решения поставленной задачи наблюдения; использование дополнительных критериев не исключает применения других критериев, т. е. возможны комбинации критериев, например дополнительные критерии 1 и 3 или 3 и 4.

Методика оптимизации распределения плотности пикселей

Введем основные определения, необходимые для понимания процесса методики оптимизации.

Коэффициент избыточности K_m — отношение реальной ПП D_p к требуемой D_p^0 в матрице в целом или m -ом сегменте. Значение коэффициента избыточности больше единицы соответствует излишне большой ПП, а меньше единицы — недостаточной для решения поставленной задачи наблюдения.

Условное разрешение m -го сегмента матрицы камеры $R_m^y = R_m/R_m^0$ — количество пикселей R_m^y сегмента матрицы, используемое для формирования части общего видеоизображения соответствующему этому сегменту. Условное разрешение определяется отношением: $R_m^y = R_m^0/K_m$, где R_m^0 — разрешение сегмента; K_m — коэффициент избыточности.

Нормированные (относительные) параметры для повышения общности результатов, так как ПП зависит от расстояния L от телевизионной камеры до СОН и разрешения R^0 матрицы телекамеры [1]. К таким параметрам можно отнести: нормированную ПП, служащую для решения задачи наблюдения $D_p^H = D_p/D_p^0$; нормированную относительную дальность $L^H = L/L_{\max}$, где L — расстояние от камеры до точки зоны обзора; $L_{\max}$ — расстояние до дальней границы зоны обзора. Тогда переход к реальным значениям осуществляется умножением нормированных параметров на значения реальных.

Возможность решения поставленной задачи с рассмотренными требованиями и ограничениями по сформулированным критериям основана на следующей методике.

1. Осуществлена фрагментация зоны обзора камеры в соответствии с подмножествами задач наблюдения, т. е. сформирована ЗЭО. Разрешение матрицы выбрано на основе требований по решению задачи наблюдения в дальней точке ЗЭО.

2. Для подмножеств задач наблюдения выбираются соответствующие m -е сегменты матрицы, формирующие изображение участков ЗЭО (рис. 2) протяженностью L_m . Для упрощения восприятия двумерное изображение может быть представлено не всей матрицей, а только одной ее вертикальной частью.
3. В соответствии с заданным критерием определены коэффициент избыточности и условное разрешение сегментов матрицы для формирования видеоизображения соответствующих фрагментов ЗЭО. Коэффициент избыточности определен по дальней границе соответствующего фрагмента ЗЭО.
4. В соответствии со значением коэффициента избыточности в каждом сегменте выделены группы пикселей размером $K_m \times K_m$, из которых может быть использован один сегмент или группа сегментов с последующей их обработкой по определенному алгоритму для преобразования в один.

Примеры оптимизации

Приведем примеры решения поставленной задачи для некоторых критериев. Для простоты рассмотрим случай трех задач наблюдения и, соответственно, использования трех сегментов матрицы.

Возьмем в качестве критерия комбинированный — два обязательных $D_p^{\min} \geq D_p^0$ и $\Delta D = D_p^{\max} - D_p^{\min}$, а также равные значения как максимумов $D_p^{n\max} = D_p^{m\max}$, так и минимумов $D_p^{n\min} = D_p^{m\min}$ ПП.

Для выполнения заданных критериев осуществим следующие действия.

1. Фрагментацию всей зоны обзора на участки, в пределах которых ПП изменяется в заданных пределах.
2. Сегментацию матрицы камеры на участки, относительные размеры которых определяются соответствующими относительными размерам фрагментов зоны обзора.
3. Изменение условного разрешения сегмента матрицы телекамеры путем объединения нескольких пикселей в группы. При этом количество пикселей в группе определяется требуемым разрешением в соответствующем фрагменте зоны обзора и учитывается коэффициентом избыточности K_m , равным количеству объединяемых соседних пикселей по горизонтали и вертикали. Переход от одного фрагмента к другому определяется достижением ПП D_p установленного порога $D_p^{\max}$ в соответствующем m -ом сегменте матрицы.

Выполним преобразования в соответствии с вышеизложенным порядком. Для примера рассмотрим типичный случай видеоконтроля протяженной зоны обзора с 9-кратным изменением максимальной и минимальной дальностей, приводящим к такому же изменению ПП по зоне обзора и разрешением R^0 матрицы телекамеры.

Рис. 4 иллюстрирует использование предложенной методики. На рис. 4, а показаны: характер исходного распределения ПП D_p по зоне, определяемый разрешением R^0 матрицы телекамеры и дальностью $L_{\max}$ (кривая 1, пунктирная линия); полученное распределение реальной ПП (кривая 2, сплошная линия) в результате сегментации пикселей матрицы и устранения избыточности; реальная ПП D_p^r для второго фрагмента

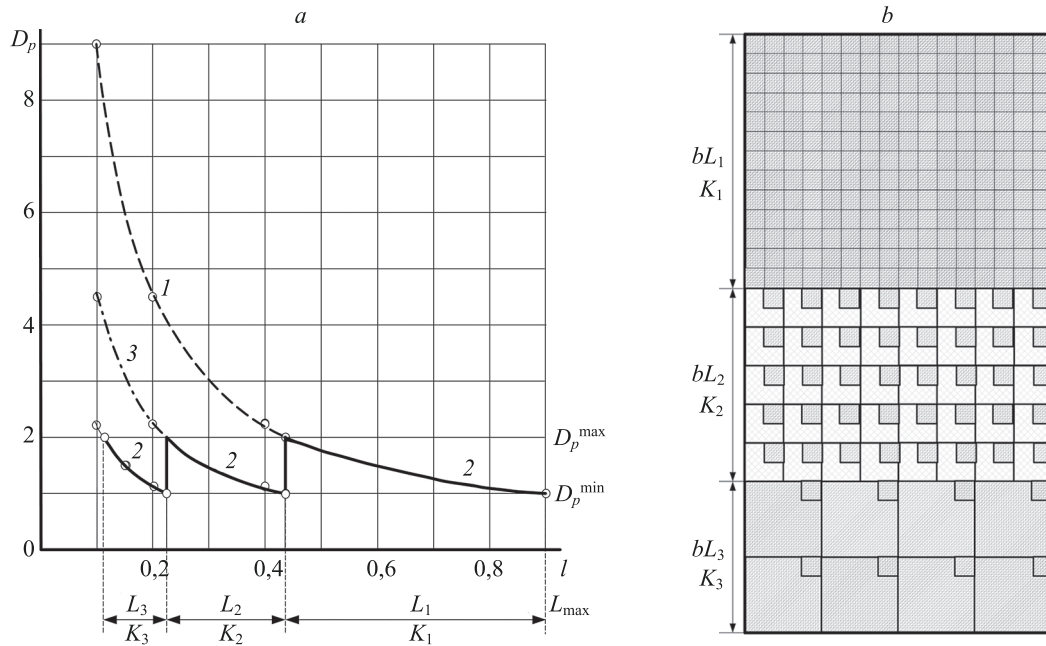


Рис. 4. Распределение плотности пикселей (а) и сегментации матрицы (b) по критерию равнозначных максимумов и минимумов

Fig. 4. Distribution of pixel density (a) and matrix segmentation (b) according to the criterion of equivalent maximums and minimums

(кривая 3, линия штрих-пунктир). На рис. 4, b показана сегментация матрицы, соответствующая фрагментации зоны обзора на участки с длинами L_1, L_2, L_3 , определяемыми выходами значений ПП за заданные пределы и соответствующие им коэффициенты избыточности $K_1 = 1, K_2 = 2, K_3 = 4$.

На рис. 4, b приведено распределение сегментов матрицы с учетом группирования пикселей на основе соответствующих коэффициентов K_n . Размеры сегмен-

тов связаны с размерами фрагментов зоны коэффициентами b_n , равными отношению размеров фрагментов и длины $\Delta L = L_{\max} - L_{\min}$ зоны обзора $b_n^s = L_n^s / \Delta L^s$, $b_n^f = L_n^f / \Delta L^f$. В первом фрагменте $D_p^r = D_p^0$. Во втором — с длиной L_2 и $K_2 = 2$ объединяются по 4 пиксела в группу (2×2). В третьем — по 16 пикселей ($K_3 = 4$). Относительный размер фрагментов зоны обзора соответствует относительным размерам сегментов матрицы $b_n^f = b_n^s = b_n$. Видно, что по сравнению с исходным

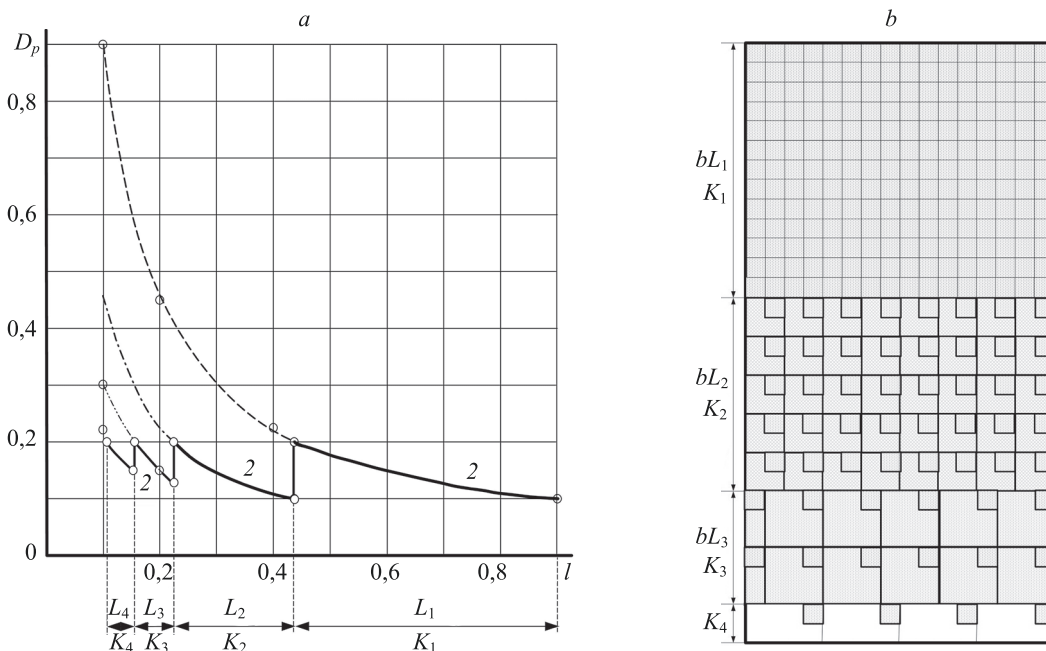


Рис. 5. Распределение плотности пикселей (а) и сегментации матрицы (b) по критерию равных значений максимумов

Fig. 5. Distribution of pixel density (a) and matrix segmentation (b) according to the criterion of equal values of maximums

распределением, ПП в разных фрагментах находится в заданных пределах, что соответствует условиям задачи и используемому критерию.

Изменим дополнительный критерий — выберем с равными значениями только максимумов ПП $D_p^{n\max} = D_p^{m\max}$. Процедура оптимизации в этом случае не будет отличаться от использованной ранее. Но в отличие от предыдущего примера минимумы не только находятся в заданном интервале от $D_p^{\min}$ до $D_p^{\max}$, но и имеют большие значения в ближних фрагментах и большую равномерность распределения рис. 5, а. Видно, что число сегментов матрицы возрастает (рис. 5, б).

Закключение

Выполнен анализ известных подходов к рассматриваемой задаче оптимизации распределения плотности пикселей и способов их реализации. Предложено представление зоны эффективного обзора как совокупности подмножеств пространств разных задач наблюдения

для различных фрагментов зоны. Предложена методика оптимизации распределения плотности пикселей по зоне эффективного обзора и сформулированы основные и дополнительные критерии оптимизации. Методика проиллюстрирована примерами ее использования с разными критериями.

Результаты работы позволяют автоматизировать процесс адаптации параметров телевизионной камеры к условиям наблюдения на основе данных о положении фрагментов зоны эффективного обзора. Тем самым минимизировать избыточность плотности пикселей и, как следствие, снизить загрузку каналов связи, объем памяти устройств хранения видеoinформации и требования к производительности устройств обработки видеосигналов. При этом автоматически решается и задача формирования непрерывного изображения контролируемой зоны.

Результаты работы могут найти применение при обработке видеосигналов в новых телекамерах систем наблюдения.

Литература

1. Волхонский В.В. Системы телевизионного наблюдения: основы проектирования и применения. М.: Горячая линия-Телеком, 2020. 392 с.
2. Волхонский В.В., Ковалевский В.А. Подход к выбору разрешения матрицы телекамеры для оптимизации параметров видеосигнала // Микро- и нанотехнологии в электронике: Материалы XII Международной научно-технической конференции. Нальчик: Каб.-Балк. ун-т, 2021. С. 412–416.
3. Волхонский В., Ковалевский В. Влияние параметров установки телекамер на информативность видеоизображения // Системы безопасности. 2022. № 2(164). С. 78–80.
4. Alpert S., Galun M., Brandt A., Basri R. Image segmentation by probabilistic bottom-up aggregation and cue integration // IEEE Transactions on Pattern Analysis and Machine Intelligence. 2012. V. 34. N 2. P. 315–327. <https://doi.org/10.1109/TPAMI.2011.130>
5. Chen C.H., Patel V.M., Chellappa R. Matrix completion for resolving label ambiguity // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015. P. 4110–4118. <https://doi.org/10.1109/CVPR.2015.7299038>
6. Wronski B., Garcia-Dorado I., Ernst M., Kelly D., Krainin M. Handheld multi-frame super-resolution // ACM Transactions on Graphics (TOG). 2019. V. 38. N 4. P. 1–18. <https://doi.org/10.1145/3306346.3323024>
7. Panda R., Roy-Chowdhury A.K. Multi-view surveillance video summarization via joint embedding and sparse optimization // IEEE Transactions on Multimedia. 2017. V. 19. N 9. P. 2010–2021. <https://doi.org/10.1109/TMM.2017.2708981>
8. Luo Z., Huang Y., Li S., Wang L., Tan T. Unfolding the alternating optimization for blind super resolution // Advances in Neural Information Processing Systems. 2020. V. 33. P. 5632–5643.
9. Su H., Yu L., Jung C. Joint contrast enhancement and noise reduction of low light images via JND transform // IEEE Transactions on Multimedia. 2020. V. 24. P. 17–32. <https://doi.org/10.1109/TMM.2020.3043106>
10. Ma L., Liu R., Wang Y., Fan X., Luo Z. Low-light image enhancement via self-reinforced retinex projection model // IEEE Transactions on Multimedia. 2022. in press. <https://doi.org/10.1109/TMM.2022.3162493>
11. Liang C.H., Chen Y.-A., Liu Y.-C., Hsu W.H. Raw image deblurring // IEEE Transactions on Multimedia. 2020. V. 24. P. 61–72. <https://doi.org/10.1109/TMM.2020.3045303>
12. Волхонский В.В., Муратов А.С. Специфика использования мультимодульных телевизионных камер // Технологии защиты. 2021. № 2. С. 52–54.
13. Dallmeier D. Monitoring Device. Patent WO 2012/164089 A1.

References

1. Volkhonsky V.V. *Video Surveillance Systems: Design and Application Fundamentals*. Moscow, Gorjachaja linija-Telekom Publ., 2020, 392 p. (in Russian)
2. Volkhonskii V.V., Kovalevskii V.A. Approach to the choice of TV-camera matrix resolution for optimizing video signal parameters. *Micro- and nanotechnologies in electronics: Proceedings of the XII International scientific and technical conference*, Nalchik, KBSU, 2021, pp. 412–416. (in Russian)
3. Volkhonsky V., Kovalevsky V. Influence of camera installation parameters on the information content of the video image. *Security and Safety*, 2022, no. 2(164), pp. 78–80. (in Russian)
4. Alpert S., Galun M., Brandt A., Basri R. Image segmentation by probabilistic bottom-up aggregation and cue integration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, vol. 34, no. 2, pp. 315–327. <https://doi.org/10.1109/TPAMI.2011.130>
5. Chen C.H., Patel V.M., Chellappa R. Matrix completion for resolving label ambiguity. *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 4110–4118. <https://doi.org/10.1109/CVPR.2015.7299038>
6. Wronski B., Garcia-Dorado I., Ernst M., Kelly D., Krainin M. Handheld multi-frame super-resolution. *ACM Transactions on Graphics (TOG)*, 2019, vol. 38, no. 4, pp. 1–18. <https://doi.org/10.1145/3306346.3323024>
7. Panda R., Roy-Chowdhury A.K. Multi-view surveillance video summarization via joint embedding and sparse optimization. *IEEE Transactions on Multimedia*, 2017, vol. 19, no. 9, pp. 2010–2021. <https://doi.org/10.1109/TMM.2017.2708981>
8. Luo Z., Huang Y., Li S., Wang L., Tan T. Unfolding the alternating optimization for blind super resolution. *Advances in Neural Information Processing Systems*, 2020, vol. 33, pp. 5632–5643.
9. Su H., Yu L., Jung C. Joint contrast enhancement and noise reduction of low light images via JND transform. *IEEE Transactions on Multimedia*, 2020, vol. 24, pp. 17–32. <https://doi.org/10.1109/TMM.2020.3043106>
10. Ma L., Liu R., Wang Y., Fan X., Luo Z. Low-light image enhancement via self-reinforced retinex projection model. *IEEE Transactions on Multimedia*, 2022, in press. <https://doi.org/10.1109/TMM.2022.3162493>
11. Liang C.H., Chen Y.-A., Liu Y.-C., Hsu W.H. Raw image deblurring. *IEEE Transactions on Multimedia*, 2020, vol. 24, pp. 61–72. <https://doi.org/10.1109/TMM.2020.3045303>
12. Volkhonskii V.V., Muratov A.S. Specific features of multimodule TV-cameras application. *Tehnologii zashchity*, 2021, no. 2, pp. 52–54. (in Russian)
13. Dallmeier D. Monitoring Device. Patent WO 2012/164089 A1.

14. Далльмайер Д. Устройство наблюдения. Патент ЕАПО 026475. 2017.
15. Stening J., Persson H., Rusz A., Bengtsson A. System for panoramic imaging. Patent US 2020/0195845 A1.

14. Dallmeier D. Monitoring Device. *Patent EAPO 026475*, 2017. (in Russian)
15. Stening J., Persson H., Rusz A., Bengtsson A. System for panoramic imaging. *Patent US 2020/0195845 A1*.

Авторы

Волхонский Владимир Владимирович — доктор технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация,  6507100798, <https://orcid.org/0000-0001-9628-2046>, volkhonski@mail.ru

Ковалевский Владислав Александрович — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-7042-7905>, Kovalvlad9@gmail.com

Authors

Vladimir V. Volkhonskiy — D. Sc., Associate Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation,  6507100798, <https://orcid.org/0000-0001-9628-2046>, volkhonski@mail.ru

Vladislav A. Kovalevskiy — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-7042-7905>, Kovalvlad9@gmail.com

Статья поступила в редакцию 25.06.2022

Одобрена после рецензирования 29.09.2022

Принята к печати 16.11.2022

Received 25.06.2022

Approved after reviewing 29.09.2022

Accepted 16.11.2022

АВТОМАТИЧЕСКОЕ УПРАВЛЕНИЕ И РОБОТОТЕХНИКА AUTOMATIC CONTROL AND ROBOTICS

doi: 10.17586/2226-1494-2022-22-6-1063-1071

Evaluation and development of a method for compensating the positioning error of computer numeric control equipment

Muhamad Albani Rizki¹✉, Yuri V. Fedosov²

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ muhamadalbanirizki@gmail.com✉, <https://orcid.org/0000-0001-7502-1699>

² Yf01@yandex.ru, <https://orcid.org/0000-0003-1869-0081>

Abstract

In the manufacturing process, Computer Numerical Control is widely used to process products that require a high level of accuracy. It is found that during product processing, Computer Numerical Control is still unable to fully counteract the influence of vibration and the presence of uneven product surfaces. In this paper, the stabilization mechanism developed, known as the Modified Stewart platform, which has a 3 Degrees of Freedom and can rotate around the X and Y axes and move translationally along the Z axis. This platform can be used to improve the accuracy and stability of the Computer Numeric Control tool. In this research, the positioning accuracy of the Modified Stewart platform has been evaluated. In this research, a mock-up or prototype and a simulation model of the Modified Stewart platform was developed. The data to be studied is the inclination angle of the platform. In the experiment, to determine the positioning error, the variable being changed acquires not only the linear movement, but also the angle of the X - Y plane. By changing the angle contained in the X - Y plane, it can be seen the influence of the X - Y angle on the position error or angle of the Z - A plane. The simulation was carried out on MATLAB. The mathematical model in this study is to find the platform position or angle. To simplify the calculation, the Modified Stewart platform was depicted in the form of a trapezoid. The results of the angle in the simulation will be compared with the result of the angle on the mock-up Modified Stewart platform. The trapezoidal parameter used in the simulation corresponds to the parameters on the mock-up modified Stewart platform. The simulation provided information about the angle of inclination, height, length of its sides, and the relative length of sides. It was found that the position or angle movement of the platform is in accordance with the calculation or simulation model that has been developed, and the positioning error data of the platform is very small and it changes constantly. It should be noted that the presented method can be used to evaluate the platform positioning error and consequential calibration of the mechanisms with spatial kinematic. The positioning error at various mobile links positions is changing, but during the movement in just one direction it remains almost constant. The position error caused by the platform mechanism can be minimized by redesigning the platform and using components that can provide a much more precise movement, moreover, using the preliminary measurements it is possible to build a table containing corrections for the control program to access the correct position of the moving platform. The accuracy and the stability of its movement can be improved and the platform can be applied to Computer Numerical Control. The method developed allows to estimate the moving platform positioning error of the mechanism with spatial kinematic. Thus, the method developed can be eliminated or compensated. It is possible to calibrate the moving platform movements in automatic mode as well.

Keywords

modified stewart platform, computer numeric control, CNC, control, vibration, stabilization

For citation: Rizki M.A., Fedosov Yu.V. Evaluation and development of a method for compensating the positioning error of computer numeric control equipment. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1063–1071. doi: 10.17586/2226-1494-2022-22-6-1063-1071

УДК 681.5.073

Оценка ошибки и разработка методики компенсации погрешности позиционирования оборудования с числовым программным управлением

Мухамад Албани Ризки¹✉, Юрий Валерьевич Федосов²

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ muhamadalbanirizki@gmail.com✉, <https://orcid.org/0000-0001-7502-1699>

² Yf01@yandex.ru, <https://orcid.org/0000-0003-1869-0081>

Аннотация

Предмет исследования. В процессе изготовления изделий с высокой степенью точности широко используется оборудование с числовым программным управлением. Известно, что в процессе изготовления изделий числовые программные устройства до сих пор не позволяют скомпенсировать влияние вибраций и связанное с ними появление неровностей на поверхности изделий. Предложена модифицированная платформа Стюарта, которая обладает тремя степенями свободы и может вращаться вдоль осей X и Y , и линейно перемещаться вдоль оси Z . Такая платформа может быть использована для увеличения точности и стабилизации в оборудовании с числовым программным управлением. Выполнена оценка точности позиционирования модифицированной платформы Стюарта. **Метод.** Разработаны макет и симуляционная модель модифицированной платформы Стюарта. В качестве данных для сравнения использованы углы наклона платформы. В ходе эксперимента для определения ошибки позиционирования изменена не только величина линейного перемещения, но и угол в плоскости $X-Y$, изменение которого показало влияние на величину ошибки выставления угла в плоскости $Z-A$. Симуляция проведена в программе MATLAB. В результате получен угол поворота платформы. Выполнено сравнение результатов расчета углов наклона модели и макета модифицированной платформы Стюарта с соответствующими размерами. Модель предоставляет данные об угле наклона, высоте, длине сторон и их относительном удлинении. **Результаты.** Обнаружено, что угловое перемещение платформы согласуется с предварительно вычисляемым положением согласно разработанной симуляционной модели, а ошибка позиционирования мала и постоянна. Предложенный метод может быть применен для оценки ошибки позиционирования платформы и дальнейшей калибровки механизмов с пространственной кинематикой. В ходе перемещения подвижного звена изменение его пространственного положения влияет на возможность перемещения остальных звеньев. При этом величина ошибки позиционирования при разных положениях подвижных звеньев изменяется. Показано, что в одном направлении погрешность остается практически неизменной. Погрешность механизма позиционирования платформы может быть уменьшена путем изменения конструкции платформы и использования деталей, обеспечивающих более точное перемещение. На основании предварительных измерений возможно построение таблицы с поправками, к которой будет обращаться управляющая программа для корректировки пространственного положения подвижной платформы. Точность и повторяемость перемещений могут быть улучшены, что позволяет применить платформу в оборудовании с числовым программным управлением. **Практическая значимость.** Разработанная методика позволяет проводить оценку погрешности позиционирования подвижной платформы механизма с пространственной кинематикой. Таким образом, ошибка позиционирования подвижной платформы может быть уменьшена, либо скомпенсирована. Представляется возможным обеспечение калибровки подвижной платформы в автоматическом режиме.

Ключевые слова

модифицированная платформа Стюарта, числовое программное управление, вибрация, стабилизация

Ссылка для цитирования: Ризки М.А., Федосов Ю.В. Оценка ошибки и разработка методики компенсации погрешности позиционирования оборудования с числовым программным управлением // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1063–1071 (на англ. яз.). doi: 10.17586/2226-1494-2022-22-6-1063-1071

Introduction

The manufacturing process is a very complex one that involves many interactions between humans and machines. In the manufacturing process of a product, the thing that is very concerned is the quality of the product [1, 2]. Currently, the manufacturing process is more modern, which means that it relies on the use of machines to process its products compared to processes carried out by humans. So the quality of the product is very dependent on the results of the production machine process. In the manufacturing process, one of the causes of decreased product quality from production machines is vibration. This vibration reduces the accuracy of the production machine so that it can reduce the quality of the manufactured product. Moreover, if the product is

measured in micron units, the accuracy of the machine is very important. This usually happens in the manufacturing process of electronic products or other sensitive products that require a high level of accuracy or a very small error. In the manufacturing process, for the processing of products that require a high level of accuracy, it is necessary to use Computer Numerical Control (CNC). However, it is found that during product processing, CNC is still unable to fully counteract the influence of vibration, thus reducing accuracy and increasing error, and reducing the quality of the product. It is also known that in addition to vibration, another factor that causes a decrease in CNC accuracy is the presence of uneven product surfaces. This is because there are differences in calculations between the program on the CNC and the actual product surface. CNC cannot know and detect the surface flatness of the product being

processed and cannot move adaptively to it [3–5]. To increase the accuracy of the CNC and reduce errors caused by vibration and uneven product surfaces, a stabilization system is needed.

Vibration stabilization systems have been widely developed and applied to various devices such as cameras. The stabilization system on the camera is needed to reduce the influence of vibration on the optical part so as to provide better results. There are several technologies applied in the development of camera optical stabilization systems. Stabilization system technology uses a voice-coil actuator mechanism [6]. In addition, there are also stabilization systems using Shape Memory Alloy actuator mechanism and piezoelectric motors [7, 8]. Previously developed stabilization systems provide good results and can reduce the influence of vibrations on the camera optic, thereby improving the image quality of the camera. However, a stabilization system that can move automatically on the camera optic mechanism has never been developed and applied to CNC. Unlike the mechanism on the camera, the stabilization system on the CNC is applied to its head. In that way it will increase the stability and resistance of the CNC to vibration so as to increase the accuracy and quality of the processed product.

CNC requires a stability mechanism that can not only withstand vibration but also can provide good results when processing products with uneven surfaces. To be able to achieve the intended stability, a mechanism that can move freely or has several degrees of freedom (DOF) is needed. It is known that there is a mechanism, called the Stewart platform that has the ability to move freely up to 6 DOFs. Stewart platform can move rotationally and translationally relating to 3 axes: X , Y , and Z [9]. With such capabilities, if the Stewart platform can be controlled properly, this mechanism can solve the problems found in CNC. Research on the Stewart platform has been carried out and applied to various types of fields. The results showed that the Stewart platform is able to provide the movement with a high level of accuracy [10–12]. Other research results show that the Stewart platform is able to provide stability and is effective in reducing the vibrations generated [13]. Therefore, based on previous research, if the Stewart platform is applied to CNC, it will increase the accuracy and stability of the vibration generated, so it will improve the quality of the processed products. However, the Stewart platform with 6 DOFs is less effective to be applied to a CNC mechanism that only moves on 2 axes (X and Y) and only requires 3 DOFs of movement from the Stewart platform, namely rotational on the X and Y axes and translational on the Z axis. Stewart platform that only has 3 DOFs are referred to as Modified Stewart platform. The Modified Stewart platform has the same working principle as the regular Stewart platform, except that the DOF movement is limited. The Modified Stewart platform has the advantage of providing simplicity and convenience in its control system [14].

Before being applied to the CNC, it is necessary to conduct further research on the mechanism of the Modified Stewart platform. Therefore, the Modified Stewart platform is the object of this research. Several experiments were conducted to determine the accuracy and stability of the

Modified Stewart platform. In this research, a mock-up or prototype and a simulation model of the Modified Stewart platform was developed. Furthermore, experiments were conducted on the mock-up and simulation model using the same parameters. The experimental data between the mock-up and simulation model will be compared and analyzed so that the Modified Stewart platform can be further evaluated. The results of the evaluation can be used to show the performance and development and improvement of the Modified Stewart platform so that later the Modified Stewart platform can be applied and improve the stability and accuracy of the CNC.

Modified Stewart platform

The Stewart platform is a manipulator device used to control position and movement. The Stewart platform consists of 2 fixed parts, the base and the platform. In addition, there are 6 legs that can be controlled in length. Originally, the Stewart platform was intended for the creation of flight simulators [15]. The Stewart platform mechanism consists of a triangular platform supported by ball joints and adjustable legs. With this mechanism, the Stewart platform has 6 DOFs. The advantage of the 6 DOFs Stewart platform is high accuracy and the ability to move easily at any desired angle and position. Therefore, Stewart platforms are widely used in structures with robotic applications such as the National Advance Driving Simulator. The angular displacement of the Stewart platform is greatly influenced by the length of the platform legs. However, the length of the platform legs also affects the length of the other legs [16]. The following figure is the scheme of the Stewart platform, Modified Stewart platform, and Design 3D of the Mock-up Modified Stewart platform (Fig. 1).

The Modified Stewart platform is an Original Stewart platform that has been modified. The Modified Stewart platform has only 4 legs, so this results in the Modified Stewart platform being able to move only in 3 DOFs. However, in general, the working principle of the Modified Stewart platform is the same as the Original Stewart platform. The movements of the Modified Stewart platform include rotating around the X and Y axes and translating along the Z axis. This is because the Modified Stewart platform will be used on CNC. The following figure is a mock-up of the Modified Stewart platform used in this research, with 3 main parts, namely: platform, legs, and base (Fig. 2).

In the mock-up Modified Stewart platform, there are several electronic components used to support its movement and mechanism, namely linear motor and Gyroscope sensor. The linear motor is used to drive the legs of the Modified Stewart platform and the Gyroscope sensor is used to detect the tilt of the platform. In the development of the Modified Stewart platform, there are parameters that have been obtained mechanically. For more details, here are the parameters of the mock-up Modified Stewart platform that has been developed (Table 1).

The linear motor movement is controlled using input steps where the legs of the Modified Stewart platform will increase in length by 0.002815 mm every 1 step. The

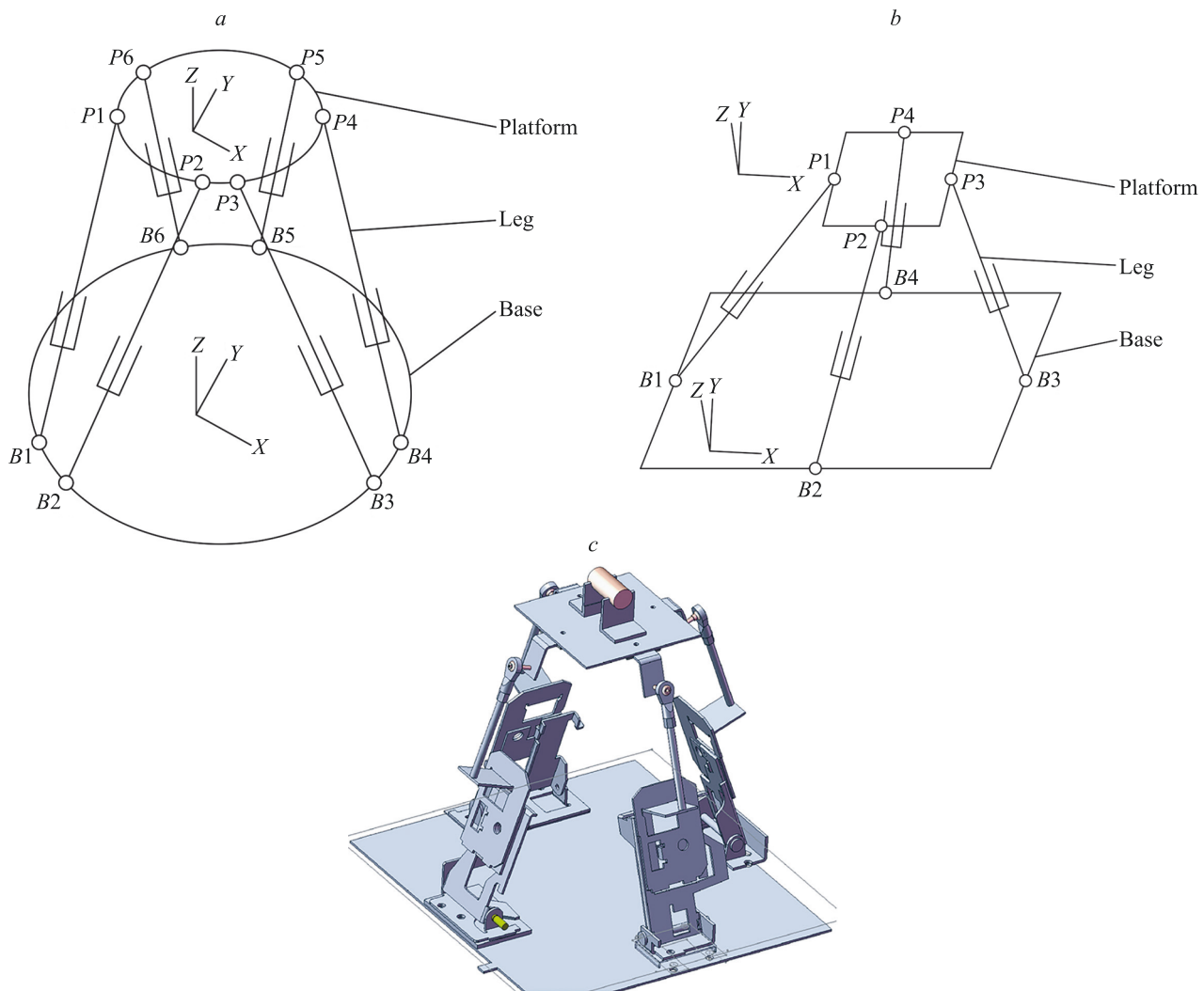


Fig. 1. Scheme of: Stewart platform (a); Modified Stewart platform (b); design 3D of mock-up Modified Stewart platform (c).
P1 – P6 is joint section in the platform, B1 – B6 is joint section in the base

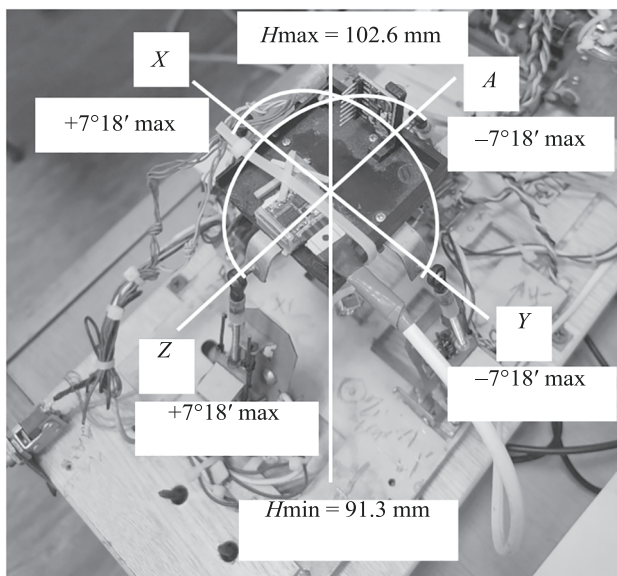


Fig. 2. Mock-up Modified Stewart platform. X, Y, Z, A: coordinates of the Modified Stewart platform with the angular rotation; Hmin and Hmax: the height of the Modified Stewart platform with translational movement

Table 1. Mock-up Modified Stewart platform parameters

Parameter	Value
Max steps	4000
Steps per, mm	0.002815
Height Maximum, mm	102.6
Height Minimum, mm	91.3
Base Length, mm	117.3
Platform Length, mm	88.6

parameters of the Modified Stewart platform will also be used as parameters of the simulation model. The platform can rotate by a maximum of 7°18' on the X and Y axes. In the Z axis, the platform can move translationally by maximum of 11.3 mm.

Mathematical model of Modified Stewart platform

To be able to make a simulation of a system, a mathematical model is needed. A mathematical model is obtained from the mechanical design and electronic system

parameters of the mock-up Modified Stewart platform. The mathematical model in this study is to find the platform position or angle. To simplify the calculation, the Modified Stewart platform is depicted in the form of a trapezoid. The following figure is a trapezoid used in the modeling of the mock-up platform (Fig. 3).

The DC and AB sides of the trapezoid always have the same size (constant). Increasing the lengths on the AD and BC sides is called delta (Δ). Delta greatly affects the angle of the trapezoid. The angle on the trapezoid is called theta (Θ). In addition to the slope of the trapezoid, the delta or the increasing of length BC and AD sides will affect the lengths of the c side. It is because the more inclined the trapezoid or platform, the longer the c side will be. The mathematical model of the modified Stewart platform can be described as follows.

To determine BC and AD sides, you can use the Pythagorean formula as follows:

$$BC^2 = (H + \Delta)^2 + (c + \Lambda)^2,$$

$$AD^2 = (H - \Delta)^2 + (c + \Lambda)^2.$$

Let us determine the height (H) of the trapezoid if the lengths of sides BC and AD are known.

$$H = \frac{\sqrt{(AD^2 - (c + \Lambda)^2)} + \sqrt{(BC^2 - (c + \Lambda)^2)}}{2}.$$

Let us now determine the angle of inclination of the platform or trapezoid Θ , then use the cosine formula. It is because the BC and AD sides have different lengths causing the platform to rotate at a fixed point. The rotation of the platform moves in the form of a circle, so determine the angle of inclination using the cosine formula. The derivation of the cosine formula from Fig. 3 of the platform is explained in the equation below:

$$\Delta^2 = \left(\frac{DC}{2}\right)^2 + (hO)^2 - 2 \cos \Theta \left(\frac{DC}{2}\right)(hO).$$

To obtain hO , we use the following formula:

$$hO = \frac{AB - 2(c + \Lambda)}{2}.$$

So, to get the angle of the platform, we use the following equation:

$$\Theta = \arccos \left(\frac{\left(\frac{DC}{2}\right)^2 + \left(\frac{AB - 2(c + \Lambda)}{2}\right)^2 - \Delta^2}{2 \left(\frac{DC}{2}\right) \left(\frac{AB - 2(c + \Lambda)}{2}\right)} \right).$$

From the equation, the inclination of the platform can be determined based on the lengths of the BC and AD sides of the trapezoid. The equation is used to create a simulation of the modified Stewart platform. The results of the angle in the simulation will be compared with the result of the angle on the mock-up Modified Stewart platform.

Simulation of Modified Stewart platform

The simulation is carried out on MATLAB which aims to be able to easily calculate and determine the angle of inclination of the platform based on the length of BC and AD sides of the trapezoid. Like on the mock-up Modified Stewart platform, input on the simulation is a signal (steps). In addition, to be able to make a simulation, it requires a reference height from the platform. In the simulation, the reference height is 97 mm. Like the mathematical model described previously, the Modified Stewart platform is calculated in the form of a trapezoid. The trapezoidal parameter used in the simulation corresponds to the parameters on the mock-up Modified Stewart platform. The following figure is a simulation program and interface of the Modified Stewart platform (Fig. 4).

The simulation can provide information about the angle of inclination, height, length of BC and AD sides, and the relative length of Bg and Ae sides. All this information is

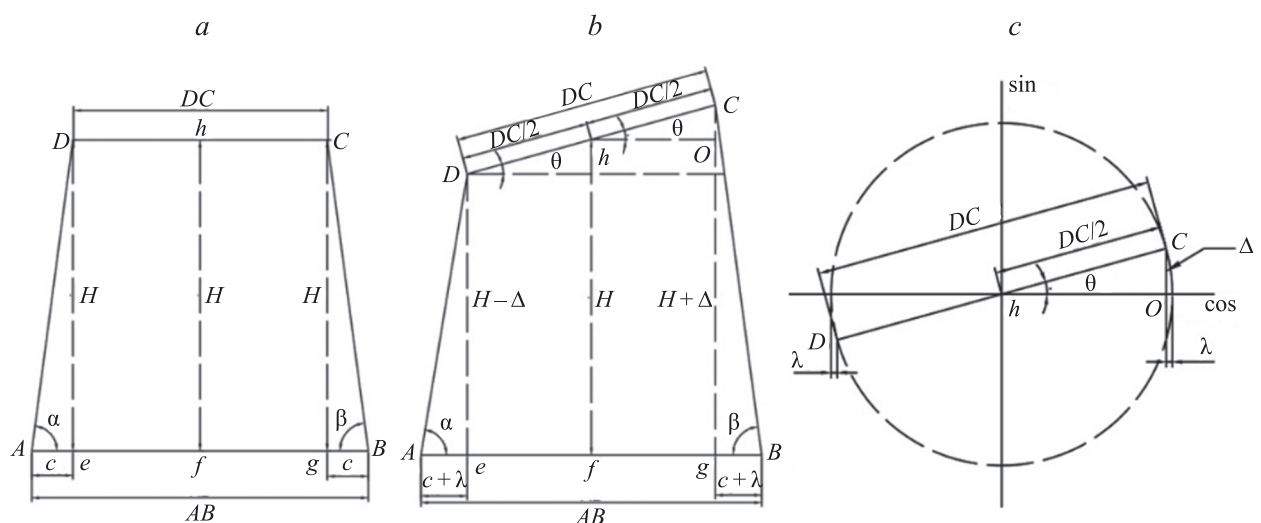


Fig. 3. Trapezoid of modeling Modified Stewart platform: normal trapezoid (a); inclined trapezoid (b); platform rotation (c); h is the center point of platform in trapezoid; H is the height of trapezoid; α is the moving angle of point DAe ; β is the moving angle of point CBg ; e , g and O are the moving point of trapezoid; f is the center point of base in trapezoid, Λ is the value that occur when platform is in inclined condition

obtained only by providing input in the form of steps. The simulation process was obtained from the mathematical model of the Modified Stewart platform that was carried out previously.

Evaluation of positioning error on Modified Stewart platform

of steps but also the angle of the X - Y plane. The value of the experimental X - Y angle is 0° ; $3^\circ 24'$; $4^\circ 12'$; $5^\circ 48'$. By changing the angle contained in the X - Y plane, it can be seen the effect of the X - Y angle on the position error or angle of the Z - A plane. The following figure is a graph of the evaluation result between the mock-up and simulation of the Modified Stewart platform with an X - Y plane (Fig. 5).

The position or angle error of the platform is caused by several factors, including mechanical factors. The Modified

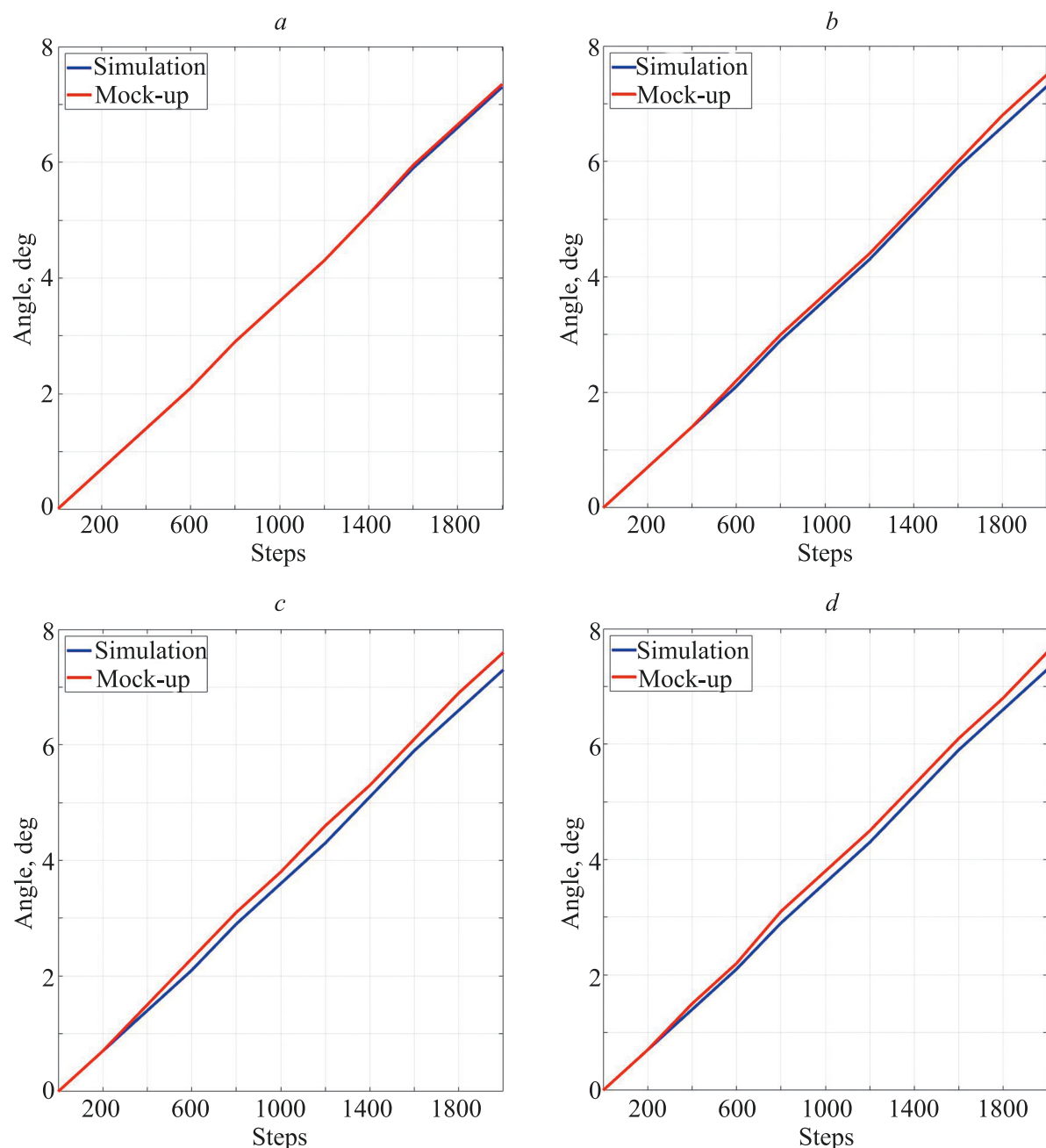


Fig. 5. Mock-up vs simulation of Modified Stewart platform with X – Y plane: 0° (a); $3^\circ 24'$ (b); $4^\circ 12'$ (c); $5^\circ 48'$ (d)

Stewart platform has 3 DOFs that can rotate around the X – Y axes and is translational along the Z axis. To perform such movements, the Modified Stewart platform has a complex mechanism. The complex movement here means that the movement of one part affects the movement of other parts. On the other hand, this complex mechanism causes inaccuracies in the position or angle of the platform itself. The position error caused by the platform mechanism can be minimized by redesigning the platform and using components that can provide a much more precise movement, in addition to calculating the mechanical connections contained in the mechanical design of the platform.

In the experiment, it was found that the angular movements X – Y affected the position or angle of Z – A , causing a position or angle error. However, the simulation

model does not allow taking into account the influence of the X – Y plane on the movement of the Z – A position. This can be possible if there is mechanical movement data of the platform or the influence of the X – Y plane on Z – A is known mathematically and is clear. However, it can be said in general that the movement generated by the platform is in accordance with the calculations and simulation models that have been developed. Admittedly, the movement of the platform is not perfect, but the error rate of the position or angle of the platform is small and it moves constantly.

The method of error compensation on the Modified Stewart platform

The Modified Stewart platform is a complex mechanism. This is obtained from the research results

where one part will affect other parts. Based on the result of the research that has been done, to be able to compensate for errors on the Modified Stewart platform, it is carried out in the following method.

1. **Calibration.** Before working with the Modified Stewart platform, all movements regarding to the X , Y , and Z axes were confirmed to be in the correct position. In this research, the X - Y and Z - A planes were at a 0° angle. Calibration cannot be done on only one side, as the position error of either side will increase the position or angle error of the Modified Stewart platform.
2. **Data acquisition.** The data collected in this research on the Modified Stewart platform is angular. From the research that has been done, data acquisition cannot be done only once. It is necessary to collect data repeatedly, intending to see the level of stability of the movement of the platform. Due to its complexity, data acquisition must be done by applying many conditions and must be done regarding to both X - Y and Z - A planes. As shown in the research, data was collected on the Z - A plane with the X - Y plane located at angles of 0° ; $3^\circ 24'$; $4^\circ 12'$; $5^\circ 48'$. This needs to be done to determine the influence of the other side on the position or angle error of the platform.
3. **Mathematical modelling and simulation.** Modelling and simulation play an important role in determining the error compensation of the Modified Stewart platform. Simulation modelling acts as a reference and verifies the angular data provided by the Modified Stewart platform. Without the results of simulation modelling, it will be very difficult to determine the position error on the platform. However, it should be noted that the Modified Stewart platform has a complex mechanism, so modelling needs to be done with in-depth calculations such as kinematic calculations and so on.
4. **Data analysis.** After the data on the Modified Stewart platform and simulation are obtained, the next step is to analyze them. Both data are compared and observed, whether angular data from the Modified Stewart platform matches the simulation data. As in the research, repeated experiments are carried out and the trend of the data generated by the Modified Stewart platform is seen. Furthermore, the trend will be analyzed to determine the equation contained in

the data. The equation will be used as a reference in developing the position control system of the Modified Stewart platform. By doing so, the position error on the platform can be reduced and the platform movement will be much more precise.

5. **Control system.** The control system is used to control the movement of the Modified Stewart platform. However, to develop the control system, equations obtained from previous experimental data are required. A good control system will provide a platform movement that matches the modelling and simulation. Also, the control system will reduce position error and increase the accuracy of the Modified Stewart platform.

Conclusion

The Modified Stewart platform is a device or platform that can move rotationally around the X and Y axes and translationally along the Z axis, in other words, it has 3 DOFs. This platform was developed to be applied to CNC to improve the accuracy and stability of the CNC head. A mock-up of the Modified Stewart platform has been developed and has been determined in accordance with the planned mechanical design. In addition, the mathematical model of the platform has been developed to determine the position or angle movement of the platform.

To determine the position or angle error of the Modified Stewart platform, an experiment was conducted by comparing the simulation model with the mock-up of the Modified Stewart platform. Several conditions were determined to identify the position error on the platform. In this research, it was found that the position or angle movement of the platform was following the calculation or simulation model. The movement of the Modified Stewart platform is so complex that the movement of one part also affects the movement of other parts. This causes position or angle errors to occur on the platform. The movement mechanism of the Modified Stewart platform needs to be considered and taken into account to minimize the position or angle error in its movement. A position error compensation method has also been developed, with the aim that it can later be used to evaluate complex mechanisms, such as the Modified Stewart platform, which can have high accuracy and stable movement that can be applied to CNC.

References

1. Shih M.-H., Sung W.-P., Chen C.-L. Vibration control and shock absorption techniques for Hi-Tech manufacturing plants. *Structural Design of Tall and Special Buildings*, 2012, vol. 21, no. 7, pp. 505–523. <https://doi.org/10.1002/tal.625>
2. Al-Shayea A., Abdullah F.M., Noman M.A., Kaid H., Abouel Nasr E. Studying and optimizing the effect of process parameters on machining vibration in turning process of AISI 1040 steel. *Advances in Materials Science and Engineering*, 2020, vol. 2020, pp. 5480614. <https://doi.org/10.1155/2020/5480614>
3. Chang C.-C., Hsia S.-Y., Huang H.-D. Improvement on CNC drilling quality using vibration suppression fixture. *Proc. of the 4th IEEE International Conference on Applied System Innovation (ICASI)*, 2018, pp. 910–913. <https://doi.org/10.1109/ICASI.2018.8394415>

Литература

1. Shih M.-H., Sung W.-P., Chen C.-L. Vibration control and shock absorption techniques for Hi-Tech manufacturing plants // *Structural Design of Tall and Special Buildings*. 2012. V. 21. N 7. P. 505–523. <https://doi.org/10.1002/tal.625>
2. Al-Shayea A., Abdullah F.M., Noman M.A., Kaid H., Abouel Nasr E. Studying and optimizing the effect of process parameters on machining vibration in turning process of AISI 1040 steel // *Advances in Materials Science and Engineering*. 2020. V. 2020. P. 5480614. <https://doi.org/10.1155/2020/5480614>
3. Chang C.-C., Hsia S.-Y., Huang H.-D. Improvement on CNC drilling quality using vibration suppression fixture // *Proc. of the 4th IEEE International Conference on Applied System Innovation (ICASI)*. 2018. P. 910–913. <https://doi.org/10.1109/ICASI.2018.8394415>

4. Wang H., Huang Z., Wang C. Design of vibration test system for CNC milling machine // *ACM International Conference Proceeding Series*, 2020, pp. 279–284. <https://doi.org/10.1145/3436286.3436406>
5. Qiu C., Chen X., Hui Y., Siddiquee T.A.R. Study of dynamic vibration characteristics and suppression of CNC machine tool during operation. *Journal of Vibroengineering*, 2020, vol. 22, no. 8, pp. 1884–1895. <https://doi.org/10.21595/jve.2020.21630>
6. Song M.-G., Baek H.-W., Park N.-C., Park K.-S., Yoon T., Park Y.-P., Lim S.-C. Development of small sized actuator with compliant mechanism for optical image stabilization. *IEEE Transactions on Magnetics*, 2010, vol. 46, no. 6, pp. 2369–2372. <https://doi.org/10.1109/TMAG.2010.2042288>
7. Kazi A., Honold M., Rimkus W., Lokner T., Bäuml M., Köpfer M. SMA actuator for optical image stabilization. *ACTUATOR 2018 — 16th International Conference and Exhibition on New Actuators and Drive Systems, Conference Proceedings*, 2018, pp. 375–378.
8. Karev P.V. Optical stabilization and microscanning with piezo actuators and piezoelectric motors. *Proc. of the International Conference Laser Optics 2018 (ICLO)*, 2018, pp. 192. <https://doi.org/10.1109/LO.2018.8435895>
9. Dasgupta B., Mruthyunjaya T.S. Stewart platform manipulator: A review. *Mechanism and Machine Theory*, 2000, vol. 35, no. 1, pp. 15–40. [https://doi.org/10.1016/S0094-114X\(99\)00006-3](https://doi.org/10.1016/S0094-114X(99)00006-3)
10. Yufei X., He L., Yanbin Z., Fei X., Jing Z. Dynamic modeling and high accuracy attitude control of a Stewart spacecraft. *Proc. of the 29th Chinese Control and Decision Conference (CCDC)*, 2017, pp. 7395–7400. <https://doi.org/10.1109/CCDC.2017.7978522>
11. Noskiewicz P., Walica D. Design and realisation of the simulation model of the Stewart platform using the MATLAB-Simulink and the Simscape Multibody library. *Proc. of the 21st IEEE International Carpathian Control Conference (ICCC)*, 2020, pp. 9257249. <https://doi.org/10.1109/ICCC49264.2020.9257249>
12. McCann C., Patel V., Dollar A. The Stewart hand: A highly dexterous, six-degrees-of-freedom manipulator based on the stewart-gough platform. *IEEE Robotics and Automation Magazine*, 2021, vol. 28, no. 2, pp. 23–36. <https://doi.org/10.1109/MRA.2021.3064750>
13. Bi F., Ma T., Wang X., Yang X., Lv Z. Research on vibration control of seating system platform based on the cubic stewart parallel mechanism. *IEEE Access*, 2019, vol. 7, pp. 155637–155649. <https://doi.org/10.1109/ACCESS.2019.2948785>
14. Fedosov Y.V., Afanasev M.Y. Design of an adaptive system for stabilization of a laser beam for CNC machine. *Proc. of the 19th Conference of Open Innovation Association, FRUCT*, 2017, pp. 31–36. <https://doi.org/10.23919/FRUCT.2016.7892180>
15. Guo H.B., Li H.R. Dynamic analysis and simulation of a six degree of freedom Stewart platform manipulator. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2006, vol. 220, no. 1, pp. 61–72. <https://doi.org/10.1243/095440605X32075>
16. Şumnu A., Güzelbey İ.H., Çakır M.V. Simulation and PID control of a Stewart platform with linear motor. *Journal of Mechanical Science and Technology*, 2017, vol. 31, no. 1, pp. 345–356. <https://doi.org/10.1007/s12206-016-1238-7>
4. Wang H., Huang Z., Wang C. Design of vibration test system for CNC milling machine // *ACM International Conference Proceeding Series*, 2020. P. 279–284. <https://doi.org/10.1145/3436286.3436406>
5. Qiu C., Chen X., Hui Y., Siddiquee T.A.R. Study of dynamic vibration characteristics and suppression of CNC machine tool during operation // *Journal of Vibroengineering*, 2020. V. 22. N 8. P. 1884–1895. <https://doi.org/10.21595/jve.2020.21630>
6. Song M.-G., Baek H.-W., Park N.-C., Park K.-S., Yoon T., Park Y.-P., Lim S.-C. Development of small sized actuator with compliant mechanism for optical image stabilization // *IEEE Transactions on Magnetics*, 2010. V. 46. N 6. P. 2369–2372. <https://doi.org/10.1109/TMAG.2010.2042288>
7. Kazi A., Honold M., Rimkus W., Lokner T., Bäuml M., Köpfer M. SMA actuator for optical image stabilization // *ACTUATOR 2018 — 16th International Conference and Exhibition on New Actuators and Drive Systems, Conference Proceedings*, 2018. P. 375–378.
8. Karev P.V. Optical stabilization and microscanning with piezo actuators and piezoelectric motors // *Proc. of the International Conference Laser Optics 2018 (ICLO)*, 2018. P. 192. <https://doi.org/10.1109/LO.2018.8435895>
9. Dasgupta B., Mruthyunjaya T.S. Stewart platform manipulator: A review // *Mechanism and Machine Theory*, 2000. V. 35. N 1. P. 15–40. [https://doi.org/10.1016/S0094-114X\(99\)00006-3](https://doi.org/10.1016/S0094-114X(99)00006-3)
10. Yufei X., He L., Yanbin Z., Fei X., Jing Z. Dynamic modeling and high accuracy attitude control of a Stewart spacecraft // *Proc. of the 29th Chinese Control and Decision Conference (CCDC)*, 2017. P. 7395–7400. <https://doi.org/10.1109/CCDC.2017.7978522>
11. Noskiewicz P., Walica D. Design and realisation of the simulation model of the Stewart platform using the MATLAB-Simulink and the Simscape Multibody library // *Proc. of the 21st IEEE International Carpathian Control Conference (ICCC)*, 2020. P. 9257249. <https://doi.org/10.1109/ICCC49264.2020.9257249>
12. McCann C., Patel V., Dollar A. The Stewart hand: A highly dexterous, six-degrees-of-freedom manipulator based on the stewart-gough platform // *IEEE Robotics and Automation Magazine*, 2021. V. 28. N 2. P. 23–36. <https://doi.org/10.1109/MRA.2021.3064750>
13. Bi F., Ma T., Wang X., Yang X., Lv Z. Research on vibration control of seating system platform based on the cubic stewart parallel mechanism // *IEEE Access*, 2019. V. 7. P. 155637–155649. <https://doi.org/10.1109/ACCESS.2019.2948785>
14. Fedosov Y.V., Afanasev M.Y. Design of an adaptive system for stabilization of a laser beam for CNC machine // *Proc. of the 19th Conference of Open Innovation Association, FRUCT*, 2017. P. 31–36. <https://doi.org/10.23919/FRUCT.2016.7892180>
15. Guo H.B., Li H.R. Dynamic analysis and simulation of a six degree of freedom Stewart platform manipulator // *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2006. V. 220. N 1. P. 61–72. <https://doi.org/10.1243/095440605X32075>
16. Şumnu A., Güzelbey İ.H., Çakır M.V. Simulation and PID control of a Stewart platform with linear motor // *Journal of Mechanical Science and Technology*, 2017. V. 31. N 1. P. 345–356. <https://doi.org/10.1007/s12206-016-1238-7>

Authors

Muhamad Albani Rizki — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-7502-1699>, muhamadbanirizki@gmail.com

Yuri V. Fedosov — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0003-1869-0081>, Yf01@yandex.ru

Авторы

Ризки Мухамад Албани — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-7502-1699>, muhamadbanirizki@gmail.com

Федосов Юрий Валерьевич — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0003-1869-0081>, Yf01@yandex.ru

Received 04.07.2022

Approved after reviewing 25.09.2022

Accepted 09.11.2022

Статья поступила в редакцию 04.07.2022

Одобрена после рецензирования 25.09.2022

Принята к печати 09.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1072-1077

УДК 519.71

Компенсация внешних возмущений по выходу для класса линейных систем с запаздыванием в канале управления

Ван Хуан Буй¹, Алексей Анатольевич Маргун²✉

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

² Институт проблем машиностроения РАН, Санкт-Петербург, 199178, Российская Федерация

¹ buinguyenkhanh201095@gmail.com, <https://orcid.org/0000-0002-6563-1909>

² alexeimargun@gmail.com✉, <https://orcid.org/0000-0002-5333-0594>

Аннотация

Предмет исследования. Рассмотрена задача компенсации внешних неизвестных возмущений по выходу при неизмеримом векторе состояния для класса линейных систем с запаздыванием в канале управления. Предполагается, что возмущение является выходом автономного линейного генератора. **Метод.** Для оценки возмущения построен специальный наблюдатель. На основе оценок наблюдателя сформирована система с расширенным вектором состояния, для которой построен регулятор, обеспечивающий компенсацию возмущения. **Основные результаты.** Представлен метод компенсации по выходу внешних возмущений для класса линейных систем с входным запаздыванием. Предложенный подход не требует идентификации параметров возмущения. Работоспособность полученного результата подтверждена с использованием компьютерного моделирования в среде MATLAB Simulink. **Практическая значимость.** Разработанный алгоритм может быть эффективно применен для класса задач, связанных с компенсацией качки в корабельных системах, в управлении движением робототехнических комплексов различного вида и других.

Ключевые слова

внешнее возмущение, линейные системы, адаптивное управление, компенсация возмущений, запаздывание

Благодарности

Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации, госзадание № 2019-0898.

Ссылка для цитирования: Буй В.Х., Маргун А.А. Компенсация внешних возмущений по выходу для класса линейных систем с запаздыванием в канале управления // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1072–1077. doi: 10.17586/2226-1494-2022-22-6-1072-1077

Compensation of output external disturbances for a class of linear systems with control delay

Van Huan Bui¹, Alexey A. Margun²✉

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

² Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

¹ buinguyenkhanh201095@gmail.com, <https://orcid.org/0000-0002-6563-1909>

² alexeimargun@gmail.com✉, <https://orcid.org/0000-0002-5333-0594>

Abstract

The paper considers the problem of the output external unknown disturbance compensation under unmeasurable state vector for a class of linear systems with the control channel delay. It is assumed that the disturbance is the output of an autonomous linear generator. A special observer was built to estimate the disturbance. A system with an extended state vector is formed on the base of the observer's estimates. A controller that provides disturbance compensation is

proposed. An algorithm for the output external disturbances compensation for a class of linear systems with input delay is presented. This method does not require identification of disturbance parameters. The performance of the proposed algorithm was confirmed using computer simulation in the MATLAB Simulink software. The developed algorithm can be effectively applied to a class of problems related to rocking compensation in ship systems, control of robotic complexes various kinds, etc.

Keywords

external disturbances, linear systems, adaptive control, disturbance compensation, delay

Acknowledgements

The study was supported by the Ministry of Science and Higher Education of the Russian Federation, state assignment No. 2019-0898.

For citation: Bui V.H., Margun A.A. Compensation of output external disturbances for a class of linear systems with control delay. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1072–1077 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1072-1077

Введение

В последние десятилетия с быстрым развитием науки и техники область применения систем автоматического управления становится все более и более широкой. Одна из важнейших задач при проектировании систем управления — задача компенсации возмущений, например, в робототехнических комплексах, беспилотных летательных аппаратах, космических аппаратах, гидравлических системах и т. д. [1–4]. Известен ряд подходов для решения данной задачи. Основные результаты исследований и разработок предполагают построение независимых алгоритмов идентификации, оценивающих фазы, частоты и амплитуды возмущений, представляемых в виде мультигармонических сигналов. Преимущество данных методов заключается в том, что работа идентификатора не зависит от регулятора, что позволяет применять различные методы управления и компенсации [5–8]. При этом на практике, при отсутствии незатухающего возбуждения регрессора, применение алгоритмов идентификации оказывается неэффективным [9]. С другой стороны, некоторые исследования используют подход прямой адаптивной компенсации [10–14]. Как правило, для этого строят наблюдатель внешних возмущений и блок регулятора, использующий оценку наблюдателя и вектор состояния объекта управления. Отметим, что часто вектор состояния недоступен для измерения, а установка дополнительных датчиков является дорогостоящей или невозможна из-за технологических ограничений.

В данной работе для решения описанной проблемы предложен метод адаптивного управления, обеспечивающий компенсацию внешних возмущений по измеримому выходу системы в условиях запаздывания в канале управления. Для оценки возмущения построен специальный наблюдатель. Компенсацию обеспечивает регулятор, использующий расширенную модель системы, и алгоритм адаптации.

Выполним построение наблюдателя внешних возмущений и синтез закона управления и алгоритма адаптации.

Для подтверждения работоспособности предложенного решения осуществим компьютерное моделирование с использованием программной среды MATLAB Simulink.

Постановка задачи

Рассмотрим класс линейных устойчивых возмущенных объектов управления вида:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t - \tau) + \mathbf{d}f(t), \\ y(t) = \mathbf{C}^T\mathbf{x}(t), \end{cases} \quad (1)$$

где $\mathbf{x}(t) \in \mathbb{R}^n$ — неизмеримый вектор состояния объекта управления; $u(t - \tau) \in \mathbb{R}$ — сигнал управления с запаздыванием; $y(t) \in \mathbb{R}$ — измеряемый выход объекта управления; $f(t) \in \mathbb{R}$ — неизмеримое ограниченное внешнее возмущение; $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{B} \in \mathbb{R}^n$, $\mathbf{C} \in \mathbb{R}^n$, $\mathbf{d} \in \mathbb{R}^n$ — известные постоянные матрицы; τ — известное постоянное запаздывание.

Предположим, что внешнее возмущение $f(t)$ представляет собой мультигармонический сигнал:

$$f(t) = \sum_{i=0}^n A_i \sin(\omega_i t + \varphi_i) + A_{0i},$$

где неизвестные $A_1, \dots, A_i$ — амплитуды; $\omega_1, \dots, \omega_i$ — частоты; $\varphi_1, \dots, \varphi_i$ — фазы; $A_{01}, \dots, A_{0i}$ — смещения.

Требуется обеспечить компенсацию внешнего возмущения по выходу с учетом следующих допущений.

Допущение 1. Пары известных постоянных матриц $(\mathbf{A}, \mathbf{B})$ и $(\mathbf{A}, \mathbf{C})$ управляемы и наблюдаемы соответственно. Матрица $\mathbf{A}$ — гурвицева.

Допущение 2. Значения вектора состояния объекта неизмеримы и доступны только сигналы $u(t)$ и $y(t)$.

Допущение 3. Внешнее возмущение $f(t)$ ограничено и может быть представлено как выход линейного автономного генератора с неизвестными параметрами [15]:

$$\begin{cases} \dot{\mathbf{z}}(t) = \mathbf{\Gamma}\mathbf{z}(t) \\ f(t) = \mathbf{h}^T\mathbf{z}(t) \end{cases} \Rightarrow \begin{cases} \dot{\xi}(t) = \mathbf{G}\xi(t) + \mathbf{L}f(t), \\ f(t) = \mathbf{\theta}^T\xi(t) \end{cases}, \quad (2)$$

где $\mathbf{z}(t) \in \mathbb{R}^q$ — неизмеримый вектор состояния генератора возмущения; $\mathbf{\Gamma} \in \mathbb{R}^{q \times q}$ — неизвестная постоянная матрица, собственные числа которой некрatны и лежат на мнимой оси; $\mathbf{h}^T \in \mathbb{R}^{1 \times q}$ — неизвестный постоянный вектор; $\xi(t)$ — регрессор; $\mathbf{G}$ — гурвицева матрица; $\mathbf{L}$ — постоянный вектор; $\mathbf{\theta}^T$ — вектор неизвестных постоянных параметров, зависящий от параметров возмущения. Без потери общности предположим, что пара $(\mathbf{\Gamma}, \mathbf{h}^T)$ полностью наблюдаема и размерность автономного генератора q известна. $\mathbf{G}$ и $\mathbf{L}$ произвольно выбраны так, чтобы пара $(\mathbf{G}, \mathbf{L})$ являлась полностью управляемой.

Таким образом, внешнее возмущение может быть рассмотрено в качестве мультигармонического сигнала с неизвестными параметрами, но с известным ограниченным количеством гармоник.

Наблюдатель возмущения

Так как значение вектора состояния объекта недоступно, известны только значения сигнала управления $u(t - \tau)$ и выхода объекта $y(t)$, построим наблюдатель состояния Люенбергера полного порядка вида:

$$\begin{cases} \dot{\hat{\mathbf{x}}} = \mathbf{A}\hat{\mathbf{x}} + \mathbf{B}u(t - \tau) + \mathbf{K}(y - \hat{y}) \\ \hat{y} = \mathbf{C}^T\hat{\mathbf{x}} \end{cases}, \quad (3)$$

где $\hat{\mathbf{x}} \in \mathbb{R}^n$ — оценка вектора состояния; $\hat{y} \in \mathbb{R}$ — оценка значения выхода; $\mathbf{K}$ — матрица наблюдателя, определяемая разработчиком.

Обозначим ошибку оценки состояния $\mathbf{e}_x = \mathbf{x} - \hat{\mathbf{x}}$. Динамическая модель ошибки наблюдения имеет вид:

$$\begin{cases} \dot{\mathbf{e}}_x = \dot{\mathbf{x}} - \dot{\hat{\mathbf{x}}} = \underbrace{(\mathbf{A} - \mathbf{K}\mathbf{C}^T)}_{\mathbf{A}_e} \mathbf{e}_x + \mathbf{d}f \\ \varepsilon = y - \hat{y} = \mathbf{C}^T \mathbf{e}_x \end{cases},$$

где $\mathbf{A}_e = \mathbf{A} - \mathbf{K}\mathbf{C}^T$, матрица $\mathbf{K}$ выбрана таким образом, чтобы матрица $\mathbf{A}_e$ являлась гурвицевой (для обеспечения устойчивости наблюдателя). Из системы уравнений (1) и (3) получим

$$\varepsilon = \mathbf{W}_e(s)[f], \quad (4)$$

где ε — отфильтрованное согласованное возмущение; $\mathbf{W}_e(s) = \mathbf{C}^T(\mathbf{s}\mathbf{I} - \mathbf{A}_e)^{-1}\mathbf{d}$ — асимптотически устойчивая передаточная функция с гурвицевым знаменателем. В (2) показано, что возмущение может быть представлено в виде линейной регрессии $f(t) = \boldsymbol{\theta}^T \boldsymbol{\xi}(t)$. Перепишем уравнение (4) в виде $\varepsilon = \mathbf{W}_e(s)[\boldsymbol{\theta}^T \boldsymbol{\xi}]$ и вынесем матрицу $\boldsymbol{\theta}^T$ за скобки:

$$\varepsilon = \boldsymbol{\theta}^T \mathbf{W}_e(s)[\boldsymbol{\xi}] = \boldsymbol{\theta}^T \boldsymbol{\xi}_f, \quad (5)$$

где $\boldsymbol{\xi}_f = \mathbf{W}_e(s)[\boldsymbol{\xi}]$ или $\boldsymbol{\xi}_f = \frac{\alpha(s)}{\beta(s)}[\boldsymbol{\xi}]$; $\alpha(s)$, $\beta(s)$ — полиномы с известными постоянными коэффициентами. С другой

стороны, $\dot{\boldsymbol{\xi}} = \frac{(\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)\boldsymbol{\xi}}{\mathbf{Q}}$ и $\dot{\boldsymbol{\xi}}_f = \frac{(\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)\boldsymbol{\xi}_f}{\mathbf{Q}}$. Используя

преобразование Лапласа, получим $\begin{cases} s\dot{\boldsymbol{\xi}} = \mathbf{Q}\dot{\boldsymbol{\xi}} \\ s\dot{\boldsymbol{\xi}}_f = \mathbf{Q}\dot{\boldsymbol{\xi}}_f \end{cases}$ тогда

$\begin{cases} \beta(s)[\dots] = \beta(\mathbf{Q})[\dots] \\ \alpha(s)[\dots] = \alpha(\mathbf{Q})[\dots] \end{cases}$. Откуда после ряда преобразова-

ний $\frac{\beta(s)}{\alpha(s)}\boldsymbol{\xi}_f = \boldsymbol{\xi}$. Обозначим $\bar{\mathbf{Q}}^{-1} = \frac{\beta(s)}{\alpha(s)}$, тогда:

$$\bar{\mathbf{Q}}^{-1}\boldsymbol{\xi}_f = \boldsymbol{\xi}. \quad (6)$$

С учетом системы уравнений (2) и выражения (5) сформируем следующий наблюдатель возмущения

$$\dot{\hat{\boldsymbol{\xi}}}_f = \mathbf{G}\hat{\boldsymbol{\xi}}_f = \mathbf{L}(y - \hat{y}), \quad (7)$$

где $\hat{\boldsymbol{\xi}}_f$ — вектор состояния наблюдателя с произвольными начальными условиями.

Наблюдатель (7) может быть представлен в виде автономной модели:

$$\dot{\hat{\boldsymbol{\xi}}}_f(t) = (\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)\hat{\boldsymbol{\xi}}_f(t). \quad (8)$$

Решая уравнение (8) [6], получим

$$\hat{\boldsymbol{\xi}}_f(t - \tau) = \mathbf{P}^{-1}\hat{\boldsymbol{\xi}}_f(t), \quad (9)$$

где $\mathbf{P} = e^{(\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)\tau}$.

Далее требуется скомпенсировать внешнее возмущение одновременно устраняя негативное влияние эффекта запаздывания.

Синтез закона управления и алгоритма адаптации

Для компенсации внешних возмущений построим регулятор на основе работы [16]. Переведем координаты внешних возмущений в систему координат вектора состояния объекта, используя матрицу преобразования $\mathbf{M}$. Ошибка параметрического отслеживания состояния объекта примет вид

$$e(t) = \mathbf{x}(t) - \mathbf{M}\boldsymbol{\xi}(t). \quad (10)$$

Продифференцировав (10), получим

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

Если спектры матриц $\mathbf{A}$ ($\text{Re}\lambda_i < 0$) и $\bar{\mathbf{G}} = (\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)(\pm j\omega_i)$ не пересекаются, то существует значение $\boldsymbol{\Psi}$ такое, что

$$\mathbf{A}\mathbf{M} - \mathbf{M}(\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T) = \mathbf{B}\boldsymbol{\Psi}^T - \mathbf{d}\boldsymbol{\theta}^T. \quad (11)$$

Выходной сигнал объекта управления имеет вид

$$y = \mathbf{C}^T \mathbf{e} + \mathbf{C}^T \mathbf{M}\boldsymbol{\xi}(t). \quad (12)$$

Существуют матрицы $\mathbf{M}$ и $\boldsymbol{\Psi}$, которые являются решениями уравнений (11)–(12) и называются уравнениями Франсиса или регулятора [12]. Следовательно, можно сделать вывод, что уравнения (11)–(12) имеют хотя бы одно решение, если спектры матриц $\mathbf{A}$ ($\text{Re}\lambda_i < 0$) и $\bar{\mathbf{G}} = (\mathbf{G} + \mathbf{L}\boldsymbol{\theta}^T)(\pm j\omega_i)$ не пересекаются.

Получим модель ошибки

$$\begin{cases} \dot{e} = \mathbf{A}e + \mathbf{B}[\boldsymbol{\Psi}^T \boldsymbol{\xi}(t) + u(t - \tau)] \\ y = \mathbf{C}^T e \end{cases}. \quad (13)$$

С учетом уравнения (6) система уравнений (13) примет вид

$$\begin{cases} \dot{e} = \mathbf{A}e + \mathbf{B}[\bar{\boldsymbol{\Psi}}^T \boldsymbol{\xi}_f(t) + u(t - \tau)] \\ y = \mathbf{C}^T e \end{cases}, \quad (14)$$

где $\bar{\boldsymbol{\Psi}}^T = \boldsymbol{\Psi}^T \bar{\mathbf{Q}}^{-1}$.

Подставив выражение (9) в (14), получим

$$\begin{cases} \dot{e} = \mathbf{A}e + \mathbf{B}[\boldsymbol{\eta}^T \hat{\boldsymbol{\xi}}_f(t - \tau) + u(t - \tau)] \\ y = \mathbf{C}^T e \end{cases},$$

где $\boldsymbol{\eta}^T = \bar{\boldsymbol{\Psi}}^T \mathbf{P}$. Выберем закон управления $u = -\hat{\boldsymbol{\eta}}^T \hat{\boldsymbol{\xi}}_f$ и получим модель ошибки замкнутой системы:

$$\begin{cases} \dot{\mathbf{e}} = \mathbf{A}\mathbf{e} + \mathbf{B}\tilde{\boldsymbol{\eta}}^T(t-\tau)\hat{\xi}_f(t-\tau), \\ y = \mathbf{C}^T\mathbf{e} \end{cases}, \quad (15)$$

где $\tilde{\boldsymbol{\eta}}^T(t-\tau) = \boldsymbol{\eta}^T - \hat{\boldsymbol{\eta}}^T(t-\tau)$ — параметрическая ошибка.

Модель ошибки и синтез алгоритма адаптации

Для устранения негативного влияния запаздывания сформируем расширенный вектор состояния:

$$\hat{\mathbf{e}} = \mathbf{e} + \boldsymbol{\chi}, \quad (16)$$

где сигнал $\boldsymbol{\chi}$ определяется по формуле

$$\dot{\boldsymbol{\chi}} = \mathbf{A}\boldsymbol{\chi} + \mathbf{B}(\hat{\boldsymbol{\eta}}^T(t-\tau) - \hat{\boldsymbol{\eta}}^T)\hat{\xi}_f(t-\tau). \quad (17)$$

Подставив систему уравнений (15) и выражение (17) в (16), после ряда преобразований получим

$$\begin{cases} \dot{\hat{\mathbf{e}}} = \hat{\mathbf{A}}\hat{\mathbf{e}} + \mathbf{B}\tilde{\boldsymbol{\eta}}^T\hat{\xi}_f(t-\tau), \\ y = \mathbf{C}^T\hat{\mathbf{e}}. \end{cases} \quad (18)$$

Воспользуемся методом расширенной ошибки. Перепишем (18) в виде

$$y = \mathbf{W}(s)[\tilde{\boldsymbol{\eta}}^T\hat{\xi}_f(t-\tau)],$$

где $\mathbf{W}(s) = \mathbf{C}^T(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$.

$$\begin{aligned} \hat{y} &= y - \tilde{y}, \\ \hat{y} &= y - \hat{\boldsymbol{\eta}}^T\mathbf{W}(s)[\hat{\xi}_f(t-\tau)] - \mathbf{W}(s)[u(t-\tau)] = \\ &= \tilde{\boldsymbol{\eta}}^T\mathbf{W}(s)[\hat{\xi}_f(t-\tau)]. \end{aligned} \quad (19)$$

На основе уравнения (19) построим стандартный алгоритм адаптации:

$$\dot{\hat{\boldsymbol{\eta}}} = \gamma\mathbf{W}(s)[\tilde{\boldsymbol{\eta}}\hat{\xi}_f(t-\tau)]\hat{y}.$$

В результате можно сделать вывод, что если внешние возмущения устойчивы, ограничены и известно

число гармоник, то с выбором коэффициента адаптации $\gamma > 0$ следует обеспечение асимптотической сходимости вектора параметрических ошибок $\tilde{\boldsymbol{\eta}}(t)$ к нулю при $t \rightarrow \infty$.

Математическое моделирование

Для проверки работоспособности и эффективности предложенного подхода выполним компьютерное моделирование с использованием программной среды MATLAB Simulink.

Рассмотрим объект управления второго порядка, имеющий вид:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}u(t-\tau) + \mathbf{d}f(t), \\ y(t) = \mathbf{C}^T\mathbf{x}(t) \end{cases}, \quad \mathbf{x}(0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

$$\text{где } \mathbf{A} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \mathbf{C}^T = [1 \quad 0].$$

Пусть возмущающее воздействие имеет вид: $f(t) = 5\sin(t)$.

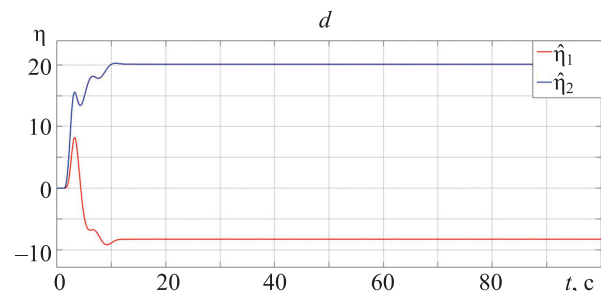
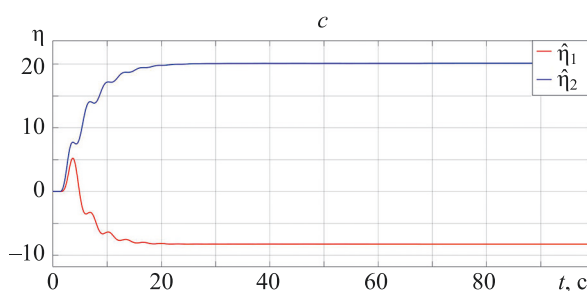
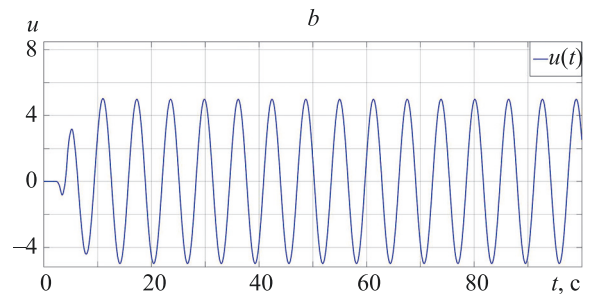
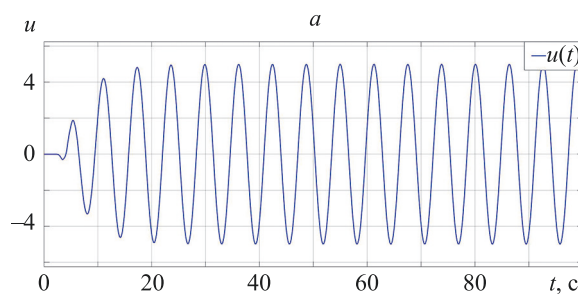
Построим наблюдатель Люенбергера (4) со следующей матрицей: $\mathbf{K} = \begin{bmatrix} -3 \\ 69 \end{bmatrix}$.

Зададим значения матрицы модели оценки состояния наблюдателя (8):

$$\mathbf{G} = \begin{bmatrix} 0 & 1 \\ -3 & -4 \end{bmatrix}, \mathbf{L} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}.$$

На рисунке показаны переходные процессы в замкнутой системе при значении запаздывания $\tau = 1$ с при коэффициентах адаптации $\gamma = 5$ (рисунок, *a, c, e*) и $\gamma = 15$ (рисунок, *b, d, f*).

Представленные результаты моделирования демонстрируют эффективность и работоспособность предложенного решения. Данный подход обеспечивает асимптотическую сходимость выходного сигнала y к нулю при $t \rightarrow \infty$. Видно, что при увеличении коэффициента



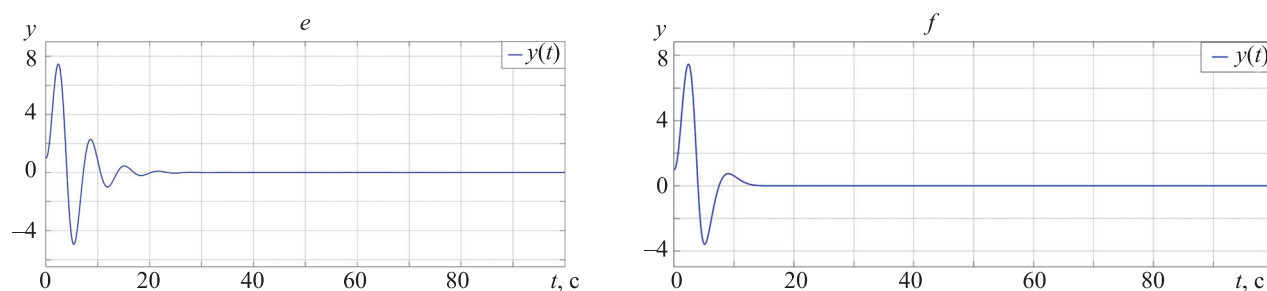


Рисунок. Графики переходных процессов при запаздывании $\tau = 1$ с для коэффициентов адаптации $\gamma = 5$ и $\gamma = 15$: сигнала управления $u(t)$ (a, b); оценок $\hat{\eta}_1, \hat{\eta}_2$ (c, d); выходного сигнала $y(t)$ (e, f)

Figura. Transient processes with delay $\tau = 1$ s for adaptation coefficients $\gamma = 5$ and $\gamma = 15$: control signal $u(t)$ (a, b); estimations $\hat{\eta}_1, \hat{\eta}_2$ (c, d); output signal $y(t)$ (e, f)

адаптации у удалось достичь лучшего быстродействия в системе.

Заключение

В работе рассмотрена задача компенсации по выходу внешних неизвестных возмущений для класса линейных систем с произвольным запаздыванием на

входе. Предложенный алгоритм обеспечивает ограниченность всех сигналов в системе и сходимость регулируемого выхода к нулю. Преимущество данного подхода заключается в том, что не требуется идентификация параметров возмущения. В дальнейшем рассмотренное решение может быть расширено на класс объектов с неизвестными параметрами и неизвестным запаздыванием.

Литература

1. Suulker C., Emirler M.T. Comparison of different time delay compensation methods for networked DC motor speed control // Proc. of the 6th International Conference on Electrical and Electronics Engineering (ICEEE). 2019. P. 225–229. <https://doi.org/10.1109/ICEEE2019.2019.00050>
2. Li K., Cai Z., Zhao J., Lou J., Wang J. Signal compensation control algorithm for quadrotor unmanned aerial vehicles // Proc. of the 36th Chinese Control Conference (CCC). 2017. P. 3266–3271. <https://doi.org/10.23919/ChiCC.2017.8027861>
3. Zheng W., Chen M. Tracking control of manipulator based on high-order disturbance observer // IEEE Access. 2018. V. 6. P. 26753–26764. <https://doi.org/10.1109/ACCESS.2018.2834978>
4. Richard J.P. Time-delay systems: an overview of some recent advances and open problems // Automatica. 2003. V. 39. N 10. P. 1667–1694. [https://doi.org/10.1016/S0005-1098\(03\)00167-5](https://doi.org/10.1016/S0005-1098(03)00167-5)
5. Пыркин А.А., Бобцов А.А., Никифоров В.О., Колубин С.А., Ведяков А.А., Борисов О.И., Громов В.С. Компенсация полигармонического возмущения, действующего на состояние и выход линейного объекта с запаздыванием в канале управления // Автоматика и телемеханика. 2015. № 12. С. 43–64.
6. Pyrkin A., Smyshlyaev A., Bekiaris-Liberis N., Krstic M. Rejection of sinusoidal disturbance of unknown frequency for linear system with input delay // Proc. of the 20th American Control Conference (ACC). 2010. P. 5688–5693. <https://doi.org/10.1109/ACC.2010.5531131>
7. Бобцов А.А., Пыркин А.А. Адаптивное и робастное управление с компенсацией неопределенностей: учебное пособие. СПб.: НИУ ИТМО, 2013. 135 с.
8. Pyrkin A.A., Bobtsov A.A., Nikiforov V., Vedyakov A., Kolyubin S., Borisov O. Output control approach for delayed linear systems with adaptive rejection of multiharmonic disturbance // IFAC Proceedings Volumes. 2014. V. 47. N 3. P. 12110–12115. <https://doi.org/10.3182/20140824-6-ZA-1003.01787>
9. Narendra K., Annaswamy A. Stable Adaptive Systems. New Jersey: Prentice Hall, 1989. 496 p.
10. Герасимов Д.Н., Парамонов А.В., Никифоров В.О. Алгоритм компенсации мультигармонических возмущений в линейных системах с произвольным запаздыванием: метод внутренней модели // Научно-технический вестник информационных технологий, механики и оптики. 2016. Т. 16. № 6. С. 1023–1030. <https://doi.org/10.17586/2226-1494-2016-16-6-1023-1030>

References

1. Suulker C., Emirler M.T. Comparison of different time delay compensation methods for networked DC motor speed control. *Proc. of the 6th International Conference on Electrical and Electronics Engineering (ICEEE)*, 2019, pp. 225–229. <https://doi.org/10.1109/ICEEE2019.2019.00050>
2. Li K., Cai Z., Zhao J., Lou J., Wang J. Signal compensation control algorithm for quadrotor unmanned aerial vehicles. *Proc. of the 36th Chinese Control Conference (CCC)*, 2017, pp. 3266–3271. <https://doi.org/10.23919/ChiCC.2017.8027861>
3. Zheng W., Chen M. Tracking control of manipulator based on high-order disturbance observer. *IEEE Access*, 2018, vol. 6, pp. 26753–26764. <https://doi.org/10.1109/ACCESS.2018.2834978>
4. Richard J.P. Time-delay systems: an overview of some recent advances and open problems. *Automatica*, 2003, vol. 39, no. 10, pp. 1667–1694. [https://doi.org/10.1016/S0005-1098\(03\)00167-5](https://doi.org/10.1016/S0005-1098(03)00167-5)
5. Pyrkin A.A., Bobtsov A.A., Nikiforov V.O., Kolyubin S.A., Vedyakov A.A., Borisov O.I., Gromov V.S. Compensation of polyharmonic disturbance of state and output of a linear plant with delay in the control channel. *Automation and Remote Control*, 2015, vol. 76, no. 12, pp. 2124–2142. <https://doi.org/10.1134/S0005117915120036>
6. Pyrkin A., Smyshlyaev A., Bekiaris-Liberis N., Krstic M. Rejection of sinusoidal disturbance of unknown frequency for linear system with input delay. *Proc. of the 20th American Control Conference (ACC)*, 2010, pp. 5688–5693. <https://doi.org/10.1109/ACC.2010.5531131>
7. Bobtsov A.A., Pyrkin A.A. *Adaptive and Robust Control with Uncertainties Compensation*. St. Petersburg, NRU ITMO, 2013, 135 p. (in Russian)
8. Pyrkin A.A., Bobtsov A.A., Nikiforov V., Vedyakov A., Kolyubin S., Borisov O. Output control approach for delayed linear systems with adaptive rejection of multiharmonic disturbance. *IFAC Proceedings Volumes*, 2014, vol. 47, no. 3, pp. 12110–12115. <https://doi.org/10.3182/20140824-6-ZA-1003.01787>
9. Narendra K., Annaswamy A. *Stable Adaptive Systems*. New Jersey, Prentice Hall, 1989, 496 p.
10. Gerasimov D.N., Paramonov A.V., Nikiforov V.O. Algorithm of multiharmonic disturbance compensation in linear systems with arbitrary delay: internal model approach. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2016,

11. Bobtsov A., Kremlev A. Adaptive compensation of biased sinusoidal disturbances with unknown frequency // *IFAC Proceedings Volumes*. 2005. V. 38. N 1. P. 131–136. <https://doi.org/10.3182/20050703-6-CZ-1902.00022>
12. Marino R., Tomei P. Output regulation for linear systems via adaptive internal model // *IEEE Transactions on Automatic Control*. 2003. V. 48. N 12. P. 2199–2202. <https://doi.org/10.1109/TAC.2003.820143>
13. Парамонов А.В. Адаптивная робастная компенсация возмущений в линейных системах с запаздыванием // *Научно-технический вестник информационных технологий, механики и оптики*. 2018. Т. 18. № 3. С. 384–391. <https://doi.org/10.17586/2226-1494-2018-18-3-384-391>
14. Никифоров В.О. Адаптивное и робастное управление с компенсацией возмущений. СПб.: Наука, 2003. 282 с.
15. Никифоров В.О. Наблюдатели внешних возмущений. 1. Объекты с известными параметрами // *Автоматика и телемеханика*. 2004. № 10. С. 13–23.
16. Krstic M., Kanellakopoulos I., Kokotovic P. *Nonlinear and Adaptive Control Design*. NY: John Wiley and Sons, Inc., 1995. 563 p.
- vol. 16, no. 6, pp. 1023–1030. (in Russian). <https://doi.org/10.17586/2226-1494-2016-16-6-1023-1030>
11. Bobtsov A., Kremlev A. Adaptive compensation of biased sinusoidal disturbances with unknown frequency. *IFAC Proceedings Volumes*, 2005, vol. 38, no. 1, pp. 131–136. <https://doi.org/10.3182/20050703-6-CZ-1902.00022>
12. Marino R., Tomei P. Output regulation for linear systems via adaptive internal model. *IEEE Transactions on Automatic Control*, 2003, vol. 48, no. 12, pp. 2199–2202. <https://doi.org/10.1109/TAC.2003.820143>
13. Paramonov A.V. Adaptive robust disturbance compensation in linear systems with delay. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2018, vol. 18, no. 3, pp. 384–391. (in Russian). <https://doi.org/10.17586/2226-1494-2018-18-3-384-391>
14. Nikiforov V.O. *Adaptive and Robust Control with Compensation of the Disturbances*. St. Petersburg, Nauka Publ., 2003, 282 p. (in Russian)
15. Nikiforov V.O. Observers of external deterministic disturbances. I. Objects with known parameters. *Automation and Remote Control*, 2004, vol. 65, no. 10, pp. 1531–1541. <https://doi.org/10.1023/B:AURC.0000044264.74470.48>
16. Krstic M., Kanellakopoulos I., Kokotovic P. *Nonlinear and Adaptive Control Design*. NY, John Wiley and Sons, Inc., 1995, 563 p.

Авторы

Буй Ван Хуан — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-6563-1909>, buinguyenkhanh201095@gmail.com

Маргун Алексей Анатольевич — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; научный сотрудник, Институт проблем машиностроения РАН, Санкт-Петербург, 199178, Российская Федерация, <https://orcid.org/0000-0002-5333-0594>, alexeimargun@gmail.com

Authors

Van Huan Bui — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-6563-1909>, buinguyenkhanh201095@gmail.com

Alexey A. Margun — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation; Scientific Researcher, Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, <https://orcid.org/0000-0002-5333-0594>, alexeimargun@gmail.com

Статья поступила в редакцию 14.06.2022
Одобрена после рецензирования 10.10.2022
Принята к печати 27.11.2022

Received 14.06.2022
Approved after reviewing 10.10.2022
Accepted 27.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

НОВЫЕ МАТЕРИАЛЫ И НАНОТЕХНОЛОГИИ MATERIAL SCIENCE AND NANOTECHNOLOGIES

doi: 10.17586/2226-1494-2022-22-6-1078-1084

Luminescence technique for studying the growth of AgInS₂ quantum dots

Ahmad K. Ahmad¹✉, Ammar H. Mohammed², Alexander A. Skaptsov³

¹ Al-Nahrain University, College of Engineering, Baghdad, 10072, Iraq

² Al-Nahrain University, College of Science, Baghdad, 10072, Iraq

³ Saratov State University, Saratov, 410012, Russian Federation

¹ ahmad.ahmad@nahrainuniv.edu.iq✉, <https://orcid.org/0000-0003-3522-4701>

² ammarhussen659@gmail.com, <https://orcid.org/0000-0003-3646-4790>

³ skaptsov@yandex.ru, <https://orcid.org/0000-0001-9336-6885>

Abstract

Although nanoparticle production techniques are well-known, getting nanoparticles with specific characteristics that enable their application as biosensors is an entirely other problem. Many issues occur as a result of employing the method for producing repeatable and time-stable nanostructures. We created AgInS₂ nanoparticles as colloidal quantum dots in a variety of methods to test the efficiency of the synthesis process on the optical characteristics of the nanoparticles, as well as their size, composition, absorption, and luminescence spectra. The capillary electrophoresis (CE) approach for AgInS₂ production was employed, with modifications in solvent and temperature, to get nanocrystal (NC) particles. The researchers discovered that Ag accumulation in the InS lattice promotes deformation which leads to structural defects. Consequently, the direction of a nanoparticle light band may now be changed. The features of mixed AgInS₂ nanoparticles have been examined with respect to different fabrication procedures, surface stability, and metal impurity incorporation. One band dominates in the luminescence spectra of Ag_xIn_{1-x}S₂ nanoparticles. The relationship between the stoichiometric ratio, luminescence amplitude, line width, and the maximum wavelength is investigated. The average size of the received nanocrystals was determined using dynamic light scattering studies. The computed nanoparticle diameter range has an average particle size of 3–3.5 nm.

Keywords

AgInS₂, quantum dot, luminescence, synthesis of nanocrystalline particles, biosensors

For citation: Ahmad A.K., Mohammed A.H., Skaptsov A.A. Luminescence technique for studying the growth of AgInS₂ quantum dots. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1078–1084. doi: 10.17586/2226-1494-2022-22-6-1078-1084

УДК 535.37:544.77.051

Люминесцентный метод исследования роста квантовых точек AgInS₂

Ахмад Карнал Ахмад¹✉, Аммар Хуссейн Мохаммед², Александр Александрович Скапцов³

¹ Аль-Нарейн Университи, Инженерный колледж, Багдад, 10072, Ирак

² Аль-Нарейн Университи, Багдад, 10072, Ирак

³ Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, 410012, Российская Федерация

¹ ahmad.ahmad@nahrainuniv.edu.iq✉, <https://orcid.org/0000-0003-3522-4701>

² ammarhussen659@gmail.com, <https://orcid.org/0000-0003-3646-4790>

³ skaptsov@yandex.ru, <https://orcid.org/0000-0001-9336-6885>

Аннотация

Методы получения наночастиц хорошо известны, однако стабильно повторяемое получение наночастиц с одинаковыми специфическими характеристиками, позволяющими их использовать в качестве биосенсоров, в настоящее время представляет собой проблему. Проблема многих методов заключается в получении результатов с хорошей воспроизводимостью оптических свойств и их стабильностью во времени. В работе получены

© Ahmad A.K., Mohammed A.H., Skaptsov A.A., 2022

наночастицы AgInS_2 в виде коллоидных квантовых точек различными методами. Выполнен анализ процесса синтеза, который влияет на оптические характеристики наночастиц, их размер, состав, спектры поглощения и люминесценции. Для получения нанокристаллических частиц при изготовлении AgInS_2 применен капиллярный электрофорез с модификациями растворителя и температуры. Обнаружено, что накопление Ag в решетке InS способствует деформации, которая приводит к структурным дефектам и, следовательно, к смещению спектральной полосы люминесценции. Исследованы характеристики смешанных наночастиц AgInS_2 при различных вариациях технологии изготовления, стабильности поверхности и включения металлических примесей. Показано, что в спектрах люминесценции наночастиц $\text{Ag}_x\text{In}_{1-x}\text{S}_2$ доминирует одна полоса. Исследована зависимость между стехиометрическим соотношением, амплитудой люминесценции, шириной линии и максимальной длиной волны. Средний размер полученных нанокристаллов определен с использованием метода динамического светорассеяния. Расчетный размер наночастиц составил в среднем 3–3,5 нм.

Ключевые слова

AgInS_2 , квантовые точки, люминесценция, синтез наноструктур, биосенсоры

Ссылка для цитирования: Ахмад А.К., Мохаммед А.Х., Скапцов А.А. Люминесцентный метод исследования роста квантовых точек AgInS_2 // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1078–1084 (на англ. яз.). doi: 10.17586/2226-1494-2022-22-6-1078-1084

Introduction

In recent years, AgInS_2 nanocrystals have garnered a great deal of interest among semiconductors of groups I, III, and VI due to their various benefits, such as their customizable optical characteristics, high extinction coefficients, high quantum yields, diverse emission spectra, and photochemical durability. Several techniques, mostly based on thermolytic one-of-a-kind metallic-sulfur complexes or direct metallic response cations and sulphur with suitable covering reagents [1, 2], have been published for the ternary synthesis of AgInS_2 with varied sizes. Metallic cations are typically coupled with amines or dodecanethiol to form complex metallic precursors, such as Colloidal Quantum Dots (CODs), which offer particles with controlled outgrowth [3]. Sulfur can be derived from elemental sulphur, thiols, dithiocarbamates, or carbon disulfide, all of which are soluble in organic solvents [4, 5]. Metallic thiolates or dithiocarbamates have been used as a source of both metal cations and sulfur [6].

Recently, it has been observed that studies and research focus on the use of CQDs. They are 2–10 nm long semiconductor nanocrystals that blend the physical and chemical properties of molecules with the optoelectronic properties of semiconductors. Quantum dots (QDs) have high fluorescence brightness and are more resistant to picture bleaching than fluorophores making them excellent candidates for long-term study. The author recommended against conjugating CQDs derived from non-toxic $\text{AgInS}_2/\text{ZnS}$ compounds with nanoparticles of other types.

Using novel synthesis techniques and materials, it is possible to increase the quantum yield of QDs significantly. CQDs play a crucial role in the current world, particularly for biomedical applications where they may be used to identify or pinpoint various illnesses and subsequently to provide medicine locally utilizing the same circuit class therapy but with a vaccine medication. In this context, the development of CQD-based labeling for less hazardous semiconductors of groups I-III and VI, such as CuInS_2 and AgInS_2 , appears promising. This form of CQD is distinguished by the ability to tune the emission band from the visible to the near-infrared region and by the comparatively long lifetime of an excited state. This study aims to investigate the feasibility of using AgInS_2 -based

QDs for fluorescence imaging of biological objects and to compare their performance with that of fluorescent dyes immobilized on the surface of nanoparticles.

Indium silver disulfide is a semiconductor that has shown promise in solar photovoltaics [7, 8]. This semiconductor has potential applications in photocells that concentrate solar energy, in cascade structures with a gradient in width, and in noisy zones. The combination of indium disulfide with ribs and, for example, indium diselenide copper with a common base of CdS can result in a tenfold improvement in efficiency [9, 10]. By matching the crystal lattices of the semiconductors mentioned above, AgInS_2 has a high optical absorption (greater than 10^4 cm^{-1}), strong radiation resistance, and low dissipation. Space charge parameters and effective carrier separation may be modified in many multilayer structures, leading to an active investigation of such combinations. AgInS_2 bulk crystals and films have received little attention. The goal of this effort is to semi-study AgInS_2 films and their electrical, physical, and optical characteristics.

These materials, like AgInS_2 , are semiconductors belonging to the III group; in the visible and near-ultraviolet spectrum they exhibit enormous gaps range and significant deactivation rates. Because of their luminous qualities, such combinations pique much interest as an alternative for a more dangerous cadmium-containing quantum. Currently, many experiments utilise luminous semiconductors (such as AgInS_2) that may be used as Nano thermometers by relying on ambient temperature. As a result, when samples containing AgInS_2 nanoparticles (4 nm in size on average) are heated, a drop in fluorescence intensity and spectrum shift is detected concurrently, both under single and two-photon excitation. The connection between fluctuations in the peak wavelength of the radiation is linear with temperature for AgInS_2 . Currently, luminous nanoparticles are utilized to assess cell temperature in various applications [11, 12]. Laser thermolysis is a significant and promising medical sector, particularly in various operations and treatments.

Photothermolysis employs a variety of photosensitizers, the most effective of which being plasmon-resonant nanoparticles (Plasmon Resonance Nanoparticles) [13, 14]. Due to limited laser radiation absorption when warming the nanoparticles delivered into the tumor

site, laser photothermolysis does little harm to hygiene cells of biological tissue around the tumor. In reality, the fundamental limitation of its efficacy is the problem of unwanted heating of adjacent tissue which also involves normal cells at risk and induces the initiation of coagulation. The heating process must thus be monitored to accomplish biological tissue coagulation in the complete tumor volume and to halt irradiation after disease eradication. Inadequate heating may promote tumor development, while excessive heating may cause healthy biological tissue to clot. As a result, the advancement of the laser photothermolysis process is hampered by a lack of trustworthy or reliable methods for determining optimal exposure settings. Direct and fast detection of thermal fields in organic tissue near Plasmon resonance nanoparticles is feasible for tissue heating control. Several non-invasive thermometry procedures include radiation ultrasound and magnetic resonance imaging.

Despite encouraging results, these approaches are often constrained by their lack of temporal resolution. During photothermolysis, when heating, the biological tissue is specifically marked by non-contact temperature scaling of the surface with a thermal imager or measuring the inside temperature at specific spots with a thermocouple. However, denying the ultimate domination of the heat process interaction and propagation of laser light in tumor materials causes uncertainty in calculating the interior temperature from the obtained surface temperature division with the aforementioned data. As a result, controlled photothermolysis becomes much more challenging. As a result, identifying in real time the spatial distribution of temperature inside biological entities remains crucial. In this scenario, the approach is to create a nanosensor that can forecast the incidence of tissue coagulation that is irreversible. We investigated how temperature affects the luminescence spectra of nanoparticles in aqueous and biological contexts. As the temperature rises, the fluorescence intensity of AgInS₂ nanoparticles in water decreases, and a temperature change of 25 °C cause a line shift by 5 nm to the luminescence peak. Furthermore, it was evident that the location of the luminescence peak of AgInS₂ nanoparticles is influenced by the medium. By heating the AgInS₂ nanoparticles in biological specimens, the peak luminescence shifts linearly to longer wavelengths, where the maximum shear rate is 0.459 ± 0.035 nm.

We created AgInS₂ nanoparticles in various methods to assess the influence of the synthesis process on the condition and optical features of nanoparticles, to explore their luminescence spectra, absorption, size, and structure. Colloidal QDs exhibit critical optical properties that are proportional to particle size. Similar to the box model in quantum mechanics, as the border dimensions drop, the energy of each state rises. Subsequent to photon absorption, the valence and conduction bands of the QDs are quantized as a result of the excited state which is an electron-hole pair with dimensions lower than the exciton Bohr radius. This quantization may be demonstrated in QD absorption spectra where discrete low-energy transitions are seen in a clean, monodisperse sample. A particle size function provides the Band edge energy of a QD (E) and is given by the following equation [15]:

$$E = E_g + \left(\frac{\hbar^2 \alpha_{ne,le}^2}{2m_e r^2} \right) + \left(\frac{\hbar^2 \alpha_{nh,lh}^2}{2m_h r^2} \right) - \frac{1.8e^2}{2\epsilon r}.$$

In the equation above, E_g is the bulk semiconductor band gap where the name originally appeared. The second and third terms, with $\alpha_{ne,le}^2$ and $\alpha_{nh,lh}^2$, are solutions (or “zeros”) to the Bessel function for quantum numbers n and l , respectively, giving information about the electron/hole kinetic energy.

The Bessel function method solves the Schrodinger equation under initial spherical conditions and computes the kinetic energy of the electron/hole. Size dependency is related to kinetic energy via QD radius (r) in Bessel function solutions. m_e and m_h incorporate effective electron/hole masses. The approximation of the effective mass compensates for the periodic potential of the semiconductor lattice by changing a particle volume and considering it as a smooth particle. The final term (fourth term) is a first-order correction for the coulomb attraction between the electron and hole where e is the charge of an electron and ϵ is the dielectric constant of the material.

Two nanoparticle solution synthesis processes have been agreed upon the high and low temperature methods. The above growth rate equation can explain why the effects of the temperature differ from those of the calculation based on the total gas equation with monomer concentrations handled with vapor pressure [16].

$$C_r = C_\infty \exp\left(\frac{2\sigma}{nNK_B T}\right) \approx C_\infty \left(1 + \frac{2\sigma}{rNK_B T}\right),$$

$$C_x = C_\infty \exp\left(\frac{2\sigma}{r_c NK_B T}\right) \approx C_\infty \left(1 + \frac{2\sigma}{r_c NK_B T}\right),$$

where C_∞ is vapor pressure; C_r is the concentration of monomers at the nanocrystal surface; and C_x is the concentration of monomers far away from the crystal surface. The dependent growth rate temperature for crystal radius r is given by the following equation [16]

$$\frac{dr}{dt} = \frac{2\sigma\beta C_\infty}{N^2 K_B T} \left(\frac{1}{r_c} - \frac{1}{r} \right), \quad (1)$$

where r_c is the critical radius at which the phase transition occurs; N is the number of atoms per unit volume; σ is the surface tension; β is the diffusion coefficient; K_B is the Boltzmann constant, and T is the temperature.

Based on the crystal radius r in Eq. (1), the word in parentheses defines whether a crystal has a positively or negatively increasing rate of size. When $r = r_c$, the crystalline phases are in equilibrium with zero crystal growth rate. When $r > r_c$, the crystal will grow. When $r < r_c$, the crystal will shrink. Ostwald ripening is a technique that aims to widen the size range of crystals while lowering the concentration of the resultant nanoparticles. Smaller crystals contract while bigger crystals develop. In Fig. 1, the form of Eq. (1) is shown. It is significant to remember that when $r = 2r_c$, the maximum growth rate may be calculated. The smaller crystals develop considerably quicker than the bigger crystals, and the average size is focused, resulting in a very narrow size distribution if the

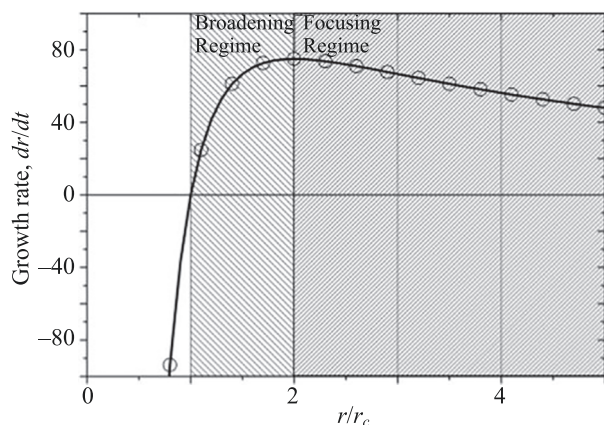


Fig. 1. An illustration form of the Eq. (1), illustrates the modes of size and dimension broadening specified by the average crystal radius $r_c < r < 2r_c$ and $r > 2r_c$, respectively [16]

entire size range is greater than this number. In reality, the typical technique repeatedly uses a monomer solution with a high enough concentration [17, 18].

Materials and Synthesis procedures

We developed a technique for producing isolated AgInS₂ nanoparticles with great size uniformity in an orthoxyle solution at high temperatures [19]. To boost luminous yield, we devised a method for synthesizing mixed AgInS₂ nanoparticles [20] which had a higher luminescence production than AgS or InS particles. The lattice period is also affected by changes in the Silver-to-Indium ratio. At a particular ratio, the structure varies from cubic for InS to hexagonal for AgS. Using diffuse transmission and reflection spectra, the kinetics of particle production and their characteristics were investigated. During synthesis durations up to 90 min, the solution absorption edge shifts to the short-wavelength area for diffuse transmission spectra. Following that, the edge gradually moves toward the long-wavelength region. As the contribution of the scattered signal is minimal in the region of significant absorption, the displacements

are generated just by variations in the absorption of the solution components. One can explain the shift to the short-wavelength area by a sulfur drop which causes a large rise in the absorption coefficient in the 300–320 nm wavelength range. As the synthesis period rises, the size of the nanoparticles induces a shift to longer wavelengths. More complex behaviour is exhibited by reflection spectra. One might cause a simultaneous alteration in particle size, concentration, and refractive index of the medium. Techniques for manufacturing AgInS₂ nanoparticles at low and high temperatures were devised. The impact of the circumstances during synthesis on the formation and properties of nanoparticles was investigated. Depending on the concentration and according to an evaluation of the acquired absorption spectra of AgInS₂ nanoparticles, the location of the nanoparticles absorption edge shifts during the duration of the synthesis period for high-temperature synthesis. With an increase the ultrasonic action duration on the sample, the absorption edge in diffuse reflection spectra shifts by a short wavelength.

For AgInS₂ synthesis, with variations in solvent and temperature, portions of cold toluene were removed at various intervals and inputted for various Nano-Crystal (NC) sizes. The reaction mixture has a total volume of 30 ml. Base metal salt aqueous solution is: 0.71 ml of 0.1 M InCl₃; Thioglycolic Acid aqueous solution is: 0.6 ml of 1 M; concentrated ammonia is 0.3 ml; ammonium sulphide aqueous solution is: 0.174 ml of 2.9 M; 10 ml of 0.005 M silver nitrate aqueous solution was dripped in at a rate of 3 ml/hour. Following the NC synthesis, the ethanol cycle was centrifuged three times with toluene and then dissolved in toluene for storage. However, importantly, not the case that QDs may induce acute toxic effects, rather than that the aggregation of QD ions in key organs, as seen in vivo, might create significant chronic toxic consequences [21]. As seen in Fig. 2, *a*, *b*, we devised a revolutionary control technique by flashing a 405 nm laser on the solution during the synthesis and luminescence spectra were analyzed every 10s. Then the spectra were analyzed. A QE Pro spectrometer (Ocean Optics, USA), a 405 nm laser (200 mW), optical fibers, and filters are used in the experiment. Every ten seconds, the luminescence spectrum

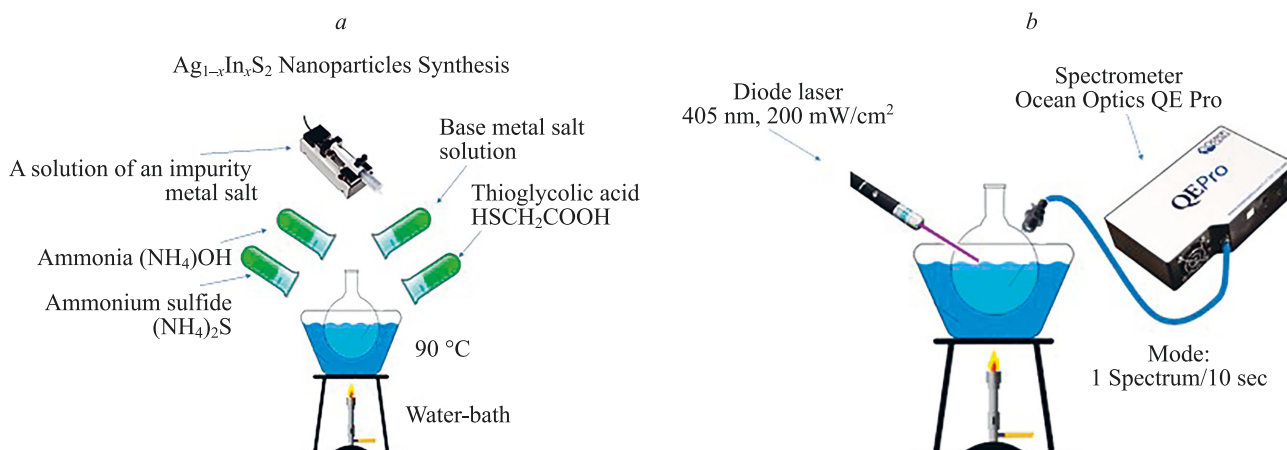


Fig. 2. Synthesis Control Methodology; HSCH₂COOH
a — aqueous nanoparticle synthesis, *b* — nanoparticle synthesis control method

of the reaction solution was recorded using a one-second laser stimulation.

Results and Discussions

The effects of the stoichiometric coefficient on the optical properties Ag_xIn_{1-x}S₂ nanoparticles, amplitude, wavelength of the peak of the luminescence band, and width of AgInS₂ nanoparticles are depicted in Fig. 3. The maximum luminescence wavelength moves to the IR region when the stoichiometric coefficient (x) changes. It is also worth noting that the amplitude begins to fall at a certain value.

The nanoparticles are identified by the altered InS structure induced by the presence of AgS impurities, according to X-ray analysis. Fig. 4 depicts the X-ray diffraction (XRD) spectra of doped QDs with various ratios of indium to silver (1:0.05, 1:0.1, 1:0.2, and 1:0.3). In general, relative intensity increases with an increase in silver ratio; nevertheless, there is no discernible shift in 2θ values. This shows that silver contributes to the diffraction of InAg/ZnS QDs doped with silver. The absence of a specific silver peak demonstrates the excellent incorporation of Ag₂S into the InAg matrix.

The X-ray fluorescence (XRF) patterns of Ag_xIn_{1-x}S₂ with different In concentrations (x : 0.03–0.18) and pure AgS are shown in Fig. 5. The structure is cubic in nature. Ag concentration is being increased.

According to our findings, one can infer the temperature sensitivity of luminescence of AgInS₂ nanoparticles in terms of both the spectrum position and intensity of the luminescence maximum. The variation is caused by the energy states distortion of light and surface defects. In Ag_xIn_{1-x}S₂ samples, a transition in the structural phase from cubic to hexagonal and a temperature alteration change to a lower region occurs with an increase in Ag concentration.

Dynamic light scattering experiments were used to measure the average size of the received nanocrystals. The measurements were taken using a laser analyzer with a particle size of SZ100 (Horiba Jobin Yvon, Kyoto, Japan) and a measuring range for nanoparticle diameters ranging

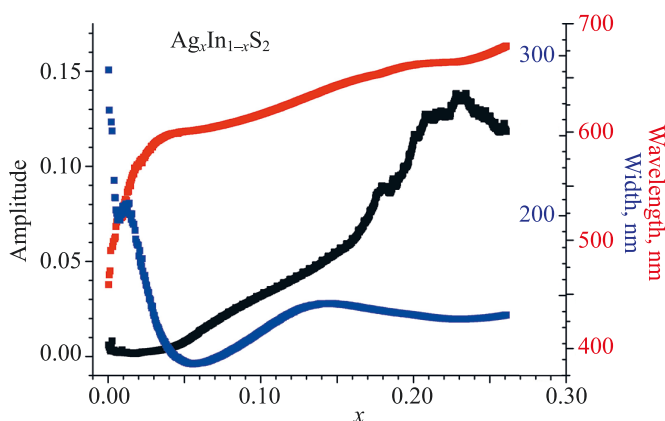


Fig. 3. Controlled synthesis of Ag_xIn_{1-x}S₂ nanoparticles. Illustrates the relation between the wavelength, the amplitude, and the width of AgInS₂ nanoparticles with the stoichiometric coefficient x

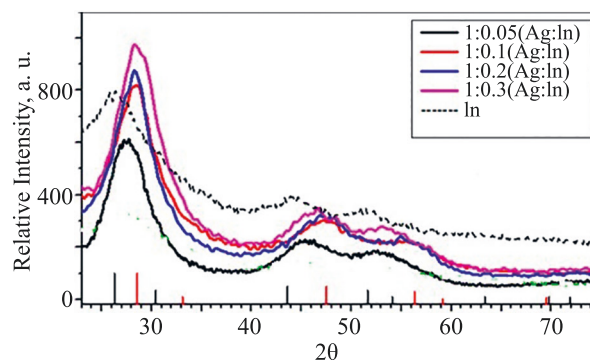


Fig. 4. XRD spectra of silver doped quantum dots with varying Ag:In ratios

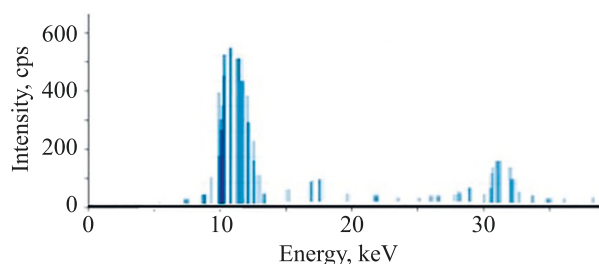


Fig. 5. Shown the AgInS₂ nanoparticles XRF analysis

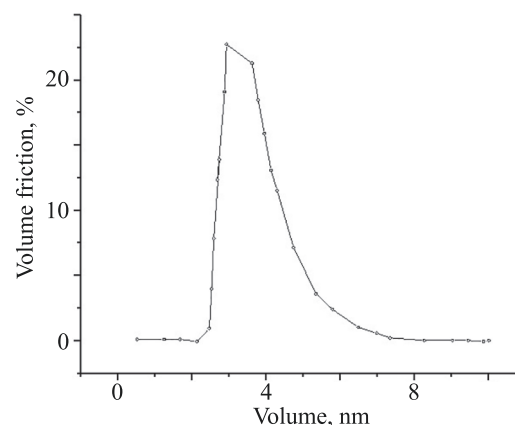


Fig. 6. The average particle size

from 0.3 nm to 8 nm. Forks have an average particle size of 3–3.5 nm (Fig. 6).

Conclusion

To increase luminescence yield, we devised a method for producing mixed AgInS₂ nanoparticles which have a higher luminescence yield than AgS or InS particles. When a certain ratio is reached, the arrangement of InS shifts to hexagonal for AgS. At lower substrate temperatures, the composition generates a hexagonal phase. As a result, it is now possible to change the orientation of a nanoparticle light band. The properties of mixed AgInS₂ nanoparticles for various production procedures, surface stability, and metal impurity introduction were examined. The link between the line width, maximum wavelength, luminescence amplitude, and the stoichiometric ratio is depicted in Fig. 3. It was

proven that adjusting the stoichiometric ratio of In and Ag in $\text{Ag}_x\text{In}_{1-x}\text{S}_2$ nanoparticles does not allow for the maximum luminescence wavelength in biological tissue transparency. According to the data, from this it can be drawn that the synthesis of $\text{Ag}_x\text{In}_{1-x}\text{S}_2$ nanoparticles (at x 0.25) does not permit the formation of nanoparticles

with maximal luminescence band wavelengths greater than 800 nm. The amplitude of the luminescence bands decreases significantly as x exceeds 0.25. This change is caused by surface defects and distortion of the light energy levels. Nanoparticle agglomeration has no impact on the reported dependency.

References

1. Du W.M., Qian X.F., Yin J., Gong Q. Shape- and phase-controlled synthesis of monodisperse, single-crystalline ternary chalcogenide colloids through a convenient solution synthesis strategy. *Chemistry — A European Journal*, 2007, vol. 13, no. 31, pp. 8840–8846. <https://doi.org/10.1002/chem.200700618>
2. Ogawa T., Kuzuya T., Hamanaka Y., Sumiyama K.J. Synthesis of Ag–In binary sulfide nanoparticles—structural tuning and their photoluminescence properties. *Journal of Materials Chemistry*, 2010, vol. 20, no. 11, pp. 2226–2231. <https://doi.org/10.1039/B920732E>
3. Torimoto T., Adachi T., Okazaki K., Sakurao M., Shibayama T., Ohtani B., Kudo A., Kuwabata S. Facile synthesis of ZnS-AgInS_2 solid solution nanoparticles for a color-adjustable luminophore. *Journal of the American Chemical Society*, 2007, vol. 129, no. 41, pp. 12388–12389. <https://doi.org/10.1021/ja0750470>
4. Wang D.S., Zheng W., Hao C.H., Peng Q., Li Y. General synthesis of I–III–VI₂ ternary semiconductor nanocrystals. *Chemical Communications*, 2008, vol. 22, pp. 2556–2558. <https://doi.org/10.1039/B800726H>
5. Xie R.G., Rutherford M., Peng X.G. Formation of high-quality I–III–VI semiconductor nanocrystals by tuning relative reactivity of cationic precursors. *Journal of the American Chemical Society*, 2009, vol. 131, no. 15, pp. 5691–5697. <https://doi.org/10.1021/ja9005767>
6. Feng Z.Y., Dai P.C., Ma X.C., Zhan J.H., Lin Z. Monodispersed cation-disordered cubic AgInS_2 nanocrystals with enhanced fluorescence. *Applied Physics Letters*, 2010, vol. 96, no. 1, pp. 013104. <https://doi.org/10.1063/1.3280372>
7. Chopra K.L., Das S.R. *Thin Film Solar Cells*. New York; London, 1983.
8. Alkhasov A.B. *Renewable Energy*. Moscow, Fizmatlit Publ., 2010. 255 p. (in Russian)
9. Ariezo M., Loferski J.J. Investigation of potentially high efficiency photovoltaic cells consisting of two heterojunctions on a common wide band gap semiconductor base. *Proc. of the 13th IEEE Photovoltaic Specialists Conference*, 1978, pp. 898–903.
10. Abdullaev M.A., Alkhasov A.B., Magomedova D.Kh. Fabrication and properties of $\text{CuInSe}_2/\text{AgInSe}_2/\text{CdS}$ double heterojunction cascade solar cells. *Inorganic Materials*, 2014, vol. 50, no. 3, pp. 228–232. <https://doi.org/10.1134/S0020168514030017>
11. Maestro L.M., Rodríguez E.M., Rodríguez F.S., Iglesias-de la Cruz M.C., Juarranz A., Naccache R., Vetrone F., Jaque D., Capobianco J.A., Solé J.G. CdSe quantum dots for two-photon fluorescence thermal imaging. *Nano Letters*, 2010, vol. 10, no. 12, pp. 5109–5115. <https://doi.org/10.1021/nl1036098>
12. Yang J.M., Yang H., Lin L. Quantum dot nano thermometers reveal heterogeneous local thermogenesis in living cells. *ACS Nano*, 2011, vol. 5, no. 6, pp. 5067–5071. <https://doi.org/10.1021/nn201142f>
13. Skaptsov A.A., Ustakov S.O., Mohammed A.H., Savenko O.A., Novikova A.S., Kozlova E.A., Kochubey V.I. Fabrication and characterization of biological tissue phantoms with embedded nanoparticles. *Journal of Physics: Conference Series*, 2017, vol. 917, no. 4, pp. 042003. <https://doi.org/10.1088/1742-6596/917/4/042003>
14. Ustakov S.O., Kozlova E.A., Savenko O.A., Mohammed A.H., Kochubey V.I., Skaptsov A.A. Influence of excitation power density on temperature dependencies of $\text{NAYF}_4:\text{Yb}$, Er nanoparticles luminescence spectra. *Proceedings of SPIE*, 2017, vol. 10336, pp. 1033614. <https://doi.org/10.1117/12.2269297>
15. Jensen R.A. *Optical studies of colloidal quantum dots: optical trapping with plasmonic nano apertures and thermal recovery from photoinduced dimming*. PhD thesis. Department of Chemistry, Massachusetts Institute of Technology, 2015.
16. Booth M. *Synthesis and characterisation of CuInS_2 quantum dots*. PhD thesis. The University of Leeds School of Physics & Astronomy, 2014.

Литература

1. Du W.M., Qian X.F., Yin J., Gong Q. Shape- and phase-controlled synthesis of monodisperse, single-crystalline ternary chalcogenide colloids through a convenient solution synthesis strategy // *Chemistry — A European Journal*. 2007. V. 13. N 31. P. 8840–8846. <https://doi.org/10.1002/chem.200700618>
2. Ogawa T., Kuzuya T., Hamanaka Y., Sumiyama K.J. Synthesis of Ag–In binary sulfide nanoparticles—structural tuning and their photoluminescence properties // *Journal of Materials Chemistry*. 2010. V. 20. N 11. P. 2226–2231. <https://doi.org/10.1039/B920732E>
3. Torimoto T., Adachi T., Okazaki K., Sakurao M., Shibayama T., Ohtani B., Kudo A., Kuwabata S. Facile synthesis of ZnS-AgInS_2 solid solution nanoparticles for a color-adjustable luminophore // *Journal of the American Chemical Society*. 2007. V. 129. N 41. P. 12388–12389. <https://doi.org/10.1021/ja0750470>
4. Wang D.S., Zheng W., Hao C.H., Peng Q., Li Y. General synthesis of I–III–VI₂ ternary semiconductor nanocrystals // *Chemical Communications*. 2008. V. 22. P. 2556–2558. <https://doi.org/10.1039/B800726H>
5. Xie R.G., Rutherford M., Peng X.G. Formation of high-quality I–III–VI semiconductor nanocrystals by tuning relative reactivity of cationic precursors // *Journal of the American Chemical Society*. 2009. V. 131. N 15. P. 5691–5697. <https://doi.org/10.1021/ja9005767>
6. Feng Z.Y., Dai P.C., Ma X.C., Zhan J.H., Lin Z. Monodispersed cation-disordered cubic AgInS_2 nanocrystals with enhanced fluorescence // *Applied Physics Letters*. 2010. V. 96. N 1. P. 013104. <https://doi.org/10.1063/1.3280372>
7. Чопра К.Л., Дас С.Р. *Тонкопленочные солнечные элементы*. М.: Мир, 1986. 440 с.
8. Алхасов А.Б. *Возобновляемая энергетика*. М.: Физматлит, 2010. 255 с
9. Ariezo M., Loferski J.J. Investigation of potentially high efficiency photovoltaic cells consisting of two heterojunctions on a common wide band gap semiconductor base // *Proc. of the 13th IEEE Photovoltaic Specialists Conference*. 1978. P. 898–903.
10. Abdullaev M.A., Alkhasov A.B., Magomedova D.Kh. Fabrication and properties of $\text{CuInSe}_2/\text{AgInSe}_2/\text{CdS}$ double heterojunction cascade solar cells // *Inorganic Materials*. 2014. V. 50. N 3. P. 228–232. <https://doi.org/10.1134/S0020168514030017>
11. Maestro L.M., Rodríguez E.M., Rodríguez F.S., Iglesias-de la Cruz M.C., Juarranz A., Naccache R., Vetrone F., Jaque D., Capobianco J.A., Solé J.G. CdSe quantum dots for two-photon fluorescence thermal imaging // *Nano Letters*. 2010. V. 10. N 12. P. 5109–5115. <https://doi.org/10.1021/nl1036098>
12. Yang J.M., Yang H., Lin L. Quantum dot nano thermometers reveal heterogeneous local thermogenesis in living cells // *ACS Nano*. 2011. V. 5. N 6. P. 5067–5071. <https://doi.org/10.1021/nn201142f>
13. Skaptsov A.A., Ustakov S.O., Mohammed A.H., Savenko O.A., Novikova A.S., Kozlova E.A., Kochubey V.I. Fabrication and characterization of biological tissue phantoms with embedded nanoparticles // *Journal of Physics: Conference Series*. 2017. V. 917. N 4. P. 042003. <https://doi.org/10.1088/1742-6596/917/4/042003>
14. Ustakov S.O., Kozlova E.A., Savenko O.A., Mohammed A.H., Kochubey V.I., Skaptsov A.A. Influence of excitation power density on temperature dependencies of $\text{NAYF}_4:\text{Yb}$, Er nanoparticles luminescence spectra // *Proceedings of SPIE*. 2017. V. 10336. P. 1033614. <https://doi.org/10.1117/12.2269297>
15. Jensen R.A. *Optical studies of colloidal quantum dots: optical trapping with plasmonic nano apertures and thermal recovery from photoinduced dimming*: PhD thesis. Department of Chemistry, Massachusetts Institute of Technology, 2015.
16. Booth M. *Synthesis and characterisation of CuInS_2 quantum dots*: PhD thesis. The University of Leeds School of Physics & Astronomy, 2014.

17. Peng X., Wickham J., Alivisatos A.P. Kinetics of II-VI and III-V colloidal semiconductor nanocrystal growth: "focusing" of size distributions. *Journal of the American Chemical Society*, 1998, vol. 120, no. 21, pp. 5343–5344. <https://doi.org/10.1021/ja9805425>
18. Mohammed A.H., Skaptsov A.A., Ahmad A.K. Luminescence method to characterize the synthesis of ZnCdS quantum dots in real time. *Materials Today: Proceedings*, 2021, vol. 42, pp. 2803–2807. <https://doi.org/10.1016/j.matpr.2020.12.725>
19. Kochubey V.I., Konyukhova Ju.G., Volkova E.K. Effect of polyacrylic acid shell on luminescence and phosphorescence of ZnCdS nanoparticles. *Book of Abstracts of the 3rd International Symposium "Molecular photonics"*, St. Petersburg, Russia, VVM Publishing Ltd, 2012, pp. 137.
20. Volkova E.K., Kochubey V.I. Synthesis and luminescent characteristics of CdS nanoparticles. *International Symposium «Nanophotonics-2011»: Collection of the reports abstracts and programs*. Ukraine, Crimea, 2011, pp. 1–2. (in Russian)
21. Volkova E.K., Yanina I.Yu., Sagaydachnaya E., Konyukhova J.G., Kochubey V.I., Tuchin V.V. Effect of luminescence transport through adipose tissue on measurement of tissue temperature by using ZnCdS nanothermometers. *Proceedings of SPIE*, 2018, vol. 10493, pp. 104931K. <https://doi.org/10.1117/12.2295620>
17. Peng X., Wickham J., Alivisatos A.P. Kinetics of II-VI and III-V colloidal semiconductor nanocrystal growth: "focusing" of size distributions // *Journal of the American Chemical Society*. 1998. V. 120. N 21. P. 5343–5344. <https://doi.org/10.1021/ja9805425>
18. Mohammed A.H., Skaptsov A.A., Ahmad A.K. Luminescence method to characterize the synthesis of ZnCdS quantum dots in real time // *Materials Today: Proceedings*. 2021. V. 42. P. 2803–2807. <https://doi.org/10.1016/j.matpr.2020.12.725>
19. Kochubey V.I., Konyukhova Ju.G., Volkova E.K. Effect of polyacrylic acid shell on luminescence and phosphorescence of ZnCdS nanoparticles // *Book of Abstracts of the 3rd International Symposium "Molecular photonics"*. St. Petersburg, Russia: VVM Publishing Ltd, 2012. P. 137.
20. Волкова Е.К., Кочубей В.И. Синтез и люминесцентные характеристики наночастиц CdS // *Международный симпозиум «Нанофотоника-2011»: сборник тезисов докладов и программ*. Украина, Крым. 2011. С. 1–2.
21. Volkova E.K., Yanina I.Yu., Sagaydachnaya E., Konyukhova J.G., Kochubey V.I., Tuchin V.V. Effect of luminescence transport through adipose tissue on measurement of tissue temperature by using ZnCdS nanothermometers // *Proceedings of SPIE*. 2018. V. 10493. P. 104931K. <https://doi.org/10.1117/12.2295620>

Authors

Ahmad K. Ahmad — PhD, Full Professor, Al-Nahrain University, College of Engineering, Baghdad, 10072, Iraq, [sc 56756921400](https://orcid.org/0000-0003-3522-4701), <https://orcid.org/0000-0003-3522-4701>, Ahmad.ahmad@nahrainuniv.edu.iq

Ammar H. Mohammed — PhD, Lecturer, Al-Nahrain University, College of Science, Baghdad, 10072, Iraq, [sc 57191197465](https://orcid.org/0000-0003-3646-4790), <https://orcid.org/0000-0003-3646-4790>, ammahussen659@gmail.com

Alexander A. Skaptsov — PhD (Physics & Mathematics), Associate Professor, Saratov State University, Saratov, 410012, Russian Federation, [sc 22951907500](https://orcid.org/0000-0001-9336-6885), <https://orcid.org/0000-0001-9336-6885>, skaptsov@yandex.ru

Received 29.07.2022

Approved after reviewing 09.09.2022

Accepted 09.11.2022

Авторы

Ахмад Ахмад Карнал — PhD, профессор, профессор, Аль-Нарейн Университи, Инженерный колледж, Багдад, 10072, Ирак, [sc 56756921400](https://orcid.org/0000-0003-3522-4701), <https://orcid.org/0000-0003-3522-4701>, ahmad.ahmad@nahrainuniv.edu.iq

Мохаммед Аммар Хуссейн — PhD, преподаватель, Аль-Нарейн Университи, Багдад, 10072, Ирак, [sc 57191197465](https://orcid.org/0000-0003-3646-4790), <https://orcid.org/0000-0003-3646-4790>, ammahussen659@gmail.com

Скапцов Александр Александрович — кандидат физико-математических наук, доцент, Саратовский национальный исследовательский государственный университет имени Н.Г. Чернышевского, Саратов, 410012, Российская Федерация, [sc 22951907500](https://orcid.org/0000-0001-9336-6885), <https://orcid.org/0000-0001-9336-6885>, skaptsov@yandex.ru

Статья поступила в редакцию 29.07.2022

Одобрена после рецензирования 09.09.2022

Принята к печати 09.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1085-1091

УДК 539.23

Особенности импульсного лазерного напыления тонких пленок InGaAsN в атмосфере активного фоновых газа

Олег Васильевич Девицкий^{1,2}✉

¹ Федеральный исследовательский центр Южный научный центр Российской академии наук, Ростов-на-Дону, 344006, Российская Федерация

² Северо-Кавказский федеральный университет, Ставрополь, 355017, Российская Федерация

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

Аннотация

Предмет исследования. Соединения III-V-N являются перспективным классом твердых растворов, которые могут активно использоваться в оптоэлектронных приборах, работающих в широком спектральном диапазоне вплоть до 3 мкм. Соединения также применяются для увеличения эффективности фотодетекторов, лазеров в волоконно-оптических линиях связи и телекоммуникационных системах. В работе исследованы особенности различных способов получения новых полупроводниковых материалов III-V-N. Тонкие пленки InGaAsN получены методом импульсного лазерного напыления на подложках GaAs (100) и Si (100) в атмосфере активного фоновых газа. **Метод.** Импульсное лазерное напыление тонких пленок InGaAsN проведено с использованием мишени $\text{In}_{0,02}\text{Ga}_{0,98}\text{As}$ в атмосфере особо чистой аргоно-азотной смеси при давлениях 2, 5 и 10 Па. В качестве источника лазерного излучения использовался лазер AYG:Nd^{3+} с длиной волны 532 нм (вторая гармоника), плотностью энергии лазерного излучения 2,3 Дж/см², частотой следования импульсов 15 Гц и длительностью импульса 10 нс. Температура подложки составила 350 °C, время напыления 60 мин. Основные результаты. Показано, что поверхность тонких пленок текстурирована микрокаплями. Установлено, что микрокапли на поверхности тонкой пленки образованы металлическим индием, а их распределение по поверхности пленки в основном упорядочено в виде линий. Подобное явление можно объяснить наличием дислокаций несоответствия. Средний размер микрокапель на поверхности пленки InGaAsN на GaAs (100) около 30 нм, а их плотность не превышала 0,076 мкм⁻². Для сравнения в пленках, полученных на Si подложках при давлении 2 Па, наибольшая плотность микрокапель — 0,26 мкм⁻². Наименьшая плотность микрокапель на поверхности (0,17 мкм⁻²) зафиксирована в образцах тонкой пленки InGaAsN на Si (100), полученной при давлении 10 Па. Отмечено, что интенсивность локальной фоновой колебательной моды (Local Vibrational Modes, LVM) InN на частоте 430 см⁻¹ возрастает с увеличением давления аргоно-азотной смеси при импульсном лазерном напылении в спектрах комбинационного рассеяния пленок InGaAsN на Si. Обнаружены фоновые моды второго порядка: LVM InN на частоте 450 см⁻¹ и LVM GaN — 470 см⁻¹. Данное обнаружение подтверждает наличие азота в тонкой пленке InGaAsN, полученной методом импульсного лазерного напыления. Показано, что увеличение давления аргоно-азотной газовой смеси при импульсном лазерном напылении способствует увеличению концентрации азота в тонких пленках InGaAsN на Si. Установлено, что концентрация азота в пленках InGaAsN, полученных при давлении 10 Па на подложках GaAs (100) и Si (100), различается незначительно и составляет 1,9 и 1,8 % соответственно. Практическая значимость. Представленные результаты могут быть использованы при создании на основе полученных тонких пленок InGaAsN высокоэффективных фотоэлектрических преобразователей, фотодетекторов ближнего и среднего инфракрасного диапазона до 3 мкм.

Ключевые слова

разбавленные нитриды, InGaAsN, импульсное лазерное напыление, многокомпонентные твердые растворы, III-V-N/Si

Благодарности

Публикация подготовлена в рамках реализации государственного задания «Разработка новых полупроводниковых материалов на основе многокомпонентных твердых растворов для фотонных, оптоэлектронных и СВЧ применений» (номер государственной регистрации 122020100326-7), а также с использованием ресурсов центра коллективного пользования Северо-Кавказского федерального университета и при финансовой поддержке

Минобрнауки России, уникальный идентификатор проекта RF-2296.61321X0029 (соглашение № 075-15-2021-687).

Ссылка для цитирования: Девицкий О.В. Особенности импульсного лазерного напыления тонких пленок InGaAsN в атмосфере активного фоновго газа // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1085–1091. doi: 10.17586/2226-1494-2022-22-6-1085-1091

Peculiarities of pulsed laser deposition of thin InGaAsN films in an active background gas atmosphere

Oleg V. Devitsky^{1,2}✉

¹ Federal Research Center the Southern Scientific Centre of the Russian Academy of Sciences, Rostov-on-Don, 344006, Russian Federation

² North-Caucasus Federal University, Stavropol, 355017, Russian Federation
v2517@rambler.ru✉, <https://orcid.org/0000-0003-3153-696X>

Abstract

III-V-N compounds are a promising class of solid solutions that have the prospect of being used in optoelectronic devices operating in a wide spectral range up to 3 μm , as well as for increasing the efficiency of photodetectors, lasers in fiber-optic communication lines and telecommunication systems. In this work, the features of various methods for obtaining new III-V-N semiconductor materials are investigated. Thin InGaAsN films were obtained by pulsed laser deposition on GaAs (100) and Si (100) substrates in an active background gas atmosphere. Pulsed laser deposition of thin InGaAsN films was carried out using an $\text{In}_{0.02}\text{Ga}_{0.98}\text{As}$ target in an atmosphere of a highly pure argon-nitrogen mixture at a pressure of 2, 5, and 10 Pa. The source of laser radiation was an AYG:Nd³⁺ laser with a wavelength of 532 nm (second harmonic), laser radiation energy density of 2.3 J/cm², pulse repetition rate of 15 Hz, and pulse duration of 10 ns. The substrate temperature was 350 °C, the deposition time was 60 minutes. It is shown that the surface of thin films is textured with microdroplets. It has been established that microdroplets on the surface of a thin film are formed by metallic indium. It has been established that the distribution of indium microdroplets over the film surface is mainly ordered in the form of lines. This phenomenon can be explained by the presence of misfit dislocations. The average size of microdroplets on the surface of the InGaAsN film on GaAs (100) was about 30 nm, and their density did not exceed 0.076 μm^{-2} . For comparison, in films obtained on Si substrates at a pressure of 2 Pa, the highest microdroplet density was 0.26 μm^{-2} . The lowest density of microdroplets on the surface (0.17 μm^{-2}) was noted in samples of a thin film of InGaAsN on Si (100) obtained at a pressure of 10 Pa. It is noted that the intensity of the local phonon vibrational mode LVM InN at a frequency of 430 cm^{-1} increases with increasing pressure of the argon-nitrogen mixture during pulsed laser deposition in the Raman scattering spectra of InGaAsN films on Si. In the Raman spectra of InGaAsN films on Si, second-order phonon modes LVM InN and LVM GaN were detected at frequencies of 450 cm^{-1} about 470 cm^{-1} , respectively. This confirms the presence of nitrogen in a thin InGaAsN film obtained by pulsed laser deposition. It is shown that an increase in the pressure of the argon-nitrogen gas mixture during pulsed laser deposition contributes to an increase in the nitrogen concentration in thin InGaAsN films on Si. It has been established that the nitrogen concentration in InGaAsN films obtained at a pressure of 10 Pa on GaAs (100) and Si (100) substrates differs insignificantly and amounts to 1.9 % and 1.8 %, respectively. The presented results will make it possible to create highly efficient photoelectric converters and photodetectors for the near and mid-infrared range up to 3 μm based on the obtained InGaAsN thin films.

Keywords

dilute nitrides, InGaAsN, pulsed laser deposition, multicomponent solid solutions, III-V-N/Si

Acknowledgments

The publication was prepared as part of the implementation of the state task “Development of new semiconductor materials based on multicomponent solid solutions for photonic, optoelectronic and microwave applications” (state registration number 122020100326-7), as well as using the resources of the North Caucasus Federal University Shared Use Center and with financial support of the Ministry of Education and Science of the Russian Federation, unique project identifier RF-2296.61321X0029 (No 075-15-2021-687).

For citation: Devitsky O.V. Peculiarities of pulsed laser deposition of thin InGaAsN films in an active background gas atmosphere. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1085–1091 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1085-1091

Введение

В начале XXI века произошло стремительное развитие методов получения эпитаксиальных тонких пленок и гетероструктур соединений III-V-N (разбавленные нитриды) с содержанием азота на уровне нескольких процентов (GaAsN, GaPN) [1]. Наиболее перспективный четырехкомпонентный твердый раствор — InGaAsN, который может быть выращен при условии согласования по постоянной решетке к подложке GaAs, или с

напряжениями на подложке Si. Приборное применение тонких пленок InGaAsN на подложках GaAs и Si нашло в диодных лазерах [2], излучающих в диапазоне 1,3–1,55 мкм, сведодиодах и солнечных элементах [3]. Основные методы получения тонких пленок разбавленных нитридов на сегодняшний день — молекулярно-лучевая эпитаксия и осаждение металлоорганических соединений из газовой фазы. С помощью этих методов получены эпитаксиальные тонкие пленки InGaAsN/GaAs с долей азота около 2 % [4]. Отметим, что при

увеличении доли азота структурные свойства тонких пленок ухудшили свое качество на столько, что это лишило их приборного применения.

В работе [5] установлено, что основная причина данного ухудшения — неравновесные условия получения разбавленных нитридов при относительно низких температурах, которые являются необходимыми для встраивания атомов азота в соединения III-V. Из-за этого в тонких пленках с повышенным содержанием азота (от 2 до 4 %) происходит достаточно сильное дефектообразование. Проблема дефектообразования в тонких пленках разбавленных нитридов — критический фактор их приборного применения, который требует простых и эффективных методов решения [6]. Одним из таких способов может стать применение метода импульсного лазерного напыления (ИЛН) [7, 8] для получения разбавленных нитридов на различных типах подложек. ИЛН, в отличие молекулярно-лучевой эпитаксии, является дискретным методом. К тому же при ИЛН разбавленных нитридов источником азота является газовая смесь особо чистого аргона и молекулярного азота (активный фоновый газ). Активация азота происходит за счет образования плазменного факела при воздействии лазерных импульсов на поверхность мишени. В работе [9] получены эпитаксиальные слои $\text{GaN}_x\text{As}_{1-x}$ на подложках GaAs методом ИЛН мишени GaAs в атмосфере аммиака. Дифракция рентгеновских лучей с высоким разрешением показала, что существует пороговое давление аммиака, выше которого концентрация азота в пленке увеличивается линейно вместе с увеличением давления аммиака. Отметим, что в современных научных работах отсутствуют исследования по получению пленок InGaAsN методом ИЛН на подложках GaAs и Si, но показана перспектива создания приборов на их основе [10].

Цель настоящей работы — получение тонких пленок InGaAsN на подложках Si и GaAs и исследование влияния давления аргоно-азотной смеси при ИЛН на структуру тонких пленок.

Материалы и методы

ИЛН тонких пленок InGaAsN на подложках GaAs (001) и Si (001) выполнено из мишени $\text{In}_{0,02}\text{Ga}_{0,98}\text{As}$, сформированной методом холодного прессования. Для получения мишени использованы порошки GaAs и InAs, которые были перетерты в течение 2 ч, а затем просеяны на ситах с размером ячейки 20 мкм. Далее при помощи одноосного прессования под давлением 207 МПа сформирована монолитная мишень $\text{In}_{0,02}\text{Ga}_{0,98}\text{As}$.

Лазерное излучение с длиной волны 532 нм сфокусировано на поверхности мишени под углом 45° . Частота повторения импульсов составила 15 Гц, длительность лазерного импульса 10 нс, расходимость лазерного излучения 1–2 мрад. При помощи линзы с фокусным расстоянием 190 мм лазерное излучение сфокусировано на поверхности мишени в пятно диаметром 0,34 мм, плотность энергии импульса при этом составила 2,3 Дж/см². Значение плотности энергии лазерного излучения выбрана исходя из экспериментальных

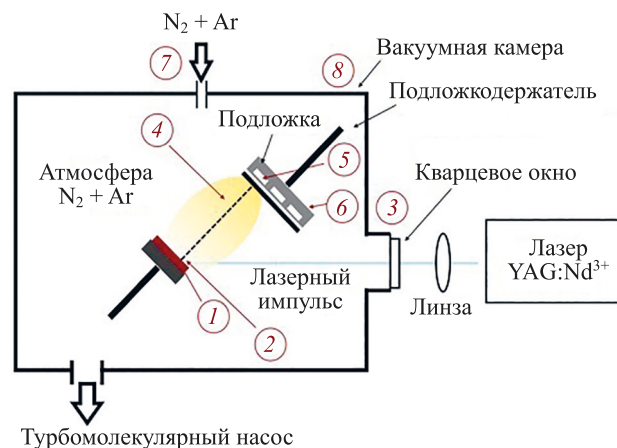


Рис. 1. Схема импульсного лазерного напыления тонких пленок InGaAsN на Si и GaAs: 1 — вращающийся держатель мишени; 2 — мишень $\text{In}_{0,02}\text{Ga}_{0,98}\text{As}$; 3 — кварцевое окно; 4 — плазменный факел; 5 — подложка Si или GaAs; 6 — подложкодержатель с нагревателем; 7 — подвод аргоно-азотной газовой смеси; 8 — вакуумная камера

Fig. 1. Scheme of pulsed laser deposition of InGaAsN thin films on Si and GaAs: 1 — rotating target holder; 2 — $\text{In}_{0,02}\text{Ga}_{0,98}\text{As}$ target; 3 — quartz window; 4 — plasma torch; 5 — Si or GaAs substrate; 6 — substrate holder with a heater; 7 — supply of argon-nitrogen gas mixture; 8 — vacuum chamber

результатов, описанных в работе [7]. Расстояние от мишени до подложки 50 мм, а толщина напыляемых слоев находится в диапазоне 150–170 нм. Процесс напыления осуществлен при температуре подложки 400 °C. Схема процесса ИЛН показана на рис. 1.

На начальном этапе объем вакуумной камеры откачен до 10^{-4} Па, далее объем камеры был изолирован, и добавлена смесь особо чистого аргона и азота. Объемная доля азота в аргоно-азотной смеси составляла 80 %, а значения давления смеси: 2, 5 и 10 Па.

С помощью сканирующего электронного микроскопа (СЭМ) MIRA3-LMN с системой определения элементного состава AZtecEnergy Standart/X-max 20, выполнено исследование СЭМ-изображений и измерение состава смеси на поверхности мишени и образцов тонких пленок.

Спектры комбинационного рассеяния света получены с использованием спектрометра inVia Raman Microscope Renishaw.

Результаты и обсуждение

На рис. 2, а, б представлены СЭМ-изображения морфологии поверхности тонкой пленки InGaAsN на подложках GaAs (001), полученной при давлении аргоно-азотной газовой смеси 10 Па. Из представленных изображений видно, что поверхность пленки сплошная и состоит из зерен, что характерно для пленок с поликристаллической структурой. Кроме этого, на поверхности всех образцов тонких пленок присутствуют микрокапли (рис. 2, б), наличие которых типично для ИЛН.

По результатам рентгеноэнергодисперсионного элементного анализа пленки установлено, что микрокапли

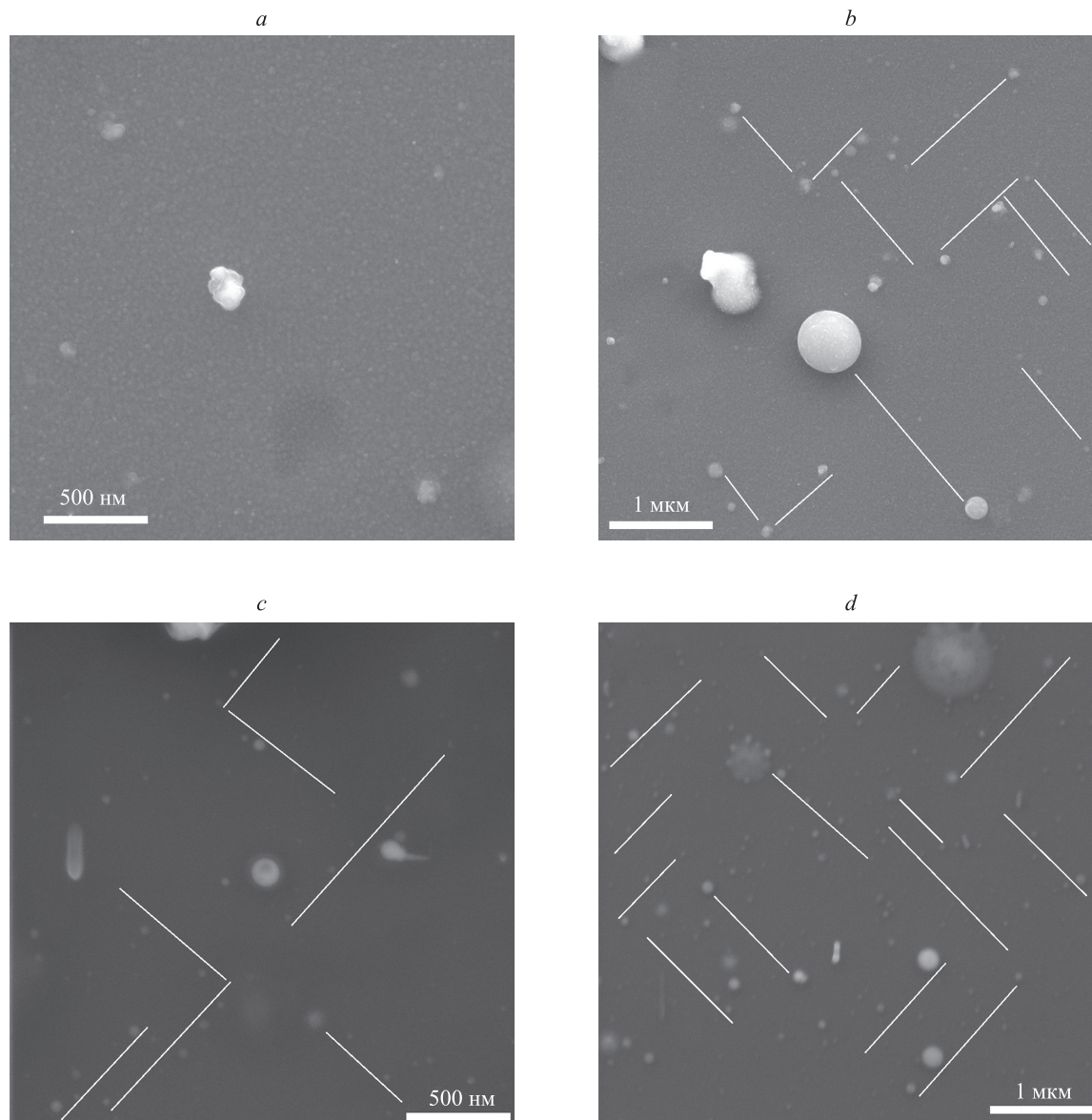


Рис. 2. СЭМ-изображения поверхности тонкой пленки InGaAsN на GaAs (a) и кремниевой (b) подложках и микрокапель индия (c, d)

Fig. 2. SEM images of the surface of a thin InGaAsN film on (a) GaAs and (b) silicon substrates and (c, d) indium microdroplets

образованы металлическим индием, а концентрация азота в пленке InGaAsN составляет 1,9 %. Средний размер микрокапель на поверхности пленки составил около 30 нм, а их плотность не превышает $0,076 \text{ мкм}^{-2}$. Установлено, что распределение микрокапель индия в основном упорядочено в виде линий, которые показаны на рисунке белым цветом. Предположим, что это связано с наличием дислокаций несоответствия, так как расогласования постоянных решетки пленки и подложки ($\Delta a/a$) составляют для InGaAsN на GaAs около 0,23 %, для InGaAsN на Si не более 3,9 %.

Для тонких пленок InGaAsN, полученных на кремниевых подложках, заметна большая плотность микрокапель на поверхности. Отметим, что при увеличении давления аргоно-азотной газовой смеси от 2 до 10 Па плотность микрокапель индия уменьшилась. Для образца пленки при давлении 2 Па получена наибольшая

плотность микрокапель — $0,26 \text{ мкм}^{-2}$. Наименьшая плотность микрокапель $0,17 \text{ мкм}^{-2}$ отмечена для образца тонкой пленки InGaAsN при давлении 10 Па (рис. 2, c).

Концентрация азота в пленках InGaAsN на Si, полученных при давлениях 2 и 5 Па, составила 1,1 %, в то время как при давлении 10 Па отмечена наибольшая концентрация азота — 1,8 %.

На рис. 3, a изображены спектры комбинационного рассеяния тонких пленок InGaAsN на подложке Si при давлении аргоно-азотной смеси при значениях 2, 5 и 10 Па. На спектрах комбинационного рассеяния наибольшей интенсивностью обладает мода длинноволнового поперечного оптического фона кремниевой подложки LO на частоте 512 см^{-1} . Кроме этого, присутствуют менее интенсивные моды кремния, связанные с поперечными акустическими фоновыми первого

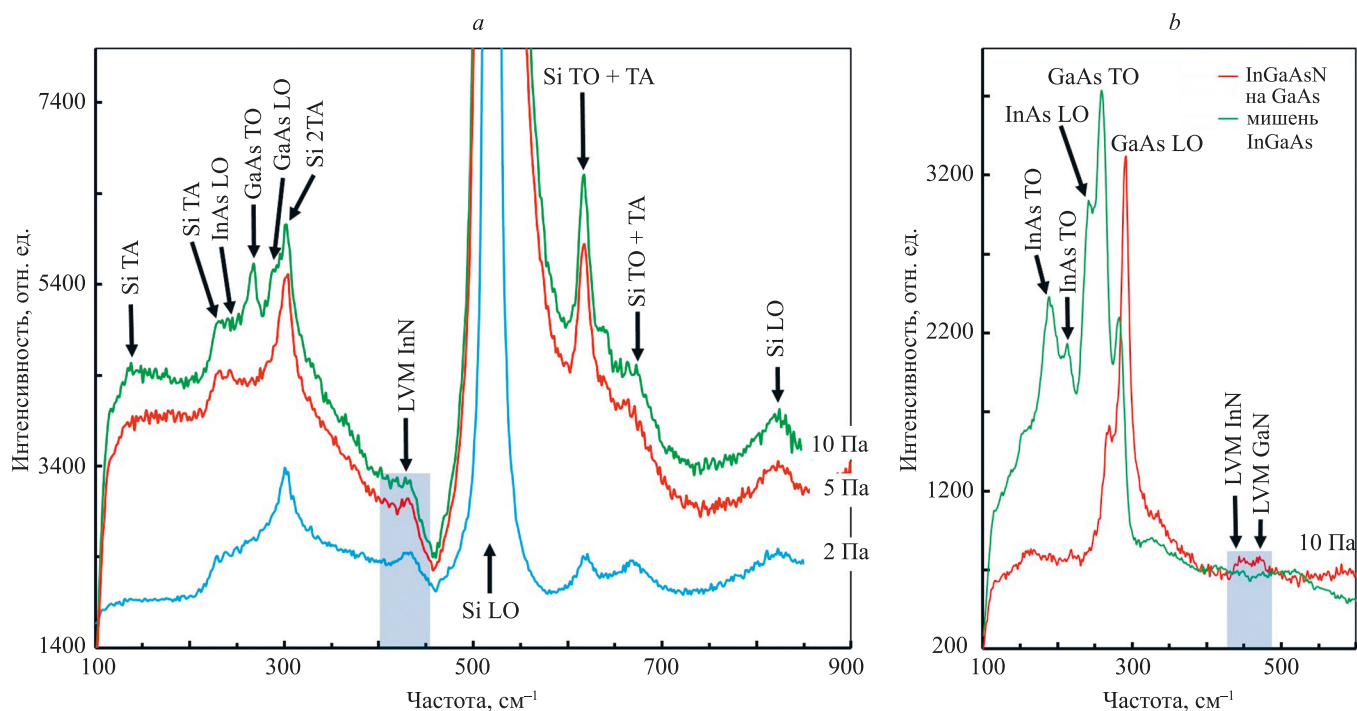


Рис. 3. Спектры комбинационного рассеяния света: тонких пленок InGaAsN на подложке Si (a), мишени InGaAs и тонкой пленки InGaAsN на подложке GaAs (b), полученные при воздействии давления аргонно-азотной газовой смеси

Fig. 3. Raman spectra of InGaAsN thin films on Si substrates (a), InGaAs target and InGaAsN thin film on GaAs (b), obtained at the pressure of an argon-nitrogen gas mixture

(transverse acoustic) TA на частоте 240 см^{-1} и второго порядков 2TA на частоте 300 см^{-1} , комбинация поперечных оптических (transverse optical, TO) и акустических мод TO + TA наблюдается на частотах 615 см^{-1} и 672 см^{-1} соответственно [11].

На спектрах рис. 3, a идентифицируются продольные LO- и поперечные TO-моды GaAs и InAs [12]. Интенсивность мод LO-InAs, LO-GaAs и TO-GaAs увеличивается при повышении давления аргонно-азотной смеси. Предположительно, это связано с наличием фазовых включений в пленках. Наличие на спектре вибрационной фононной колебательной моды (Local Vibrational Modes, LVM) InN на частоте 430 см^{-1} подтверждает вхождение атомов азота в InGaAs [13]. В связи с низким содержанием азота в пленке интенсивность LVM InN относительно невелика, однако видно, что с ростом давления она возрастает. Полученные данные совпадают с результатами, представленными в работах [14, 15].

На рис. 3, b представлены спектры комбинационного рассеяния мишени $\text{In}_{0,02}\text{Ga}_{0,98}\text{As}$ и тонкой пленки InGaAsN, полученной при давлении аргонно-азотной смеси 10 Па. На спектрах тонкой пленки InGaAsN хорошо заметны высокоинтенсивные LO-GaAs 291 см^{-1}

и TO-GaAs 283 см^{-1} , связанные с подложкой GaAs. Также можно идентифицировать моду LVM InN второго порядка на частоте 450 см^{-1} и LVM GaN — 470 см^{-1} , возможно они связаны с такими конфигурациями, как In_4N и Ga_4N [16], которые присутствуют в полученной тонкой пленке InGaAsN.

Заключение

Анализируя экспериментальные данные, полученные в рамках настоящей работы, можно сделать вывод, что увеличение давления аргонно-азотной газовой смеси при импульсном лазерном напылении способствует увеличению концентрации азота в тонких пленках InGaAsN на подложке Si. Показано, что концентрация азота в пленках InGaAsN, полученных при давлении 10 Па на подложках GaAs и Si, различается незначительно и составляет 1,9 и 1,8 % соответственно. В спектрах комбинационного рассеяния света тонких пленок обнаружены локальные колебательные моды LVM InN и LVM GaN, что подтверждает наличие азота в тонкой пленке InGaAsN, полученной методом импульсного лазерного напыления.

Литература

References

1. Park Y., Cich M.J., Zhao R., Specht P., Feick H., Weber E.R. AFM study of lattice matched and strained InGaAsN layers on GaAs // *Physica B: Condensed Matter*. 2001. V. 308–310. P. 98–101. [https://doi.org/10.1016/S0921-4526\(01\)00669-X](https://doi.org/10.1016/S0921-4526(01)00669-X)
2. Kim T., Garrod T.J., Kim K., Lee J., Mawst L.J., Kuech T.F., LaLumondiere S.D., Sin Y., Lotshaw W.T., Moss S.C. Characteristics of bulk InGaAsN and InGaAsSbN materials grown by metal organic vapor phase epitaxy (MOVPE) for solar cell application // *Proceedings of SPIE*. 2012. V. 8256. P. 82561D. <https://doi.org/10.1117/12.906961>
3. Levillayer M., Arnoult A., Massiot I., Duzellier S., Nuns T., Inguimbert C., Aicardi C., Parola S., Carrère H., Balocchi A., Vaissiere N., Decobert J., Almuneau G., Artola L. As-grown InGaAsN subcells for multijunction solar cells by molecular beam epitaxy // *IEEE Journal of Photovoltaics*. 2021. V. 11. N 5. P. 1271–1277. <https://doi.org/10.1109/JPHOTOV.2021.3093048>
4. Pavelescu E.M., Wagner J., Komsa H.P., Rantala T.T., Dumitrescu M., Pessa M. Nitrogen incorporation into GaInNAs lattice-matched to GaAs: The effects of growth temperature and thermal annealing // *Journal of Applied Physics*. 2005. V. 98. N 8. P. 83524. <https://doi.org/10.1063/1.2112173>
5. Bank S.R., Wistey M.A., Yuen H.B., Goddard L.L., Bae H., Harris J.S. Molecular-beam epitaxy growth of low-threshold cw GaInNAsSb lasers at 1.5 μm // *Journal of Vacuum Science & Technology B*. 2005. V. 23. N 3. P. 1337–1340. <https://doi.org/10.1116/1.1914825>
6. Baranov A.I., Gudovskikh A.S., Kudryashov D.A., Lazarenko A.A., Morozov I.A., Mozharov A.M., Nikitina E.V., Pirogov E.V., Sobolev M.S., Zelentsov K.S., Egorov A.Yu., Darga A., Le Gall S., Kleider J.-P. Defect properties of InGaAsN layers grown as submonolayer digital alloys by molecular beam epitaxy // *Journal of Applied Physics*. 2018. V. 123. N 16. P. 161418. <https://doi.org/10.1063/1.5011371>
7. 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>
8. Девицкий О.В., Никулин Д.А., Сысоев И.А. Импульсное лазерное напыление тонких пленок нитрида алюминия на сапфировые подложки // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 2. С. 177–184. <https://doi.org/10.17586/2226-1494-2020-20-2-177-184>
9. Hung W.K., Chern M.Y., Fan J.C., Lin T.Y., Chen Y.F. Pulsed laser deposition of epitaxial GaN_xAs_{1-x} on GaAs // *Applied Physics Letters*. 1999. V. 74. N 26. P. 3951–3953. <https://doi.org/10.1063/1.124234>
10. Девицкий О.В., Санакулов С.О. Численное моделирование функциональных характеристик солнечных элементов на основе гетероструктур InGaAsN/Si // Научно-технический вестник информационных технологий, механики и оптики. 2021. Т. 21. № 2. С. 191–197. <https://doi.org/10.17586/2226-1494-2021-21-2-191-197>
11. Iatsunskiy I., Jurga S., Smyntyna V., Pavlenko M., Myndrul V., Zaleska A. Raman spectroscopy of nanostructured silicon fabricated by metal-assisted chemical etching // *Proceedings of SPIE*. 2014. V. 9132. P. 913217. <https://doi.org/10.1117/12.2051489>
12. Chafi A., Pagès O., Postnikov A.V., Gleize J., Sallet V., Rzepka E., Li L.H., Jusserand B., Harmand J.C. Combined Raman study of InGaAsN from the N-impurity and InGaAs-matrix sides // *Applied Physics Letters*. 2007. V. 91. N 5. P. 051910. <https://doi.org/10.1063/1.2767244>
13. Milanova M., Donchev V., Kostov K.L., Alonso-Álvarez D., Valcheva E., Kirilov K., Asenova I., Ivanov I.G., Georgiev S., Ekins-Daukes N. Experimental study of the effect of local atomic ordering on the energy band gap of melt grown InGaAsN alloys // *Semiconductor Science and Technology*. 2017. V. 32. N 8. P. 085005. <https://doi.org/10.1088/1361-6641/aa7404>
14. Alt H.C., Gomeniuk Y.V. Local mode frequencies of the NAs–InGa nearest-neighbor pair in (Ga,In)(As,N) alloys // *Physical Review B*. 2004. V. 70. N 16. P. 161314(R). <https://doi.org/10.1103/PhysRevB.70.161314>
15. Kurtz S., Webb J., Gedvilas L., Friedman D., Geisz J., Olson J., King R., Joslin D., Karam N. Structural changes during annealing of
1. Park Y., Cich M.J., Zhao R., Specht P., Feick H., Weber E.R. AFM study of lattice matched and strained InGaAsN layers on GaAs. *Physica B: Condensed Matter*, 2001, vol. 308–310, pp. 98–101. [https://doi.org/10.1016/S0921-4526\(01\)00669-X](https://doi.org/10.1016/S0921-4526(01)00669-X)
2. Kim T., Garrod T.J., Kim K., Lee J., Mawst L.J., Kuech T.F., LaLumondiere S.D., Sin Y., Lotshaw W.T., Moss S.C. Characteristics of bulk InGaAsN and InGaAsSbN materials grown by metal organic vapor phase epitaxy (MOVPE) for solar cell application. *Proceedings of SPIE*, 2012, vol. 8256, pp. 82561D. <https://doi.org/10.1117/12.906961>
3. Levillayer M., Arnoult A., Massiot I., Duzellier S., Nuns T., Inguimbert C., Aicardi C., Parola S., Carrère H., Balocchi A., Vaissiere N., Decobert J., Almuneau G., Artola L. As-grown InGaAsN subcells for multijunction solar cells by molecular beam epitaxy. *IEEE Journal of Photovoltaics*, 2021, vol. 11, no. 5, pp. 1271–1277. <https://doi.org/10.1109/JPHOTOV.2021.3093048>
4. Pavelescu E.M., Wagner J., Komsa H.P., Rantala T.T., Dumitrescu M., Pessa M. Nitrogen incorporation into GaInNAs lattice-matched to GaAs: The effects of growth temperature and thermal annealing. *Journal of Applied Physics*, 2005, vol. 98, no. 8, pp. 83524. <https://doi.org/10.1063/1.2112173>
5. Bank S.R., Wistey M.A., Yuen H.B., Goddard L.L., Bae H., Harris J.S. Molecular-beam epitaxy growth of low-threshold cw GaInNAsSb lasers at 1.5 μm . *Journal of Vacuum Science & Technology B*, 2005, vol. 23, no. 3, pp. 1337–1340. <https://doi.org/10.1116/1.1914825>
6. Baranov A.I., Gudovskikh A.S., Kudryashov D.A., Lazarenko A.A., Morozov I.A., Mozharov A.M., Nikitina E.V., Pirogov E.V., Sobolev M.S., Zelentsov K.S., Egorov A.Yu., Darga A., Le Gall S., Kleider J.-P. Defect properties of InGaAsN layers grown as submonolayer digital alloys by molecular beam epitaxy. *Journal of Applied Physics*, 2018, vol. 123, no. 16, pp. 161418. <https://doi.org/10.1063/1.5011371>
7. 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>
8. Devitsky O.V., Nikulin D.A., Sysoev I.A. Pulsed laser deposition of aluminum nitride thin films onto sapphire substrates. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 2, pp. 177–184. (in Russian). <https://doi.org/10.17586/2226-1494-2020-20-2-177-184>
9. Hung W.K., Chern M.Y., Fan J.C., Lin T.Y., Chen Y.F. Pulsed laser deposition of epitaxial GaN_xAs_{1-x} on GaAs. *Applied Physics Letters*, 1999, vol. 74, no. 26, pp. 3951–3953. <https://doi.org/10.1063/1.124234>
10. Devitsky O.V., Sanakulov S.O. Numerical simulation of functional characteristics of solar elements InGaAsN/Si. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 2, pp. 191–197. (in Russian). <https://doi.org/10.17586/2226-1494-2021-21-2-191-197>
11. Iatsunskiy I., Jurga S., Smyntyna V., Pavlenko M., Myndrul V., Zaleska A. Raman spectroscopy of nanostructured silicon fabricated by metal-assisted chemical etching. *Proceedings of SPIE*, 2014, vol. 9132, pp. 913217. <https://doi.org/10.1117/12.2051489>
12. Chafi A., Pagès O., Postnikov A.V., Gleize J., Sallet V., Rzepka E., Li L.H., Jusserand B., Harmand J.C. Combined Raman study of InGaAsN from the N-impurity and InGaAs-matrix sides. *Applied Physics Letters*, 2007, vol. 91, no. 5, pp. 051910. <https://doi.org/10.1063/1.2767244>
13. Milanova M., Donchev V., Kostov K.L., Alonso-Álvarez D., Valcheva E., Kirilov K., Asenova I., Ivanov I.G., Georgiev S., Ekins-Daukes N. Experimental study of the effect of local atomic ordering on the energy band gap of melt grown InGaAsN alloys. *Semiconductor Science and Technology*, 2017, vol. 32, no. 8, pp. 085005. <https://doi.org/10.1088/1361-6641/aa7404>
14. Alt H.C., Gomeniuk Y.V. Local mode frequencies of the NAs–InGa nearest-neighbor pair in (Ga,In)(As,N) alloys. *Physical Review B*, 2004, vol. 70, no. 16, pp. 161314(R). <https://doi.org/10.1103/PhysRevB.70.161314>
15. Kurtz S., Webb J., Gedvilas L., Friedman D., Geisz J., Olson J., King R., Joslin D., Karam N. Structural changes during annealing of

- GaInAsN // *Applied Physics Letters*. 2001. V. 78. N 6. P. 748–750. <https://doi.org/10.1063/1.1345819>
16. De La Mare M., Das S.C., Das T.D., Dhar S., Krier A. N incorporation and photoluminescence in In-rich InGaAsN grown on InAs by liquid phase epitaxy // *Journal of Physics D: Applied Physics*. 2011. V. 44. N 31. P. 315102. <https://doi.org/10.1088/0022-3727/44/31/315102>

Автор

Девицкий Олег Васильевич — кандидат технических наук, старший 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

Статья поступила в редакцию 22.08.2022
Одобрена после рецензирования 15.09.2022
Принята к печати 09.11.2022

- GaInAsN. *Applied Physics Letters*, 2001, vol. 78, no. 6, pp. 748–750. <https://doi.org/10.1063/1.1345819>
16. De La Mare M., Das S.C., Das T.D., Dhar S., Krier A. N incorporation and photoluminescence in In-rich InGaAsN grown on InAs by liquid phase epitaxy. *Journal of Physics D: Applied Physics*, 2011, vol. 44, no. 31, pp. 315102. <https://doi.org/10.1088/0022-3727/44/31/315102>

Author

Oleg V. Devitsky — PhD, Senior 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

Received 22.08.2022
Approved after reviewing 15.09.2022
Accepted 09.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1092-1097

УДК 538.911, 621.315.592

Нахождение распределения электронов в сверхрешетках AlGaAs/GaAs с узкими барьерами методом вольт-фарадного профилирования

Елена Игоревна Василькова^{1✉}, Евгений Викторович Пирогов²,
Максим Сергеевич Соболев³, Артем Игоревич Баранов⁴,
Александр Сергеевич Гудовских⁵, Алексей Дмитриевич Буравлев⁶

^{1,2,3,4,5} Санкт-Петербургский национальный исследовательский Академический университет им. Ж.И. Алферова РАН, Санкт-Петербург, 194021, Российская Федерация

^{5,6} Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация

⁶ Университет при МПА ЕвразЭС, Санкт-Петербург, 194044, Российская Федерация

⁶ Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация

¹ elenvasilkov@gmail.com ✉, <https://orcid.org/0000-0002-0349-7134>

² zzzavr@gmail.com, <https://orcid.org/0000-0001-7186-3768>

³ sobolevsm@gmail.com, <https://orcid.org/0000-0001-8629-2064>

⁴ itiomchik@yandex.ru, <https://orcid.org/0000-0002-4894-6503>

⁵ gudovskikh@spbau.ru, <https://orcid.org/0000-0002-7632-3194>

⁶ bour@mail.ioffe.ru, <https://orcid.org/0000-0002-7432-2091>

Аннотация

Предмет исследования. Исследовано распределение концентрации электронов в однородно легированных кремнием сверхрешетках AlGaAs/GaAs с толщинами слоев 1,5/10 нм и различным количеством квантовых ям. **Метод.** Структуры с одинаковыми параметрами слоев и уровнем легирования, содержащие 3, 5 и 25 периодов, выращены на установке молекулярно-пучковой эпитаксии. Профили концентрации носителей в структурах определены методом вольт-фарадного профилирования экспериментально и с помощью численного моделирования. **Основные результаты.** В результате анализа экспериментальных вольт-фарадных характеристик получено, что концентрация носителей заряда растет с увеличением числа квантовых ям от $7,1 \cdot 10^{16} \text{ см}^{-3}$ (для трех ям) до $9,2 \cdot 10^{16} \text{ см}^{-3}$ (для 25-ти ям) при уровне легирования 10^{17} см^{-3} . На профилях концентрации у части образцов наблюдаются плоские участки насыщения в областях, соответствующих сверхрешетке. Концентрации, полученные из компьютерного моделирования, соответствуют экспериментальным данным с точностью в пределах 10 %. **Практическая значимость.** Вольт-фарадное профилирование позволило определить профиль концентрации носителей по глубине в сверхрешетках с узкими барьерами. Несмотря на то, что метод дает представление о распределении «кажущейся» концентрации носителей, его можно использовать для оценки характера распределения легирующей примеси в гетероструктурах с сильно связанными квантовыми ямами.

Ключевые слова

вольт-фарадное профилирование, сверхрешетки AlGaAs/GaAs, распределение носителей, молекулярно-пучковая эпитаксия

Благодарности

Работа частично поддержана Министерством образования и науки Российской Федерации (FSRM-2020-0008) и Российским фондом фундаментальных исследований (19-29-12053).

Ссылка для цитирования: Василькова Е.И., Пирогов Е.В., Соболев М.С., Баранов А.И., Гудовских А.С., Буравлев А.Д. Нахождение распределения электронов в сверхрешетках AlGaAs/GaAs с узкими барьерами методом вольт-фарадного профилирования // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1092–1097. doi: 10.17586/2226-1494-2022-22-6-1092-1097

Determination of the electron distribution in thin barrier AlGaAs/GaAs superlattices by capacitance-voltage profiling

Elena I. Vasilkova¹✉, Evgeny V. Pirogov², Maxim S. Sobolev³, Artem I. Baranov⁴,
Alexander S. Gudovskikh⁵, Alexei D. Bouravlev⁶

^{1,2,3,4,5} Alferov Saint Petersburg National Research Academic University of the Russian Academy of Sciences, Saint Petersburg, 194021, Russian Federation

^{5,6} Saint Petersburg Electrotechnical University “LETI”, Saint Petersburg, 197022, Russian Federation

⁶ University under the Inter-Parliamentary Assembly of EurAsEC, Saint Petersburg, 194044, Russian Federation

⁶ Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation

¹ elenvasilkov@gmail.com✉, <https://orcid.org/0000-0002-0349-7134>

² zzzavr@gmail.com, <https://orcid.org/0000-0001-7186-3768>

³ sobolevsm@gmail.com, <https://orcid.org/0000-0001-8629-2064>

⁴ itiomchik@yandex.ru, <https://orcid.org/0000-0002-4894-6503>

⁵ gudovskikh@spbau.ru, <https://orcid.org/0000-0002-7632-3194>

⁶ bour@mail.ioffe.ru, <https://orcid.org/0000-0002-7432-2091>

Abstract

Electron density distribution in uniformly doped AlGaAs/GaAs superlattices with respective layer thicknesses 1.5/10 nm and a different number of quantum wells was investigated. Experimental samples containing 3, 5 and 25 periods with the same layer parameters were grown by molecular beam epitaxy. Capacitance-voltage profiling was used to determine the carrier concentration profiles in the structures both numerically and experimentally. During the analysis of experimental capacitance-voltage characteristics it was found that the maximum electron concentration increases with an increase in the number of quantum wells starting from $7,1 \cdot 10^{16} \text{ cm}^{-3}$ for 3 wells up to $9,2 \cdot 10^{16} \text{ cm}^{-3}$ for 25 wells with overall superlattice doping level of 10^{17} cm^{-3} . In some samples saturation areas are observed on the concentration profiles, that are associated with the region of superlattice. Concentration values, obtained from computer modeling, correspond to the experimental data with an error of less than 10 %. Capacitance-voltage profiling is a suitable technique for determining the carrier concentration profiles in thin barrier superlattices. Despite the fact that the method provides distribution of the “apparent” carrier concentration profile, it can be used to estimate the dopant atoms distribution in the strongly coupled quantum well heterostructures.

Keywords

capacitance-voltage profiling, AlGaAs/GaAs superlattices, molecular beam epitaxy

Acknowledgments

The work was partly supported by the Ministry of Science and Higher Education of the Russian Federation (FSRM-2020-0008) and the Russian Foundation for Basic Research (RFBR) (19-29-12053).

For citation: Vasilkova E.I., Pirogov E.V., Sobolev M.S., Baranov A.I., Gudovskikh A.S., Bouravlev A.D. Determination of the electron distribution in thin barrier AlGaAs/GaAs superlattices by capacitance-voltage profiling. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1092–1097 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1092-1097

Введение

На базе A_3B_5 гетероструктур с квантовыми ямами уже долгие годы реализуется множество приложений полупроводниковой сверхвысокочастотной электроники и оптоэлектроники [1, 2]. Среди применений эпитаксиальных квантоворазмерных гетероструктур отметим сверхмногопериодные сверхрешетки AlGaAs/GaAs с сильно связанными квантовыми ямами, излучающие в терагерцовом диапазоне [3]. При изготовлении подобных структур требуется осуществлять прецизионный контроль уровня легирования слоев и профиля распределения атомов примеси, так как от них напрямую зависит частота излучения и эффективность генерации в терагерцовой области спектра.

Наибольшую точность определения профиля примеси в эпитаксиальных слоях GaAs предоставляют классическое вольт-фарадное (C-V, capacitance-voltage) профилирование и метод на основе эффекта Холла [4]. Отметим, что измерение эффекта Холла дает представление лишь о среднем значении концентрации носителей, а не распределению их по структуре, которое

принципиально важно для многопериодных толстых гетероструктур, ориентированных на создание терагерцовых источников. В связи с этим в настоящей работе для исследования распределения электронов в сверхрешетке выбран метод C-V профилирования. Метод основан на зависимости ширины W и емкости C области пространственного заряда, образующейся на границе металл-полупроводник, от приложенного напряжения смещения. Профиль концентрации носителей n (а также профиль залегания примеси) косвенно определяется из вольт-фарадной характеристики по формуле:

$$n(W) = \frac{2}{q\epsilon_s\epsilon_0 S^2} \left(\frac{d(1/C^2)}{dV} \right)^{-1}, \quad (1)$$

$$W = \frac{\epsilon_s\epsilon_0 S}{C},$$

где q — заряд электрона; ϵ_s — относительная диэлектрическая проницаемость материала; ϵ_0 — диэлектрическая проницаемость вакуума; S^2 — площадь контакта металл-полупроводник; V — приложенное напряжение смещения.

Формально C-V профилирование предоставляет информацию о некой эффективной (или «кажущейся», apparent) концентрации основных носителей. Ее профиль представляет собой усредненную функцию, разрешение которой ограничивается дебаевской длиной экранирования. Кажущаяся концентрация в общем случае не равна ни истинной концентрации электронов, ни концентрации легирующей примеси в структуре, однако, их можно считать приближенно совпадающими [5]. Например, C-V профилирование применяется для нахождения профиля залегания примеси в GaAs-сверхрешетках в основном с широкими барьерами [6–8]. Исследования распределения носителей в сверхрешетках с сильно связанными квантовыми ямами с помощью C-V профилирования в научных работах представлены реже [8]. Заметим, что ранее данным методом не исследовалась зависимость концентрации электронов от числа периодов сверхрешетки.

Среди существующих экспериментальных способов определения уровня легирования в гетероструктурах метод C-V профилирования выделяется относительной трудозатратностью, так как требует наличия перехода металл-полупроводник, а значит предварительного формирования на экспериментальном образце диодной структуры типа Шоттки. Кроме того, при работе с толстыми структурами, такими как гетероструктуры с большим количеством квантовых ям, следует учитывать, что максимальная глубина профилирования ограничена вследствие возникновения электрического пробоя при высоких напряжениях обратного смещения.

Образцы и методика эксперимента

Рассмотрим A_3B_5 сверхрешетки, состоящие из однородно легированных кремнием 10^{17} см^{-3} барьеров $Al_{0.2}Ga_{0.8}As$ и квантовых ям GaAs толщиной 1,5 и 10 нм соответственно. Данные сверхрешетки имеют конфигурацию слоев, оптимальную для генерации терагерцового излучения [9], но меньшее количество периодов — 3, 5 и 25. Слоевой дизайн структур включает сильно легированный 10 нм GaAs контактный слой, а также верхний 400 нм и нижний 200 нм GaAs слои-обкладки с легированием $3 \cdot 10^{16} \text{ см}^{-3}$. Экспериментальные образцы выращены методом молекулярно-пучковой эпитаксии на установке Riber MBE49 на двухдюймовых «*epi-ready*» проводящих подложках GaAs с ориентацией (100).

Профили концентрации свободных электронов по глубине в данных структурах определены с использованием методики C-V профилирования по численно смоделированным и по экспериментальным вольт-фарадным характеристикам. Компьютерное моделирование C-V кривых выполнено с использованием программного обеспечения AFORS-HET 2.5 [10]. Экспериментальные вольт-фарадные характеристики данных образцов получены с помощью прецизионного LCR-метра KEYSIGHT E4980A-001. Измерения выполнены при комнатной температуре на частоте 1 МГц в диапазоне обратных напряжений смещения от 0 до –12 В. Для осуществления эксперимента были подготовлены тестовые структуры с массивами диодов

Шоттки, перед изготовлением которых произведено травление сильно легированного контактного слоя GaAs. Далее методом вакуумного термического напыления на поверхность открывшегося слоя-обкладки GaAs напылены круговые Au/Ge контакты Шоттки, а с тыльной стороны образцов к подложке *n*-типа сформирован сплошной омический контакт.

Результаты и обсуждение

Вид распределения концентрации носителей в выращенных образцах был теоретически предсказан на основе смоделированных вольт-фарадных характеристик и уравнения (1). На рисунке, *a* представлено сравнение профилей кажущейся концентрации, полученных из расчетных C-V кривых, для сверхрешеток с тремя, пятью и 25-ю периодами. Согласно результатам моделирования, распределение электронов имеет максимум в области сверхрешетки, где происходит аккумуляция носителей. Осцилляции значений концентрации, соответствующие областям барьеров и квантовых ям, не наблюдались, так как толщины барьеров не превысили дебаевской длины, и разрешить их по глубине невозможно. Показано, что несмотря одинаковую степень легирования во всех образцах, значения кажущейся концентрации электронов увеличивались с ростом числа периодов в структуре. Максимальные концентрации электронов в сверхрешетке по результатам моделирования для структур составили: с тремя периодами — $6,6 \cdot 10^{16} \text{ см}^{-3}$, с пятью — $8,6 \cdot 10^{16} \text{ см}^{-3}$, с 25-ю — $9,8 \cdot 10^{16} \text{ см}^{-3}$.

Для образца с 25-ю периодами распределение электронов приходит в насыщение при значении 10^{17} см^{-3} , что совпадает со степенью преднамеренного легирования слоев, образующих сверхрешетку. Исходя из протяженности участка насыщения и разрешения метода, электроны должны быть делокализованы практически по всей области сверхрешетки.

Профили кажущейся концентрации носителей заряда, представленные на рисунок, *b*, рассчитаны по экспериментальным вольт-фарадным характеристикам. Измерения вольт-фарадных характеристик для структуры с 25-ю периодами выполнены в ограниченном диапазоне толщин до 550 нм от поверхности. Это связано с ростом тока утечки при больших напряжениях обратного смещения (свыше 10 В), затрудняющим корректные измерения. Для всех образцов экспериментальная концентрация электронов в верхнем слое-обкладке GaAs с высокой точностью совпадает с заложенной степенью легирования, что позволяет говорить о достоверности измерений и возможности провести анализ профилей концентрации в интересующей нас области сверхрешетки.

Профили кажущейся концентрации имеют единственный ярко выраженный пик, соответствующий накоплению электронов в квантовых ямах, что согласуется с результатами моделирования. Уширение на левом плече распределения носителей образца с 3-мя квантовыми ямами в области толщин от 300 до 400 нм предположительно связано с диффузией электронов в 400-нм слой GaAs. Подобные «хвосты» наблюдались

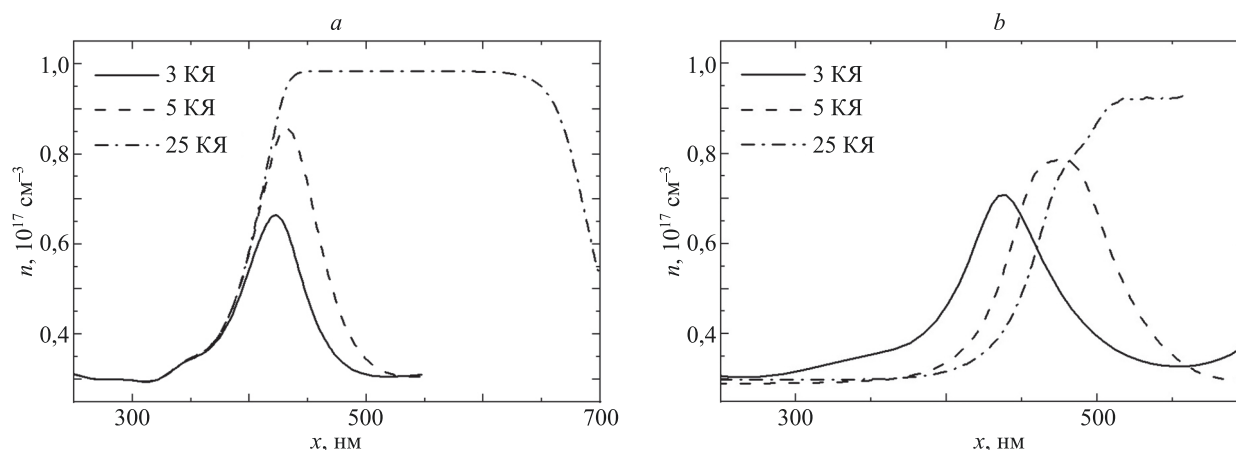


Рисунок. Профили кажущихся концентраций сверхрешеток $\text{Al}_{0.2}\text{Ga}_{0.8}\text{As}/\text{GaAs}$, содержащих 3, 5 и 25 квантовых ям, полученные из смоделированных (а) и экспериментальных (б) вольт-фарадных характеристик.

x — глубина от поверхности образца, КЯ — квантовая яма

Figure. Apparent concentration profiles of $\text{Al}_{0.2}\text{Ga}_{0.8}\text{As}/\text{GaAs}$ superlattices with 3, 5 and 25 periods, obtained from numerically simulated (a) and experimental (b) capacitance-voltage characteristics. x represents the distance from the sample surface.

КЯ — quantum well

на кривых, полученных из моделирования (рисунок, а), для каждого из образцов. Смещение пиков распределения, соответствующих сверхрешеткам с различным количеством периодов, вероятно связано с тем, что в результате травления контактного слоя GaAs итоговые толщины стравленных слоев у разных образцов отличались.

В соответствии с результатами теоретического моделирования на кривых наблюдается рост абсолютных значений концентрации пропорционально увеличению числа квантовых ям в структуре от n равного $7,1 \cdot 10^{16} \text{ см}^{-3}$ (для трех периодов) до $9,2 \cdot 10^{16} \text{ см}^{-3}$ (для 25-ти периодов). Абсолютные значения n для всех образцов соответствуют значениям, предсказанным моделированием, с погрешностью в пределах 10 %. У структур, содержащих 5 и 25 ям, наступает насыщение концентрации носителей на глубине порядка 460 и 500 нм от поверхности соответственно. В областях насыщения распределение электронов имеет плоский участок, который указывает, что электроны делокализованы и распределены по сверхрешетке однородно. Дополнительных особенностей в виде максимумов на профилях концентрации не наблюдается, что говорит об отсутствии флуктуаций толщин эпитаксиальных слоев [8]. Таким образом, метод C-V профилирования может являться дополнительным методом контроля структурного качества сверхрешеток.

Заключение

В работе выполнено исследование распределения свободных носителей заряда в $\text{AlGaAs}/\text{GaAs}$ сверхрешетках с сильно связанными квантовыми ямами с использованием методики вольт-фарадного профилирования. Из экспериментальных профилей кажущейся концентрации носителей выявлено, что свободные электроны эффективно накапливаются в квантовых ямах и их концентрация увеличивается с ростом числа периодов сверхрешетки, что также согласуется с результатами моделирования. Для части образцов наблюдается насыщение концентрации электронов, при этом электроны распределены равномерно по области сверхрешетки. При сравнении теоретических профилей распределения носителей с экспериментальными выяснено, что абсолютные значения концентраций, полученные из численного моделирования, являются завышенными, тем не менее отклонение не превышает 10 %. Таким образом, моделирование вольт-фарадных характеристик является эффективным подходом для предсказания характера распределения носителей в сверхрешетках с узкими барьерами.

Отметим, что вольт-фарадное профилирование является методом, подходящим для оценки не только распределения носителей, но и примеси. При этом следует учесть, что в гетероструктурах со сложным зонным строением смоделированная или измеренная методом C-V профилирования концентрация может не давать точного вида профиля залегания примеси.

Литература

1. Del Alamo J.A. Nanometre-scale electronics with III–V compound semiconductors // *Nature*. 2011. V. 479. N 7373. P. 317–323. <https://doi.org/10.1038/nature10677>
2. Fox M., Ispasoiu R. Quantum wells, superlattices, and band-gap engineering // *Springer Handbook of Electronic and Photonic Materials*. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-48933-9_40
3. Goray L., Pirogov E., Sobolev M., Ilkiv I., Dashkov A., Nikitina E., Ubyivovk E., Gerchikov L., Ipatov A., Vainer Y., Svechnikov M., Yunin P., Chkhalo N., Bouravlev A. Matched characterization of super-multiperiod superlattices // *Journal of Physics D: Applied Physics*. 2020. V. 53. N 45. P. 455103. <https://doi.org/10.1088/1361-6463/aba4d6>
4. Мохов Д.В., Березовская Т.Н., Кузьменков А.Г., Малеев Н.А., Тимошнев С.Н., Устинов В.М. Прецизионная калибровка уровня легирования кремнием эпитаксиальных слоев арсенида галлия // Письма в журнал технической физики. 2017. Т. 43. № 19. С. 87–94. <https://doi.org/10.21883/PJTF.2017.19.45086.16931>
5. Schroder D.K. *Semiconductor Material and Device Characterization*. 3rd ed. Piscataway, Hoboken, NJ: Wiley-IEEE Press, 2015. 800 с.
6. Tschirner B.M., Morier-Genoud F., Martin D., Reinhart F.K. Capacitance-voltage profiling of quantum well structures // *Journal of Applied Physics*. 1996. V. 79. N 9. P. 7005–7013. <https://doi.org/10.1063/1.361466>
7. Bobylev B.A., Kovalevskaja T.E., Marchishin I.V., Ovsyuk V.N. Capacitance-voltage profiling of multiquantum well structures // *Solid-State Electronics*. 1997. V. 41. N 3. P. 481–486. [https://doi.org/10.1016/S0038-1101\(96\)00186-4](https://doi.org/10.1016/S0038-1101(96)00186-4)
8. Chiquito A.J., Pusep Yu.A., Mergulhão S., Galzerani J.C. Carrier confinement in an ultrathin barrier GaAs/AlAs superlattice probed by capacitance-voltage measurements // *Physica E: Low-dimensional Systems and Nanostructures*. 2002. V. 13. N 1. P. 36–42. [https://doi.org/10.1016/S1386-9477\(01\)00222-3](https://doi.org/10.1016/S1386-9477(01)00222-3)
9. Герчиков Л.Г., Дашков А.С., Горай Л.И., Буравлев А.Д. Разработка дизайна сверхмногопериодных излучающих структур терагерцевого диапазона, выращиваемых методом молекулярно-пучковой эпитаксии // Журнал экспериментальной и теоретической физики. 2021. Т. 160. № 2. С. 197–205. <https://doi.org/10.31857/S0044451021080058>
10. Varache R., Leendertz C., Gueunier-Farret M.E., Haschke J., Muñoz D., Korte L. Investigation of selective junctions using a newly developed tunnel current model for solar cell applications // *Solar Energy Materials and Solar Cells*. 2015. V. 141. P. 14–23. <https://doi.org/10.1016/j.solmat.2015.05.014>

Авторы

Василькова Елена Игоревна — инженер, Санкт-Петербургский национальный исследовательский Академический университет им. Ж.И. Алфёрова РАН, Санкт-Петербург, 194021, Российская Федерация, [sc 57433507900](https://orcid.org/0000-0002-0349-7134), <https://orcid.org/0000-0002-0349-7134>, elenvasilkov@gmail.com

Пирогов Евгений Викторович — младший научный сотрудник, Санкт-Петербургский национальный исследовательский Академический университет им. Ж.И. Алфёрова РАН, Санкт-Петербург, 194021, Российская Федерация, [sc 24468711600](https://orcid.org/0000-0001-7186-3768), <https://orcid.org/0000-0001-7186-3768>, zzzavr@gmail.com

Соболев Максим Сергеевич — кандидат физико-математических наук, научный сотрудник, заведующий лабораторией, Санкт-Петербургский национальный исследовательский Академический университет им. Ж.И. Алфёрова РАН, Санкт-Петербург, 194021, Российская Федерация, [sc 57205203666](https://orcid.org/0000-0001-8629-2064), <https://orcid.org/0000-0001-8629-2064>, sobolevsm@gmail.com

Баранов Артем Игоревич — кандидат физико-математических наук, младший научный сотрудник, Санкт-Петербургский национальный исследовательский Академический университет им. Ж.И. Алфёрова РАН, Санкт-Петербург, 194021, Российская Федерация, [sc 57195761820](https://orcid.org/0000-0002-4894-6503), <https://orcid.org/0000-0002-4894-6503>, itiomchik@yandex.ru

Гудовских Александр Сергеевич — доктор технических наук, ведущий научный сотрудник, Санкт-Петербургский национальный исследовательский Академический университет им. Ж.И. Алфёрова РАН, Санкт-Петербург, 194021, Российская Федерация; профессор,

References

1. Del Alamo J.A. Nanometre-scale electronics with III–V compound semiconductors. *Nature*, 2011, vol. 479, no. 7373, pp. 317–323. <https://doi.org/10.1038/nature10677>
2. Fox M., Ispasoiu R. Quantum wells, superlattices, and band-gap engineering. *Springer Handbook of Electronic and Photonic Materials*. Cham, Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-48933-9_40
3. Goray L., Pirogov E., Sobolev M., Ilkiv I., Dashkov A., Nikitina E., Ubyivovk E., Gerchikov L., Ipatov A., Vainer Y., Svechnikov M., Yunin P., Chkhalo N., Bouravlev A. Matched characterization of super-multiperiod superlattices. *Journal of Physics D: Applied Physics*, 2020, vol. 53, no. 45, pp. 455103. <https://doi.org/10.1088/1361-6463/aba4d6>
4. Mokhov D.V., Berezovskaya T.N., Kuzmenkov A.G., Maleev N.A., Timoshnev S.N., Ustinov V.M. Precision calibration of the silicon doping level in gallium arsenide epitaxial layers. *Technical Physics Letters*, 2017, vol. 43, no. 10, pp. 909–911. <https://doi.org/10.1134/S1063785017100091>
5. Schroder D.K. *Semiconductor Material and Device Characterization*. 3rd ed. Piscataway, Hoboken, NJ, Wiley-IEEE Press, 2015, 800 c.
6. Tschirner B.M., Morier-Genoud F., Martin D., Reinhart F.K. Capacitance-voltage profiling of quantum well structures. *Journal of Applied Physics*, 1996, vol. 79, no. 9, pp. 7005–7013. <https://doi.org/10.1063/1.361466>
7. Bobylev B.A., Kovalevskaja T.E., Marchishin I.V., Ovsyuk V.N. Capacitance-voltage profiling of multiquantum well structures. *Solid-State Electronics*, 1997, vol. 41, no. 3, pp. 481–486. [https://doi.org/10.1016/S0038-1101\(96\)00186-4](https://doi.org/10.1016/S0038-1101(96)00186-4)
8. Chiquito A.J., Pusep Yu.A., Mergulhão S., Galzerani J.C. Carrier confinement in an ultrathin barrier GaAs/AlAs superlattice probed by capacitance-voltage measurements. *Physica E: Low-dimensional Systems and Nanostructures*, 2002, vol. 13, no. 1, pp. 36–42. [https://doi.org/10.1016/S1386-9477\(01\)00222-3](https://doi.org/10.1016/S1386-9477(01)00222-3)
9. Gerchikov L.G., Dashkov A.S., Goray L.I., Bouravlev A.D. Development of the design of super-multiperiod structures grown by molecular-beam epitaxy and emitting in the terahertz range. *Journal of Experimental and Theoretical Physics*, 2021, vol. 133, no. 2, pp. 161–168. <https://doi.org/10.1134/S1063776121070037>
10. Varache R., Leendertz C., Gueunier-Farret M.E., Haschke J., Muñoz D., Korte L. Investigation of selective junctions using a newly developed tunnel current model for solar cell applications. *Solar Energy Materials and Solar Cells*, 2015, vol. 141, pp. 14–23. <https://doi.org/10.1016/j.solmat.2015.05.014>

Authors

Elena I. Vasilkova — Engineer, Alferov Saint Petersburg National Research Academic University of the Russian Academy of Sciences, Saint Petersburg, 194021, Russian Federation, [sc 57433507900](https://orcid.org/0000-0002-0349-7134), <https://orcid.org/0000-0002-0349-7134>, elenvasilkov@gmail.com

Evgeny V. Pirogov — Junior Researcher, Alferov Saint Petersburg National Research Academic University of the Russian Academy of Sciences, Saint Petersburg, 194021, Russian Federation, [sc 24468711600](https://orcid.org/0000-0001-7186-3768), <https://orcid.org/0000-0001-7186-3768>, zzzavr@gmail.com

Maxim S. Sobolev — PhD (Physics & Mathematics), Scientific Researcher, Head of Laboratory, Alferov Saint Petersburg National Research Academic University of the Russian Academy of Sciences, Saint Petersburg, 194021, Russian Federation, [sc 57205203666](https://orcid.org/0000-0001-8629-2064), <https://orcid.org/0000-0001-8629-2064>, sobolevsm@gmail.com

Artem I. Baranov — PhD (Physics & Mathematics), Junior Researcher, Alferov Saint Petersburg National Research Academic University of the Russian Academy of Sciences, Saint Petersburg, 194021, Russian Federation, [sc 57195761820](https://orcid.org/0000-0002-4894-6503), <https://orcid.org/0000-0002-4894-6503>, itiomchik@yandex.ru

Alexander S. Gudovskikh — D. Sc., Leading Researcher, Alferov Saint Petersburg National Research Academic University of the Russian Academy of Sciences, Saint Petersburg, 194021, Russian Federation; Professor, Saint Petersburg Electrotechnical University “LETI”, Saint

Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» имени В.И. Ульянова (Ленина), Санкт-Петербург, 197022, Российская Федерация, [sc 6602958574](https://orcid.org/0000-0002-7632-3194), <https://orcid.org/0000-0002-7632-3194>, gudovskikh@spbau.ru

Буравлев Алексей Дмитриевич — доктор физико-математических наук, заведующий лабораторией, Университет при МПА ЕврАзЭС, Санкт-Петербург, 194044, Российская Федерация; профессор, Санкт-Петербургский электротехнический университет «ЛЭТИ», Санкт-Петербург, 197022, Российская Федерация; ведущий научный сотрудник, Институт аналитического приборостроения РАН, Санкт-Петербург, 198095, Российская Федерация, [sc 6603227351](https://orcid.org/0000-0002-7432-2091), <https://orcid.org/0000-0002-7432-2091>, bour@mail.ioffe.ru

Petersburg, 197022, Russian Federation, [sc 6602958574](https://orcid.org/0000-0002-7632-3194), <https://orcid.org/0000-0002-7632-3194>, gudovskikh@spbau.ru

Alexei D. Bouravlev — D. Sc. (Physics & Mathematics), Head of Laboratory, University under the Inter-Parliamentary Assembly of EurAsEC, Saint Petersburg, 194044, Russian Federation; Professor, Saint Petersburg Electrotechnical University “LETI”, Saint Petersburg, 197022, Russian Federation; Leading Researcher, Institute for Analytical Instrumentation of the Russian Academy of Sciences, Saint Petersburg, 198095, Russian Federation, [sc 6603227351](https://orcid.org/0000-0002-7432-2091), <https://orcid.org/0000-0002-7432-2091>, bour@mail.ioffe.ru

Статья поступила в редакцию 18.08.2022
Одобрена после рецензирования 31.10.2022
Принята к печати 27.11.2022

Received 18.08.2022
Approved after reviewing 31.10.2022
Accepted 27.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1098-1103

УДК 538.958

Спектральные и кинетические свойства квантовых точек сульфида серебра во внешнем электрическом поле

Даниил Саузович Дайбаге¹✉, Сергей Александрович Амброзевич²,
Алексей Сергеевич Перепелица³, Иван Александрович Захарчук⁴,
Анна Владимировна Осадченко⁵, Дарья Михайловна Безверхняя⁶,
Антон Игоревич Авраменко⁷, Александр Сергеевич Селюков⁸

^{1,2,4,5,6} Московский государственный технический университет имени Н.Э. Баумана, Москва, 105005,
Российская Федерация

^{1,5,8} Московский политехнический университет, Москва, 107023, Российская Федерация

^{1,2,6,8} Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация

³ Воронежский государственный университет, Воронеж, 394018, Российская Федерация

⁷ Всероссийский институт научной и технической информации Российской академии наук, Москва, 125190,
Российская Федерация

¹ daibagya@mail.ru✉, <https://orcid.org/0000-0003-1944-1546>

² s.ambrozevich@mail.ru, <https://orcid.org/0000-0002-3906-0735>

³ a-perpelitsa@yandex.ru, <https://orcid.org/0000-0002-1264-0107>

⁴ zakharchukia@yandex.ru, <https://orcid.org/0000-0002-1502-6460>

⁵ osadchenkoav@student.bmstu.ru, <https://orcid.org/0000-0001-9556-4885>

⁶ d.bezverkhnyaya@mail.ru, <https://orcid.org/0000-0003-1937-4689>

⁷ anton1905@internet.ru, <https://orcid.org/0000-0002-0374-6428>

⁸ selyukov@lebedev.ru, <https://orcid.org/0000-0002-4007-6291>

Аннотация

Предмет исследования. Исследовано влияние внешнего электрического поля на люминесцентные характеристики наночастиц сульфида серебра, внедренных в пленку на основе оптически пассивной диэлектрической матрицы.

Метод. Исследование люминесцентных характеристик выполнено методами оптической спектроскопии, а также времязрешенной спектроскопии с применением техники времякоррелированного счета одиночных фотонов. Морфология наночастиц изучена при помощи просвечивающей электронной микроскопии. **Основные**

результаты. Показано, что помещение наночастиц сульфида серебра во внешнее электрическое поле приводит к увеличению интенсивности полосы рекомбинационной люминесценции, а также к ускорению процессов релаксации электронного возбуждения. Этот эффект можно объяснить тем, что электрическое поле увеличивает скорость транспорта свободных дырок к электронным ловушкам, которые играют роль центров излучательной рекомбинации. **Практическая значимость.** Показано, что наночастицы сульфида серебра могут быть эффективно использованы в качестве активных слоев органических светоизлучающих диодов, где внешнее поле порядка 500 кВ/см не приведет к ухудшению их рабочих люминесцентных характеристик.

Ключевые слова

полупроводниковые наночастицы, сульфид серебра, рекомбинационная люминесценция, кинетика люминесценции, внешнее электрическое поле

Благодарности

Исследование проведено в рамках проекта РФФИ 20-02-00222 А.

Авторы выражают благодарность декану Физического факультета Воронежского государственного университета О.В. Овчинникову, а также доценту кафедры оптики и спектроскопии Воронежского государственного университета М.С. Смирнову за полезное обсуждение результатов.

Ссылка для цитирования: Дайбаге Д.С., Амброзевич С.А., Перепелица А.С., Захарчук И.А., Осадченко А.В., Безверхняя Д.М., Авраменко А.И., Селюков А.С. Спектральные и кинетические свойства квантовых точек сульфида серебра во внешнем электрическом поле // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1098–1103. doi: 10.17586/2226-1494-2022-22-6-1098-1103

Spectral and kinetic properties of silver sulfide quantum dots in an external electric field

Daniil S. Daibagya^{1✉}, Sergey A. Ambrozevich², Aleksey S. Perepelitsa³,
Ivan A. Zakharchuk⁴, Anna V. Osadchenko⁵, Daria M. Bezverkhnyaya⁶,
Anton I. Avramenko⁷, Alexandr S. Selyukov⁸

^{1,2,4,5,6} Bauman Moscow State Technical University, Moscow, 105005, Russian Federation

^{1,5,8} Moscow Polytechnic University, Moscow, 107023, Russian Federation

^{1,2,6,8} Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation

³ Voronezh State University, Voronezh, 394018, Russian Federation

⁷ Russian Institute for Scientific and Technical Information, Moscow, 125190, Russian Federation

¹ daibagya@mail.ru✉, <https://orcid.org/0000-0003-1944-1546>

² s.ambrozevich@mail.ru, <https://orcid.org/0000-0002-3906-0735>

³ a-perepelitsa@yandex.ru, <https://orcid.org/0000-0002-1264-0107>

⁴ zakharchukia@yandex.ru, <https://orcid.org/0000-0002-1502-6460>

⁵ osadchenkoav@student.bmstu.ru, <https://orcid.org/0000-0001-9556-4885>

⁶ d.bezverkhnyaya@mail.ru, <https://orcid.org/0000-0003-1937-4689>

⁷ anton1905@internet.ru, <https://orcid.org/0000-0002-0374-6428>

⁸ selyukov@lebedev.ru, <https://orcid.org/0000-0002-4007-6291>

Abstract

The effect of an external electric field on the luminescence characteristics of silver sulfide nanoparticles embedded in a film based on an optically passive dielectric matrix has been studied. The luminescence characteristics were studied using methods of optical and time-resolved spectroscopy involving the time-correlated single-photon counting technique. The morphology of the nanoparticles was studied using transmission electron microscopy. It was shown that in an external electric field, an increase in the intensity of the recombination luminescence band is observed for silver sulfide nanoparticles, together with an increase in the electronic relaxation rate. This effect is explained by the fact that the electric field enhances the transport of free holes to electron traps which are radiative recombination centers. The observed effects indicate that silver sulfide nanoparticles can be effectively used as active layers of organic light-emitting diodes, where an external field of the order of 500 kV/cm will not lead to a deterioration in their operating luminescence characteristics.

Keywords

semiconductor nanoparticles, silver sulfide, recombination luminescence, luminescence decays, external electric field

Acknowledgements

The study was carried out within the RFBR project no. 20-02-00222 A.

Authors are grateful to the Dean of the Faculty of Physics of Voronezh State University O.V. Ovchinnikov as well as to Associate Professor of the Department of Optics and Spectroscopy of Voronezh State University M.S. Smirnov for helpful discussions.

For citation: Daibagya D.S., Ambrozevich S.A., Perepelitsa A.S., Zakharchuk I.A., Osadchenko A.V., Bezverkhnyaya D.M., Avramenko A.I., Selyukov A.S. Spectral and kinetic properties of silver sulfide quantum dots in an external electric field, *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1098–1103 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1098-1103

Введение

В настоящее время перспективными источниками света являются органические светоизлучающие диоды [1–5]. Такие устройства экономичны, обладают низкой себестоимостью и могут быть нанесены на гибкие подложки. Больших успехов в создании органических светодиодов добились при использовании активных слоев на основе полупроводниковых нанокристаллов [6]. Такие люминесцентные материалы позволяют создать светоизлучающие устройства с высокой квантовой эффективностью и в меньшей степени подвержены деградации по сравнению с органическими люминофорами. Отметим, что имеющиеся на сегодняшний день успехи достигнуты для видимой области спектра. Применение полупроводниковых нанокристаллов для создания инфракрасных светодиодов исследовано существенно меньше, хотя такие работы оказываются актуальными и имеют множество применений, например в сфере машинного зрения [7]. Наиболее подходящими матери-

алами для использования в качестве люминесцентных слоев в органических светодиодах для инфракрасной области спектра являются наночастицы на основе узкозонных полупроводников, и, в частности, на основе коллоидных сферических квантовых точек сульфида серебра (Ag_2S) [8–11].

Создание органических светодиодов требует всестороннего изучения материала активного излучающего слоя. В этом случае важны не только люминесцентные свойства, но и их зависимость от внешнего электрического поля, возникающего в структуре светоизлучающего устройства [12, 13]. Для наночастиц Ag_2S исследования зависимости люминесцентных свойств от внешнего электрического поля являются единичными и не позволяют получить полную картину происходящих в наночастицах процессов.

В настоящей работе исследовано влияние внешнего электрического поля на люминесцентные характеристики наночастиц Ag_2S , полученных методом фотоиндуцированного синтеза в этиленгликоле.

Техника эксперимента

Синтез наночастиц Ag_2S был проведен с помощью техники фотоиндуцированного роста в этиленгликоле, обеспечивающей пассивацию их поверхности с помощью 2-меркаптопропионовой кислоты [14]. В качестве источника серебра использован нитрат серебра, а источника серы — 2-меркаптопропионовая кислота, взятая в молярном соотношении, в два раза превышающим количество нитрата серебра. Синтез проведен в стеклянной колбе, термостатированной при температуре 25 °C и при постоянном перемешивании магнитной мешалкой. В 30 мл этиленгликоля помещен 2,4 ммоль нитрата серебра, а затем добавлен 4,8 ммоль 2-меркаптопропионовой кислоты также в виде раствора в этиленгликоле. Рост наночастиц осуществлен в темном помещении в течение суток. Уменьшение дисперсии результирующих наночастиц по размерам реализовано с помощью фотолиза при облучении полученного раствора излучением лазера с длиной волны 405 нм и средней мощностью 100 мВт при постоянном перемешивании в течение суток. После проведения процедуры фотолиза наночастицы были выдержаны в темноте в течение трех суток для стабилизации.

Наночастицы были внедрены с помощью инфльтрации полученного раствора в оптически пассивную диэлектрическую органическую матрицу на основе полупроницаемой мембраны из регенерированной целлюлозы толщиной 5 мкм; характерные размеры пор мембраны составили порядка 3 нм.

Контроль размеров синтезированных наночастиц проведен с помощью просвечивающей электронной микроскопии с применением системы Carl Zeiss Libra 120.

Для определения влияния электрического поля на люминесцентные свойства наночастиц Ag_2S изготовлена серия образцов, состоящих из двух стеклянных подложек с нанесенными полупрозрачными электродами на основе оксида индия-олова, покрытыми пленкой полистирола толщиной 5 мкм. Между электродами помещен сэндвич из пленки тефлона толщиной 5 мкм, мембраны с внедренными наночастицами Ag_2S , и еще одной аналогичной пленки тефлона. В результате образцы имеют структуру, соответствующую плоскому конденсатору толщиной 25 мкм, имеющему полупрозрачные обкладки, в объем диэлектрика которого помещены исследуемые наночастицы. Напряжение смещения на полученные конденсаторы подано с высоковольтного источника постоянного напряжения Tesla TV-2, позволяющего задавать напряжение от 0 до 1500 В. При этом в объеме диэлектрика конденсатора было создано постоянное электрическое поле от 0 до 500 кВ/см с учетом диэлектрической проницаемости используемых материалов. Указанные значения поля являются характерными для органических светоизлучающих диодов.

Спектры люминесценции исследованных нанокристаллов измерены с помощью волоконного спектрометра Ocean Optics Maya 2000 Pro, чувствительного в диапазоне 200–1100 нм. Возбуждение было осуществлено с помощью лазера PicoQuant LDH-C-400, излучающего в области 405 нм со средней мощностью

5 мВт. Измерение кинетики люминесценции было проведено методом времякоррелированного счета одиночных фотонов с использованием коррелятора PicoQuant TimeHarp 100. При этом сигнал люминесценции регистрировали с помощью фотоэлектронного умножителя PicoQuant PMA-C 192-N-M, чувствительного в области 250–850 нм. Полоса детектирования шириной порядка 5 нм вырезана с помощью установленного перед фотоэлектронным умножителем монохроматора.

Результаты и их обсуждение

Изображение ансамбля наночастиц Ag_2S , полученное с помощью просвечивающего электронного микроскопа, представлено на рис. 1. Анализ изображения показал, что средний размер наночастиц составляет 2,7 нм, а дисперсия по размерам не превышает 20 %.

Спектры фотолюминесценции ансамбля наночастиц Ag_2S , помещенных в электрическое поле с различной напряженностью, представлены на рис. 2, а. В спектре наблюдается ярко выраженный максимум в области энергий 1,4 эВ, отвечающий рекомбинационной люминесценции [15–17], при этом ширина спектра на полувысоте составила около 320 мэВ. Помещение наночастиц в электрическое поле привело к скачкообразному увеличению интенсивности полосы люминесценции. Отметим, что с ростом поля интенсивность монотонно уменьшается, а форма спектров фотолюминесценции и положение ее максимума существенно не изменяются. Для наглядности на рис. 2, б представлена зависимость интеграла спектра рекомбинационной люминесценции по энергии для наночастиц сульфида серебра от напряженности электрического поля. Увеличение электрического поля от 100 до 500 кВ/см привело к практически линейному уменьшению интенсивности рекомбинационной полосы от 100 % до 92 %. Дальнейшее увеличение напряженности электрического поля не проводилось в силу электрического пробоя образца.

Кинетические зависимости интенсивности фотолюминесценции наночастиц Ag_2S при различных

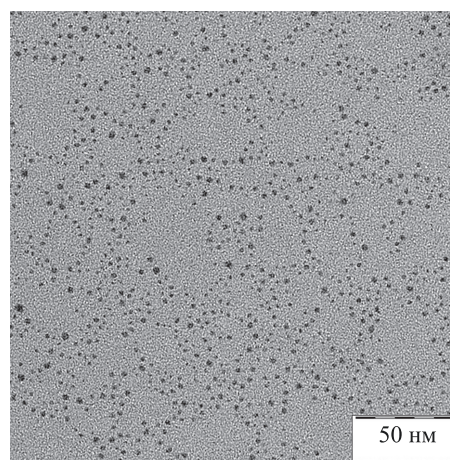


Рис. 1. Изображение наночастиц Ag_2S , полученное с помощью просвечивающего электронного микроскопа

Fig. 1. Transmission electron microscope image of Ag_2S nanoparticles

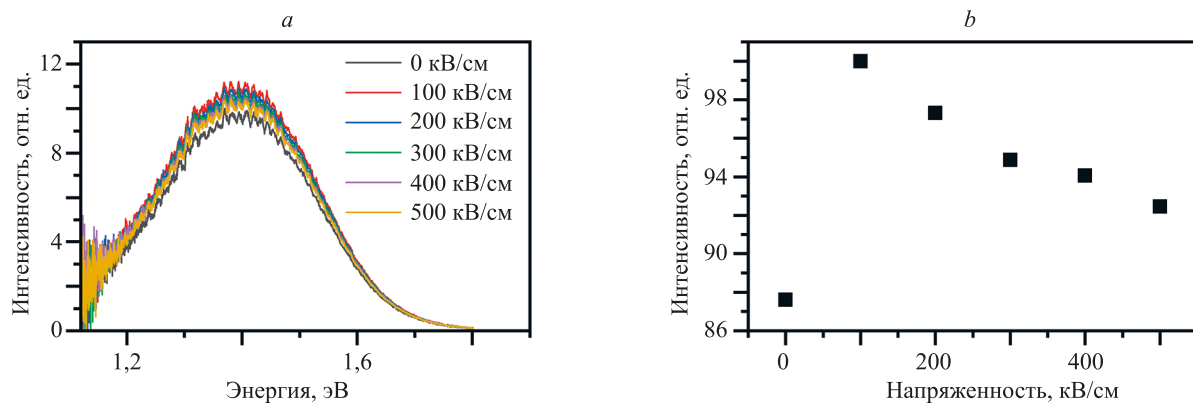


Рис. 2. Спектры фотолюминесценции наночастиц Ag_2S во внешнем электрическом поле (a); зависимость интенсивности фотолюминесценции наночастиц Ag_2S от внешнего электрического поля (b)

Fig. 2. Photoluminescence spectra for Ag_2S nanoparticles in an external electric field (a); dependence of the photoluminescence intensity of Ag_2S nanoparticles vs. the external electric field (b)

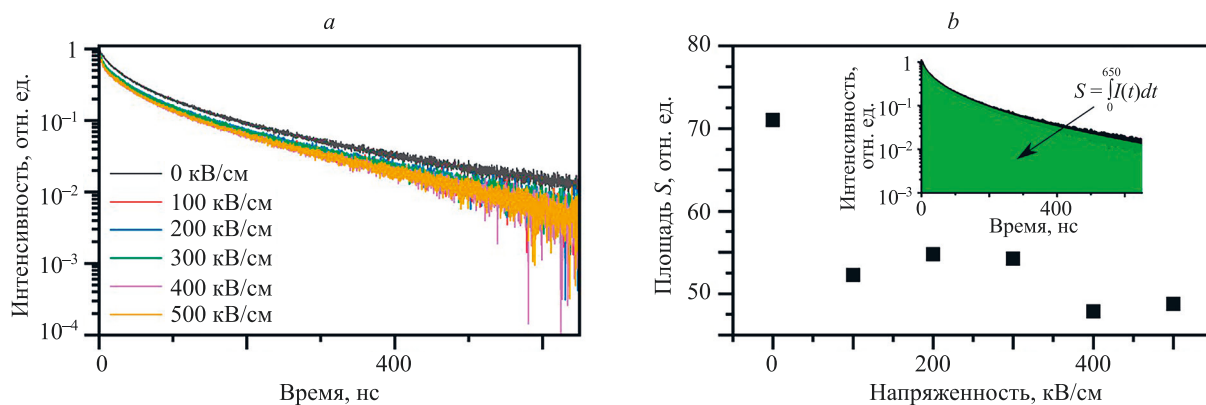


Рис. 3. Кинетические зависимости $I(t)$ фотолюминесценции наночастиц Ag_2S , измеренные при различных значениях внешнего электрического поля (a); зависимость интеграла S кинетических кривых $I(t)$ по времени t в пределах 0–650 нс от величины внешнего электрического поля (b).

На врезке черным цветом изображена зависимость $I(t)$, зеленым — геометрическая интерпретация интеграла S

Fig. 3. Photoluminescence decays $I(t)$ for Ag_2S nanoparticles measured at different values of the external electric field (a); dependence of integrated decays S within range 0–650 ns vs. the external electric field (b). The inset shows the $I(t)$ dependence (black); the green field shows the geometric interpretation of the integral S

величинах внешнего электрического поля представлены на рис. 3, a. Зависимости имеют существенно неэкспоненциальный вид. Видно, что при включении электрического поля происходит ускорение процессов релаксации электронного возбуждения. Для того чтобы установить качественное поведение кинетических зависимостей при изменении величины внешнего электрического поля, построены зависимости интеграла S от кинетических кривых люминесценции $I(t)$ наночастиц по времени t в пределах от 0 до 650 нс при различных значениях напряженности электрического поля (рис. 3, b). В области нулевого электрического поля величина интеграла максимальна; при полях от 100 до 500 кВ/см эта величина уменьшается и изменяется с ростом поля немонотонно.

Наблюдаемые экспериментальные результаты можно объяснить следующим образом. Ускорение процессов релаксации электронного возбуждения и увеличение интенсивности люминесценции определяется тем, что электрическое поле увеличивает скорость транспор-

та свободных дырок к электронным ловушкам, играющим роль центров излучательной рекомбинации [18]. Снижение интегральной интенсивности фотолюминесценции с ростом поля может быть объяснено общей фотодеградацией интенсивности люминесценции наночастиц [19] в результате длительного комбинированного воздействия возбуждающего лазерного излучения и внешнего электрического поля.

Закключение

В работе экспериментально показано, что внешнее электрическое поле может изменять интенсивность люминесценции наночастиц сульфида серебра, а также влиять на скорость происходящих в них процессов релаксации электронного возбуждения. Тем не менее, полученные изменения люминесцентных свойств в полях от 100 до 500 кВ/см не являются существенными для того, чтобы утверждать, что наночастицы сульфида серебра не подходят для создания на их основе орга-

нических светоизлучающих диодов. Таким образом, полученные результаты свидетельствуют о том, что наночастицы сульфида серебра могут быть использованы

для создания органических светоизлучающих диодов нового поколения.

Литература

1. Luo J., Rong X.-F., Ye Y.-Y., Li W.-Z., Wang X.-Q., Wang W. Research progress on triarylmethyl radical-based high-efficiency OLED // *Molecules*. 2022. V. 27. N 5. P. 1632. <https://doi.org/10.3390/molecules27051632>
2. Corrêa Santos D., Vieira Marques M.D.F. Blue light polymeric emitters for the development of OLED devices // *Journal of Materials Science: Materials in Electronics*. 2022. V. 33. N 16. P. 12529–12565. <https://doi.org/10.1007/s10854-022-08333-3>
3. Ващенко А.А., Осадченко А.В., Селюков А.С., Амброзевич С.А., Захарчук И.А., Дайбаге Д.С., Шляхтун О., Володин Н.Ю., Чепцов Д.А., Долотов С.М., Травень В.Ф. Электролюминесценция кумариновых красителей // *Краткие сообщения по физике ФИАН*. 2022. Т. 49. № 3. С. 13–18.
4. Ващенко А.А., Витухновский А.Г., Лебедев В.С., Селюков А.С., Васильев Р.Б., Соколик М.С. Органический светоизлучающий диод на основе плоского слоя полупроводниковых нанопластинок CdSe в качестве эмиттера // *Письма в Журнал экспериментальной и теоретической физики*. 2014. Т. 100. № 2. С. 94–98. <https://doi.org/10.7868/S0370274X14140045>
5. Селюков А.С., Витухновский А.Г., Лебедев В.С., Ващенко А.А., Васильев Р.Б., Соколик М.С. Электролюминесценция коллоидных квазидвумерных полупроводниковых наноструктур CdSe в гибридном светоизлучающем диоде // *Журнал экспериментальной и теоретической физики*. 2015. Т. 147. № 4. С. 687–701. <https://doi.org/10.7868/S0044451015040035>
6. Bauri J., Choudhary R.B., Mandal G. Recent advances in efficient emissive materials-based OLED applications: a review // *Journal of Materials Science*. 2021. V. 56. N 34. P. 18837–18866. <https://doi.org/10.1007/s10853-021-06503-y>
7. Wu P., He T., Zhu H., Wang Y., Li Q., Wang, Z., Fu X., Wang F., Wang P., Shan C., Fan Z., Liao L., Zhou P., Hu W. Next-generation machine vision systems incorporating two-dimensional materials: Progress and perspectives // *InfoMat*. 2022. V. 4. N 1. P. e12275. <https://doi.org/10.1002/inf2.12275>
8. Jiang P., Tian Z.-Q., Zhu C.-N., Zhang Z.-L., Pang D.-W. Emission-tunable near-infrared Ag₂S quantum dots // *Chemistry of Materials*. 2012. V. 24. N 1. P. 3–5. <https://doi.org/10.1021/cm202543m>
9. Grevtseva I.G., Ovchinnikov O.V., Smirnov M.S., Perepelitsa A.S., Chevychelova T.A., Derepko V.N., Osadchenko A.V., Selyukov A.S. The structural and luminescence properties of plexcitonic structures based on Ag₂S/L-Cys quantum dots and Au nanorods // *RSC Advances*. 2022. V. 12. N 11. P. 6525–6532. <https://doi.org/10.1039/D1RA08806H>
10. Lin S., Feng Y., Wen X., Zhang P., Woo S., Shrestha S., Conibeer G., Huang S. Theoretical and experimental investigation of the electronic structure and quantum confinement of wet-chemistry synthesized Ag₂S nanocrystals // *The Journal of Physical Chemistry*. 2015. V. 119. N 1. P. 867–872. <https://doi.org/10.1021/jp511054g>
11. Grevtseva I., Ovchinnikov O., Smirnov M., Perepelitsa A., Chevychelova T., Derepko V., Osadchenko A., Selyukov A. IR luminescence of plexcitonic structures based on Ag₂S/L-Cys quantum dots and Au nanorods // *Optics Express*. 2022. V. 30. N 4. P. 4668–4679. <https://doi.org/10.1364/OE.447200>
12. Bozyigit D., Yarema O., Wood V. Origins of low quantum efficiencies in quantum dot LEDs // *Advanced Functional Materials*. 2013. V. 23. N 24. P. 3024–3029. <https://doi.org/10.1002/adfm.201203191>
13. Vitukhnovsky A.G., Selyukov A.S., Solovey V.R., Vasiliev R.B., Lazareva E.P. Photoluminescence of CdTe colloidal quantum wells in external electric field // *Journal of Luminescence*. 2017. V. 186. P. 194–198. <https://doi.org/10.1016/j.jlumin.2017.02.041>
14. Ovchinnikov O.V., Aslanov S.V., Smirnov M.S., Grevtseva I.G., Perepelitsa A.S. Photostimulated control of luminescence quantum yield for colloidal Ag₂S/2-MPA quantum dots // *RSC Advances*. 2019. V. 9. N 64. P. 37312–37320. <https://doi.org/10.1039/C9RA07047H>
15. Кацаба А.В., Федянин В.В., Амброзевич С.А., Витухновский А.Г., Лобанов А.Н., Селюков А.С., Васильев Р.Б., Саматов И.Г., Брунков П.Н. Характеризация дефектов в коллоидных нанокри-

References

1. Luo J., Rong X.-F., Ye Y.-Y., Li W.-Z., Wang X.-Q., Wang W. Research progress on triarylmethyl radical-based high-efficiency OLED. *Molecules*, 2022, vol. 27, no. 5, pp. 1632. <https://doi.org/10.3390/molecules27051632>
2. Corrêa Santos D., Vieira Marques M.D.F. Blue light polymeric emitters for the development of OLED devices. *Journal of Materials Science: Materials in Electronics*, 2022, vol. 33, no. 16, pp. 12529–12565. <https://doi.org/10.1007/s10854-022-08333-3>
3. Vashchenko A.A., Osadchenko A.V., Selyukov A.S., Ambrozevich S.A., Zakharchuk I.A., Daibagya D.S., Shliakhtun O., Volodin N.Y., Cheptsov D.A., Dolotov S.M., Traven V.F. Electroluminescence of coumarin-based dyes. *Bulletin of the Lebedev Physics Institute*, 2022, vol. 49, no. 3, pp. 74–77. <https://doi.org/10.3103/S106833562203006X>
4. Vashchenko A.A., Vitukhnovskii A.G., Lebedev V.S., Selyukov A.S., Vasiliev R.B., Sokolikova M.S. Organic light emitting diode with an emitter based on a planar layer of cdse semiconductor nanoplatelets. *JETP Letters*, 2014, vol. 100, no. 2, pp. 86–90. <https://doi.org/10.1134/S0021364014140124>
5. Selyukov A.S., Vitukhnovskii A.G., Lebedev V.S., Vashchenko A.A., Vasiliev R.B., Sokolikova M.S. Electroluminescence of colloidal quasi-two-dimensional semiconducting CdSe nanostructures in a hybrid light-emitting diode. *Journal of Experimental and Theoretical Physics*, 2015, vol. 120, no. 4, pp. 595–606. <https://doi.org/10.1134/S1063776115040238>
6. Bauri J., Choudhary R.B., Mandal G. Recent advances in efficient emissive materials-based OLED applications: a review. *Journal of Materials Science*, 2021, vol. 56, no. 34, pp. 18837–18866. <https://doi.org/10.1007/s10853-021-06503-y>
7. Wu P., He T., Zhu H., Wang Y., Li Q., Wang, Z., Fu X., Wang F., Wang P., Shan C., Fan Z., Liao L., Zhou P., Hu W. Next-generation machine vision systems incorporating two-dimensional materials: Progress and perspectives. *InfoMat*, 2022, vol. 4, no. 1, pp. e12275. <https://doi.org/10.1002/inf2.12275>
8. Jiang P., Tian Z.-Q., Zhu C.-N., Zhang Z.-L., Pang D.-W. Emission-tunable near-infrared Ag₂S quantum dots. *Chemistry of Materials*, 2012, vol. 24, no. 1, pp. 3–5. <https://doi.org/10.1021/cm202543m>
9. Grevtseva I.G., Ovchinnikov O.V., Smirnov M.S., Perepelitsa A.S., Chevychelova T.A., Derepko V.N., Osadchenko A.V., Selyukov A.S. The structural and luminescence properties of plexcitonic structures based on Ag₂S/L-Cys quantum dots and Au nanorods. *RSC Advances*, 2022, vol. 12, no. 11, pp. 6525–6532. <https://doi.org/10.1039/D1RA08806H>
10. Lin S., Feng Y., Wen X., Zhang P., Woo S., Shrestha S., Conibeer G., Huang S. Theoretical and experimental investigation of the electronic structure and quantum confinement of wet-chemistry synthesized Ag₂S nanocrystals. *The Journal of Physical Chemistry*, 2015, vol. 119, no. 1, pp. 867–872. <https://doi.org/10.1021/jp511054g>
11. Grevtseva I., Ovchinnikov O., Smirnov M., Perepelitsa A., Chevychelova T., Derepko V., Osadchenko A., Selyukov A. IR luminescence of plexcitonic structures based on Ag₂S/L-Cys quantum dots and Au nanorods. *Optics Express*, 2022, vol. 30, no. 4, pp. 4668–4679. <https://doi.org/10.1364/OE.447200>
12. Bozyigit D., Yarema O., Wood V. Origins of low quantum efficiencies in quantum dot LEDs. *Advanced Functional Materials*, 2013, vol. 23, no. 24, pp. 3024–3029. <https://doi.org/10.1002/adfm.201203191>
13. Vitukhnovsky A.G., Selyukov A.S., Solovey V.R., Vasiliev R.B., Lazareva E.P. Photoluminescence of CdTe colloidal quantum wells in external electric field. *Journal of Luminescence*, 2017, vol. 186, pp. 194–198. <https://doi.org/10.1016/j.jlumin.2017.02.041>
14. Ovchinnikov O.V., Aslanov S.V., Smirnov M.S., Grevtseva I.G., Perepelitsa A.S. Photostimulated control of luminescence quantum yield for colloidal Ag₂S/2-MPA quantum dots. *RSC Advances*, 2019, vol. 9, no. 64, pp. 37312–37320. <https://doi.org/10.1039/C9RA07047H>
15. Katsaba A.V., Fedyanin V.V., Ambrozevich S.A., Vitukhnovsky A.G., Lobanov A.N., Selyukov A.S., Vasiliev R.B., Samatov I.G.,

- сталлах CdSe модифицированным методом термостимулированной люминесценции // Физика и техника полупроводников. 2013. Т. 47. № 10. С. 1339–1343.
16. Ovchinnikov O.V., Grevtseva I.G., Smirnov M.S., Kondratenko T.S. Reverse photodegradation of infrared luminescence of colloidal Ag₂S quantum dots // *Journal of Luminescence*. 2019. V. 207. P. 626–632. <https://doi.org/10.1016/j.jlumin.2018.12.019>
 17. Derepko V.N., Ovchinnikov O.V., Smirnov M.S., Grevtseva I.G., Kondratenko T.S., Selyukov A.S., Turishchev S.Y. Plasmon-exciton nanostructures, based on CdS quantum dots with exciton and trap state luminescence // *Journal of Luminescence*. 2022. V. 248. P. 118874. <https://doi.org/10.1016/j.jlumin.2022.118874>
 18. Smirnov M.S., Ovchinnikov O.V. IR luminescence mechanism in colloidal Ag₂S quantum dots // *Journal of Luminescence*. 2020. V. 227. P. 117526. <https://doi.org/10.1016/j.jlumin.2020.117526>
 19. Смирнов М.С., Овчинников О.В., Гревцева И.Г., Звягин А.И., Перепелица А.С., Ганеев Р.А. Фотоиндуцированная деградация оптических свойств коллоидных квантовых точек Ag₂S и CdS, пассивированных тиогликолевой кислотой // *Оптика и спектроскопия*. 2018. Т. 124. № 5. С. 648–653. <https://doi.org/10.21883/OS.2018.05.45946.312-17>
 - Brunkov P.N. Characterization of defects in colloidal cdse nanocrystals by the modified thermostimulated luminescence technique. *Semiconductors*, 2013, vol. 47, no. 10, pp. 1328–1332. <https://doi.org/10.1134/S1063782613100138>
 16. Ovchinnikov O.V., Grevtseva I.G., Smirnov M.S., Kondratenko T.S. Reverse photodegradation of infrared luminescence of colloidal Ag₂S quantum dots. *Journal of Luminescence*, 2019, vol. 207, pp. 626–632. <https://doi.org/10.1016/j.jlumin.2018.12.019>
 17. Derepko V.N., Ovchinnikov O.V., Smirnov M.S., Grevtseva I.G., Kondratenko T.S., Selyukov A.S., Turishchev S.Y. Plasmon-exciton nanostructures, based on CdS quantum dots with exciton and trap state luminescence. *Journal of Luminescence*, 2022, vol. 248, pp. 118874. <https://doi.org/10.1016/j.jlumin.2022.118874>
 18. Smirnov M.S., Ovchinnikov O.V. IR luminescence mechanism in colloidal Ag₂S quantum dots. *Journal of Luminescence*, 2020, vol. 227, pp. 117526. <https://doi.org/10.1016/j.jlumin.2020.117526>
 19. Smirnov M.S., Ovchinnikov O.V., Grevtseva I.G., Zvyagin A.I., Perepelitsa A.S., Ganeev R.A. Photoinduced degradation of the optical properties of colloidal Ag₂S and CdS quantum dots passivated by thioglycolic acid. *Optics and Spectroscopy*, 2018, vol. 124, no. 5, pp. 681–686. <https://doi.org/10.1134/S0030400X18050211>

Авторы

Дайбаге Даниил Саузович — студент, Московский государственный технический университет имени Н.Э. Баумана, Москва, 105005, Российская Федерация; ассистент, Московский политехнический университет, Москва, 107023, Российская Федерация; младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, [sc 57673090900](https://orcid.org/0000-0003-1944-1546), <https://orcid.org/0000-0003-1944-1546>, daibagya@mail.ru

Амброзович Сергей Александрович — кандидат физико-математических наук, старший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация; доцент, Московский государственный технический университет имени Н.Э. Баумана, Москва, 105005, Российская Федерация, [sc 12789274600](https://orcid.org/0000-0002-3906-0735), <https://orcid.org/0000-0002-3906-0735>, s.ambrozevich@mail.ru

Перепелица Алексей Сергеевич — кандидат физико-математических наук, старший преподаватель, Воронежский государственный университет, Воронеж, 394018, Российская Федерация, [sc 55793662400](https://orcid.org/0000-0002-1264-0107), <https://orcid.org/0000-0002-1264-0107>, a-perepelitsa@yandex.ru

Захарчук Иван Александрович — студент, Московский государственный технический университет имени Н.Э. Баумана, Москва, 105005, Российская Федерация, [sc 57672815700](https://orcid.org/0000-0002-1502-6460), <https://orcid.org/0000-0002-1502-6460>, zakharchukia@yandex.ru

Осадченко Анна Владимировна — студент, Московский государственный технический университет имени Н.Э. Баумана, Москва, 105005, Российская Федерация; ассистент, Московский политехнический университет, Москва, 107023, Российская Федерация, [sc 57439684100](https://orcid.org/0000-0001-9556-4885), <https://orcid.org/0000-0001-9556-4885>, osadchenkoav@student.bmstu.ru

Безверхняя Дарья Михайловна — студент, Московский государственный технический университет имени Н.Э. Баумана, Москва, 105005, Российская Федерация; лаборант, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, <https://orcid.org/0000-0003-1937-4689>, d.bezverkhnaya@mail.ru

Авраменко Антон Игоревич — научный редактор, Всероссийский институт научной и технической информации Российской академии наук, Москва, 125190, Российская Федерация, <https://orcid.org/0000-0002-0374-6428>, anton1905@internet.ru

Селюков Александр Сергеевич — кандидат физико-математических наук, младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация; старший преподаватель, Московский политехнический университет, Москва, 107023, Российская Федерация, [sc 55787344500](https://orcid.org/0000-0002-4007-6291), <https://orcid.org/0000-0002-4007-6291>, selyukov@lebedev.ru

Authors

Daniil S. Daibagya — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation; Assistant, Moscow Polytechnic University, Moscow, 107023, Russian Federation; Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, [sc 57673090900](https://orcid.org/0000-0003-1944-1546), <https://orcid.org/0000-0003-1944-1546>, daibagya@mail.ru

Sergey A. Ambrozevich — PhD (Physics & Mathematics), Senior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation; Associate Professor, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation, [sc 12789274600](https://orcid.org/0000-0002-3906-0735), <https://orcid.org/0000-0002-3906-0735>, s.ambrozevich@mail.ru

Aleksey S. Perepelitsa — PhD (Physics & Mathematics), Senior Lecturer, Voronezh State University, Voronezh, 394018, Russian Federation, [sc 55793662400](https://orcid.org/0000-0002-1264-0107), <https://orcid.org/0000-0002-1264-0107>, a-perepelitsa@yandex.ru

Ivan A. Zakharchuk — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation, [sc 57672815700](https://orcid.org/0000-0002-1502-6460), <https://orcid.org/0000-0002-1502-6460>, zakharchukia@yandex.ru

Anna V. Osadchenko — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation; Assistant, Moscow Polytechnic University, Moscow, 107023, Russian Federation, [sc 57439684100](https://orcid.org/0000-0001-9556-4885), <https://orcid.org/0000-0001-9556-4885>, osadchenkoav@student.bmstu.ru

Daria M. Bezverkhnaya — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation; Laboratory Assistant, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, <https://orcid.org/0000-0003-1937-4689>, d.bezverkhnaya@mail.ru

Anton I. Avramenko — Science Editor, Russian Institute for Scientific and Technical Information, Moscow, 125190, Russian Federation, <https://orcid.org/0000-0002-0374-6428>, anton1905@internet.ru

Alexandr S. Selyukov — PhD (Physics & Mathematics), Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation; Senior Lecturer, Moscow Polytechnic University, Moscow, 107023, Russian Federation, [sc 55787344500](https://orcid.org/0000-0002-4007-6291), <https://orcid.org/0000-0002-4007-6291>, selyukov@lebedev.ru

Статья поступила в редакцию 18.08.2022
Одобрена после рецензирования 06.10.2022
Принята к печати 20.11.2022

Received 18.08.2022
Approved after reviewing 06.10.2022
Accepted 20.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1104-1111

УДК 543.428

Влияние наноразмерных горизонтальных неоднородностей на послойный анализ поверхности методом рентгеновской фотоэлектронной спектроскопии

Денис Сергеевич Лукьянцев¹✉, Александр Владимирович Лубенченко²,
Дмитрий Александрович Иванов³, Ольга Игоревна Лубенченко⁴,
Александр Сергеевич Федотов⁵

^{1,2,3,4,5} Национальный исследовательский университет «МЭИ», Москва, 111250, Российская Федерация

¹ stavden1996@mail.ru✉, <https://orcid.org/0000-0003-4350-8034>

² LubenchenkoAV@mpei.ru, <https://orcid.org/0000-0002-5998-5495>

³ iv_prof@mail.ru, <https://orcid.org/0000-0003-3760-0598>

⁴ IvanovaOI@mpei.ru, <https://orcid.org/0000-0002-6317-2841>

⁵ courage_man@mail.ru, <https://orcid.org/0000-0002-7641-3173>

Аннотация

Предмет исследования. Количественный послойный анализ поверхностных слоев тонких пленок проводят с помощью метода рентгеновской фотоэлектронной спектроскопии по расчетной модели, которая предполагает, что поверхностные слои образца однородны и плоскопараллельны. Однако практически любая поверхность ультратонкой пленки неровная. Исследование такой поверхности по модели плоскопараллельных слоев может привести к некорректным результатам. В настоящей работе для анализа ультратонкой пленки предложено использовать модель неоднородного стохастического наноструктурированного поверхностного слоя. **Метод.** Поверхностные стохастические наноструктурированные неоднородности описаны функцией нормального распределения Гаусса и определяются тремя параметрами: дисперсия (разброс толщин по слою), средняя и максимальная толщины поверхностного слоя. Впервые определен вид рентгеновского фотоэлектронного спектра неоднородной стохастической наноструктурированной поверхности с помощью функций рождения и прохождения фотоэлектронов через поверхностный слой. Разработанная модель основана на следующих предположениях: фотоэлектроны рождаются в веществе и движутся прямо-вперед (приближение Straight Line Approximation) по нормали к поверхности, плотность их потока ослабевает в слое по закону Бугера–Ламберта; фотоэлектроны с различными энергиями теряют энергию по-разному; потери энергии фотоэлектронов в объеме и на поверхности слоя различны. **Основные результаты.** Выполнено моделирование рентгеновских фотоэлектронных спектров окисленной металлической пленки по следующим моделям: однородные плоскопараллельные слои, островковый и неоднородный стохастический наноструктурированные поверхностные слои. Определены границы применимости моделей однородных плоскопараллельных слоев и простого периодически островкового наноструктурированного поверхностного слоя для анализа неоднородной стохастической наноструктурированной поверхности. При некоторых параметрах неоднородного стохастического поверхностного слоя модель однородных плоскопараллельных слоев показывает удовлетворительные результаты послойного анализа. Показано, что модель простого периодически наноструктурированного островкового слоя дает неадекватные результаты при анализе неоднородной стохастической поверхности. **Практическая значимость.** Проведенные в работе исследования показали, что для более точного послойного анализа неоднородной ультратонкой пленки необходимо учитывать неоднородность реальной поверхности, в противном случае полученные результаты будут не соответствовать фактическим.

Ключевые слова

послойный фазовый анализ, горизонтальные неоднородности, шероховатость поверхности, наноразмерные пленки, РФЭС

Ссылка для цитирования: Лукьянцев Д.С., Лубенченко А.В., Иванов Д.А., Лубенченко О.И., Федотов А.С. Влияние наноразмерных горизонтальных неоднородностей на послойный анализ поверхности методом рентгеновской фотоэлектронной спектроскопии // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1104–1111. doi: 10.17586/2226-1494-2022-22-6-1104-1111

© Лукьянцев Д.С., Лубенченко А.В., Иванов Д.А., Лубенченко О.И., Федотов А.С., 2022

Influence of nano-sized horizontal inhomogeneities on surface profiling by means of XPS

Denis S. Lukyantsev¹✉, Alexander V. Lubenchenko², Dmitriy A. Ivanov³,
Olga I. Lubenchenko⁴, Alexandr S. Fedotov⁵

^{1,2,3,4,5} National Research University "Moscow Power Engineering Institute", Moscow, 111250, Russian Federation

¹ stavden1996@mail.ru✉, <https://orcid.org/0000-0003-4350-8034>

² LubenchenkoAV@mpei.ru, <https://orcid.org/0000-0002-5998-5495>

³ iv_prof@mail.ru, <https://orcid.org/0000-0003-3760-0598>

⁴ IvanovaOI@mpei.ru, <https://orcid.org/0000-0002-6317-2841>

⁵ courage_man@mail.ru, <https://orcid.org/0000-0002-7641-3173>

Abstract

Quantitative analysis of thin films surface is performed by means of X-ray electron spectroscopy (XPS) according to a calculation model assuming surface layers of the target to be homogeneous and parallel. However, almost every surface of an ultra-thin film is rough. A study of such surface using the plane-parallel layer model will lead to incorrect results. This work proposes to use the model of inhomogeneous stochastic nano-structured surface layer for ultra-thin film profiling. Surface stochastic nano-structured inhomogeneities are described by the normal Gauss distribution function. To determine these inhomogeneities, three parameters are specified: dispersion (spread of thicknesses by the layer), mean and maximal thickness of the surface layer. For the first time, the type of X-ray photoelectron spectrum of an inhomogeneous stochastic nano-structured surface is found that is determined by functions of photoelectron production and transmission through that surface layer. The designed model is based on the following assumptions: photoelectrons are produced in substance and travel straight-forward (Straight Line Approximation) along the surface, photoelectron flux density decreases in the layer according to the Bouguer–Lambert law, photoelectrons of different energies lose energy differently, photoelectron energy losses in bulk and on surface differ. Modeling of X-ray photoelectron spectra of an oxidized metal film is performed using different models: homogeneous plane-parallel layers, an island nano-structured surface layer and an inhomogeneous stochastic nano-structured surface layer. Ranges of applicability of plane-parallel layer models and simple periodical nano-structured island surface layer for inhomogeneous stochastic nano-structured surface profiling are determined. The model of homogeneous plane-parallel layers shows satisfactory profiling results by some values of parameters of an inhomogeneous stochastic surface layer. It is shown that the model of a simple periodically nano-structured island layer leads to inadequate results by profiling of an inhomogeneous stochastic surface. The investigation shows that for more accurate profiling of an inhomogeneous ultra-thin film, it is necessary to consider inhomogeneity of a real surface, otherwise the calculated results would not match the true profile.

Keywords

phase profiling, horizontal inhomogeneities, surface roughness, nano-sized films, XPS

For citation: Lukyantsev D.S., Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Fedotov A.S. Influence of nano-sized horizontal inhomogeneities on surface profiling by means of XPS. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1104–1111 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1104-1111

Введение

Поверхность ультратонкой пленки можно исследовать неразрушающим методом рентгеновской фотоэлектронной спектроскопии (РФЭС). При послойном фазовом анализе таких пленок толщины слоев обычно определяют по однослойной [1] или многослойной [2] моделям, основанным на допущении, что все слои образца являются однородными и плоскопараллельными. Послойный анализ горизонтально ровных ультратонких пленок, проведенный по многослойной модели, дает хорошие результаты [3–5]. Если поверхность представлена в виде одинаковых островковых наноструктур, то для анализа можно применить расчетную модель [6]. В работах [7–10] получены модельные рентгеновские фотоэлектронные спектры для различных двумерных неровностей на поверхности, в том числе для плотно укомплектованных одинаковых сфер. Модель одинаковых сфер показала хорошие результаты при послойном фазовом анализе наноразмерных порошкообразных химических катализаторов [11].

В работе [12] проведен послойный фазовый анализ рифленых поверхностей окисленного кремния и алюминия с помощью модели, в которой шероховатость поверхности смоделирована набором малых много-

гранников. В модель помещена гистограмма угла наклона многогранников, извлеченная из карты высот поверхности, полученной с помощью атомно-силовой микроскопии, и вычислены интенсивности для каждого отдельного элемента поверхности. Данная модель при подавляющем большинстве граней с более высоким углом наклона показала более точные результаты по сравнению с моделью плоскопараллельных слоев. Похожий подход использован в работе [13], но по более упрощенному алгоритму, который привел к большей неточности результатов.

В [14] показано, что реальная поверхность ультратонкой пленки неровная и может иметь неоднородности различных размеров и формы. В [15] исследованы методом РФЭС поверхности различной степени неоднородности и полученные результаты показали, что вид поверхности образца влияет на интенсивность и форму рентгеновского фотоэлектронного сигнала. Таким образом, можно сделать вывод, что применение модели однородных плоскопараллельных слоев для анализа неоднородной поверхности ультратонкой пленки приводит к неадекватным результатам.

В настоящей работе для анализа ультратонкой пленки предложено использовать модель неоднородного стохастического наноструктурного поверхностного

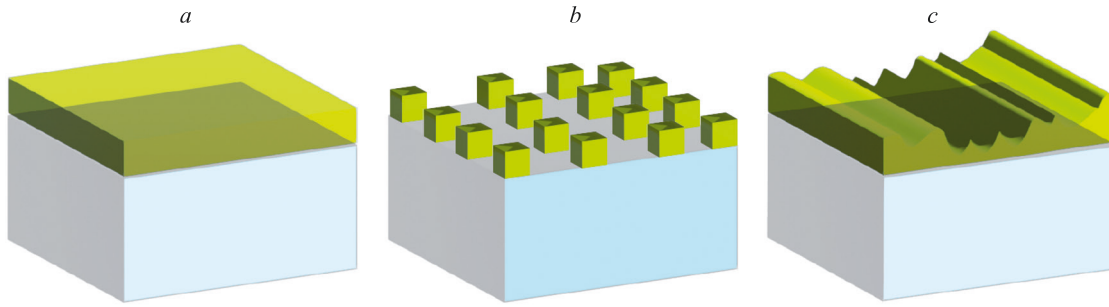


Рис. 1. Модели поверхностных слоев: однородного плоскопараллельного (модель 1) (a); островкового наноструктурированного (модель 2) (b); стохастического наноструктурированного (модель 3) (c)

Fig. 1. Surface layer models: homogeneous plane-parallel (model 1) (a); islet nano-structured (model 2) (b); stochastic nano-structured (c)

слоя, которая, предположительно, лучше подходит для описания случайной шероховатости поверхности. В работе рассмотрен нормальный вылет фотоэлектронов из поверхности образца. Для определения вида рентгеновского фотоэлектронного спектра от стохастического поверхностного слоя решена задача прохождения фотоэлектронов сквозь этот слой.

Рассмотрим три модели поверхностных слоев (рис. 1), и по результатам послойного фазового анализа ультратонкой пленки получим толщины слоев для одной и той же поверхности.

Теоретическое описание фотоэлектронной эмиссии

Основные предположения и допущения решения.

Предположим, что при определении интенсивности потока фотоэлектронов, проходящих сквозь поверхностный слой вещества, фотоэлектроны могут рождаться в любой точке поверхностного слоя и подложки. После рождения фотоэлектроны в веществе движутся прямо-вперед (приближение Straight Line Approximation) по нормали к поверхности. При движении плотность потока фотоэлектронов ослабевает по экспоненциальному закону Бугера–Ламберта. Причем фотоэлектроны, рожденные с разной кинетической энергией, теряют энергию по-разному. Это учитывается модельным дифференциальным сечением неупругого рассеяния [2]. При движении к поверхности фотоэлектроны проходят границу раздела двух сред (например, подложка – поверхностный слой), на которой наблюдается ослабление потока фотоэлектронов. Для учета данной особенности введем разделение плазменных потерь энергии на объемные и поверхностные [2].

Функция фотоэлектронной эмиссии. Интенсивность потока фотоэлектронов, рожденных в подложке-полубесконечном слое, определим функцией фотоэлектронной эмиссии [16]:

$$Q_{\text{inf}} = n_0 \sigma_{\text{ion}} \lambda_{\text{in}} (\mathbf{E} - \mathbf{A})^{-1},$$

где n_0 — атомная концентрация частиц в веществе; σ_{ion} — дифференциальное сечение ионизации; λ_{in} — длина свободного пробега между неупругими актами соударения; $\mathbf{E}$ — единичная матрица; $\mathbf{A} = \text{Toeplitz}(\mathbf{X}_{\text{in}})$, $\mathbf{X}_{\text{in}}$ — дифференциальное сечение неупругого рассеяния [16].

Для модели 1 (рис. 2, a) фотоэлектроны могут рождаться в подложке, проходить сквозь однородный плоскопараллельный слой (Q_{12}) и рождаться в поверхностном слое (Q_2). В таком случае получим суммарную интенсивность потока фотоэлектронов:

$$Q_{\Sigma} = Q_{12} + Q_2, \quad (1)$$

$$Q_{12} = Q_{\text{inf}} \times T_{\text{in}}(d) \text{ и } Q_2 = Q_{\text{inf}} \times (\mathbf{E} - T_{\text{in}}(d)),$$

где $T_{\text{in}}(d)$ — функция пропускания фотоэлектронов сквозь однородный плоскопараллельный слой вещества толщиной d [17].

Если поверхность представлена в виде островковых наноструктур (модель 2) (рис. 2, b), то функция фотоэлектронной эмиссии будет определена фотоэлектронами, рожденными в подложке и вылетевшими сквозь островок поверхностного слоя (Q_{12}) или минуя этот слой (Q_{inf}), рожденными в островке поверхностного слоя (Q_2). Интенсивность потока фотоэлектронов для островковой модели имеет вид:

$$Q_{\Sigma} = (1 - \alpha) \times Q_{\text{inf}} + \alpha \times (Q_{12} + Q_2), \quad (2)$$

где $\alpha = \frac{S_0}{S_1}$ — степень заполнения поверхностного слоя; S_0 — суммарная площадь островков; S_1 — площадь всей поверхности.

Функцию пропускания для моделей 1 и 2 определим выражением [16]:

$$T_{\text{in}}(d) = \int_0^d e^{-(\mathbf{E} - \mathbf{X}_{\text{in}}) \frac{z}{\lambda_{\text{in}}}} dz = \lambda_{\text{in}} (\mathbf{E} - \mathbf{A})^{-1} e^{-\frac{d}{\lambda_{\text{in}}}} e^{\mathbf{A} \frac{d}{\lambda_{\text{in}}}}.$$

В случае с моделью со стохастическим наноструктурным поверхностным слоем (модель 3) (рис. 2, c) интенсивность потока фотоэлектронов определяется выражением:

$$Q_{\Sigma} = Q_{12} + Q_2,$$

$$Q_{12} = Q_{\text{inf}} \times T_{\text{in}}(d_{\text{max}}, \sigma, d_{\text{cp}})$$

$$\text{и } Q_2 = Q_{\text{inf}} \times (\mathbf{E} - T_{\text{in}}(d_{\text{max}}, \sigma, d_{\text{cp}})), \quad (3)$$

где $T_{\text{in}}(d_{\text{max}}, \sigma, d_{\text{cp}})$ — функция пропускания фотоэлектронов сквозь стохастический слой вещества, которая

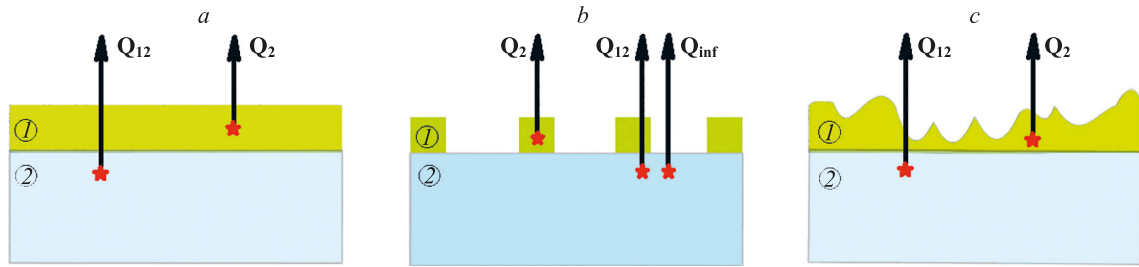


Рис. 2. Схематичное прохождение электронов через поверхностные слои для модели 1 (а), модели 2 (б) и модели 3 (с).

1 — подложка; 2 — поверхностный слой

Fig. 2. Schematic passage of electrons through surface layers for model 1 (a), model 2 (b) and model 3 (c).

1 — substrate; 2 — surface layer

определяется средней (d_{cp}) и максимальной (d_{max}) толщинами слоя, а также дисперсией толщины слоя (σ).

Функция пропускания фотоэлектронов сквозь стохастический слой

Для стохастической наноструктурированной модели 3 (рис. 2, с) случайный характер неоднородностей поверхностного слоя задается функцией плотности вероятности нормального распределения Гаусса:

$$P(d) = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(d-d_{cp})^2}{(2\sigma)^2}},$$

которая определяется средней толщиной слоя (d_{cp}), дисперсией — разбросом высот по поверхности (σ) и случайной величиной (d). Данную функцию необходимо ограничить в пределах толщины поверхностного слоя (рис. 3). Для этого введем условие нормировки:

$$\int_0^{d_{max}} P(z) dz = 1,$$

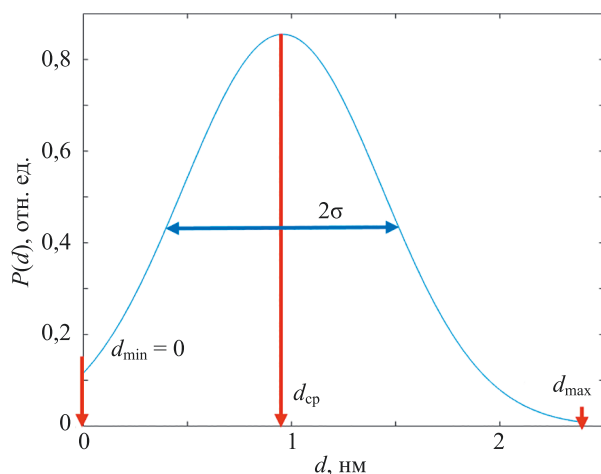


Рис. 3. Вид функции плотности вероятности нормального распределения Гаусса, нормированной в пределах от 0 до d_{max}

Fig. 3. Normal distribution normalized within the limits from 0 to d_{max}

где d_{max} ограничивается максимальной толщиной слоя; z — случайная величина.

Для модели 3 (рис. 2, с) функцию пропускания фотоэлектронов сквозь стохастический поверхностный слой определим выражением:

$$T_{in}(d_{max}, \sigma, d_{cp}) = \int_0^{d_{max}} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(z-d_{cp})^2}{(2\sigma)^2}} \times e^{-(E-X_{in})\frac{z}{\lambda_{in}}} dz. \quad (4)$$

Для получения аналитического решения выражения (4) удобно ввести безразмерную толщину слоя $\tau = \frac{d}{\lambda_{in}}$. Получим аналитическое выражение функции пропускания фотоэлектронов сквозь стохастический поверхностный слой:

$$T_{in}(\tau_{max}, \sigma, \tau_{cp}) = \frac{G_1}{\sqrt{G_2}} e^{mG_2} \times \left[\sum_{k=1}^{50} L(\sqrt{G_2}(\mathbf{D} + \tau_{max}\mathbf{E}))^{2k-1} - \sum_{k=1}^{50} L(\sqrt{G_2}\mathbf{D})^{2k-1} \right], \quad (5)$$

$$G_1 = \frac{\lambda_{in}}{\sigma\sqrt{2\pi}}, \quad G_2 = \frac{1}{\left(2\frac{\sigma}{\lambda_{in}}\right)^2}, \quad \mathbf{D} = \frac{\mathbf{E} - \mathbf{A}}{2G_2} - \tau_{cp}\mathbf{E},$$

$$\mathbf{m} = \mathbf{D}^2 - (\tau_{cp}\mathbf{E})^2, \quad L = \frac{(-1)^{k+1}}{(2k-1)(k-1)!}.$$

Выражение (5) содержит ряды, сумму которых достаточно ограничить 50 членами, в таком случае ряд сходится и погрешность расчета составляет менее 0,01 %

Моделирование рентгеновских фотоэлектронных спектров. Моделирование рентгеновских фотоэлектронных спектров выполнено для трех моделей мишеней (рис. 2) по выражениям (1)–(3). В качестве материала подложки выбран ниобий, поверхностный слой для всех моделей состоит из пентаоксида ниобия Nb_2O_5 . Первоначально для определенных параметров функции нормального распределения Гаусса построен спектр модели 3 со стохастическим наноструктурным слоем (рис. 2, с), который принят за действительный. В результате процедуры фитинга по алгоритму

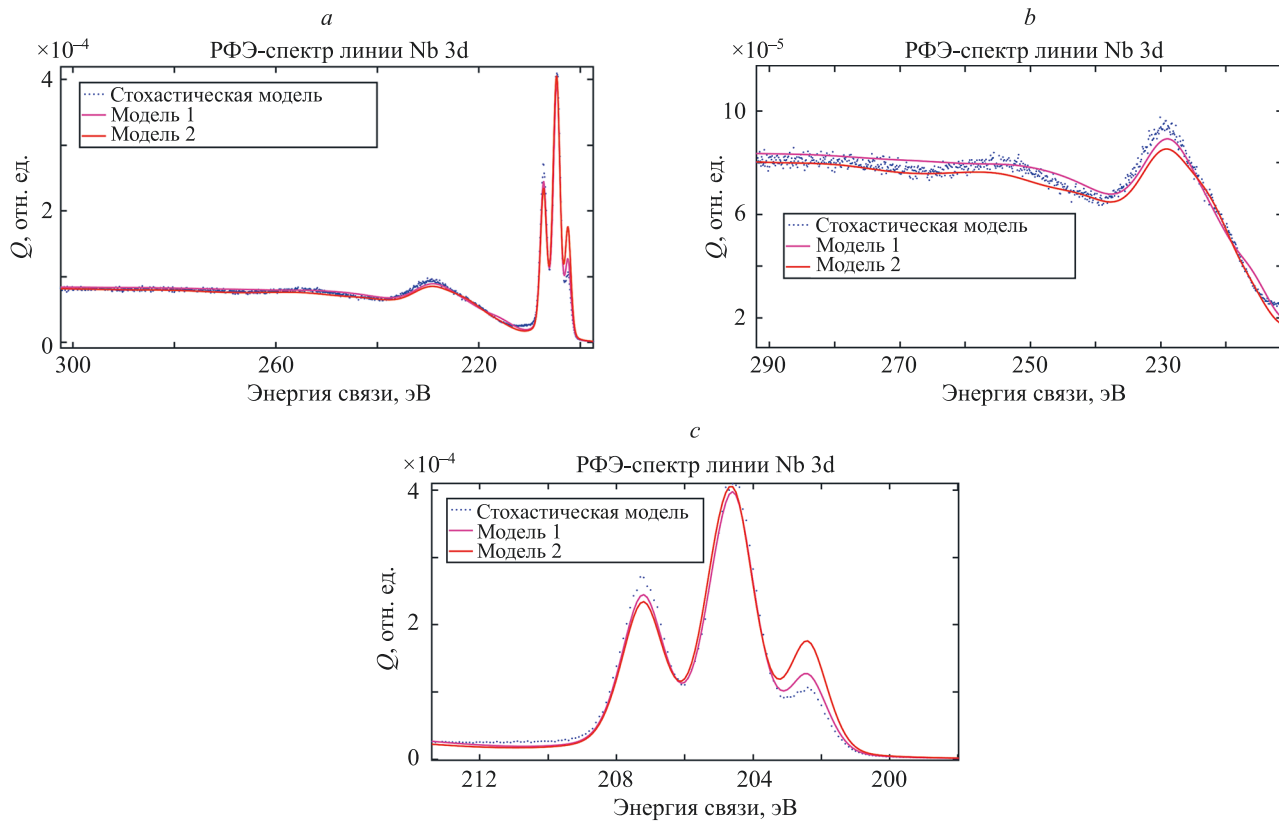


Рис. 4. Рентгеновский фотоэлектронный спектр линии Nb 3d: спектры моделей (a); приближенные части спектров для диапазонов от 214 до 290 эВ (b) и от 198 до 213 эВ (c)

Fig. 4. X-ray photoelectron spectrum of line Nb 3d: spectra of models (a); enlarged parts of the spectra for the ranges from 214 up to 290 eV (b); and from 198 up to 213 eV (c)

Левенгерда–Маккаварда действительный спектр описан по моделям 1 (рис. 2, a) и 2 (рис. 2, b). В качестве примера на рис. 4 представлены результаты моделирования рентгеновских фотоэлектронных (РФЭ) спектров линии Nb 3d для параметров стохастического слоя $d_{\max} = 9$ нм, $\sigma = 2$, $d_{\text{ср}} = 4$ нм. Спектры, полученные по моделям 1 и 2, недостаточно описывают спектр неоднородной стохастической поверхности. Для подробного исследования применимости моделей 1 и 2 выполнена серия моделирований для различных параметров стохастического слоя ($d_{\max}$, σ , $d_{\text{ср}}$).

Результаты и обсуждение

Зададим неоднородную поверхность нормальным распределением Гаусса для разных дисперсий толщин (рис. 5, b), максимальной (рис. 5, d) и средней толщины (рис. 5, f) поверхностного слоя. Сравнение результатов послойного анализа для различных функций стохастической поверхности приведено на рис. 5, a, c, e. Для точности описания рентгеновского фотоэлектронного спектра введем среднее квадратичное отклонение. Значение равно 1 свидетельствует о полном описании действительного спектра выбранной моделью. Выполним сравнение толщины слоя, полученной по моделям 1 и 2 ($d_{\text{моделей}}$), со средней толщиной слоя

стохастической модели ($d_{\text{стохастической}}$). Если отношение толщин слоев равно 1, то можно говорить о правильном описании неоднородной поверхности выбранной моделью.

Результаты анализа для различной дисперсии толщины слоя (рис. 5, a) показывают, что при увеличении разброса высот модель 2 неудовлетворительно описывает спектр стохастической модели, отклонения толщин составляет до 30 % при среднем квадратичном отклонении около 32 %. В таком случае описание поверхности моделью 1 приводит к завышенному значению толщины на 30 % при точности 95 %.

Послойный анализ поверхности для различной максимальной толщины слоя (рис. 5, c) показал, что при утолщении слоя спектр модели 1 практически полностью повторяет действительный, при этом рассчитанная толщина оказывается примерно на 40 % выше фактической. Описание стохастической поверхности моделью 2 снова дает некорректные результаты, можно предположить, что это связано с сильно неоднородной поверхностью. Модель 1 хорошо описывает поверхность при максимальной толщине слоя 5 нм, а при больших толщинах дает сильно завышенные результаты.

На рис. 5, e приведены результаты послойного анализа для различной средней толщины слоя. При сред-

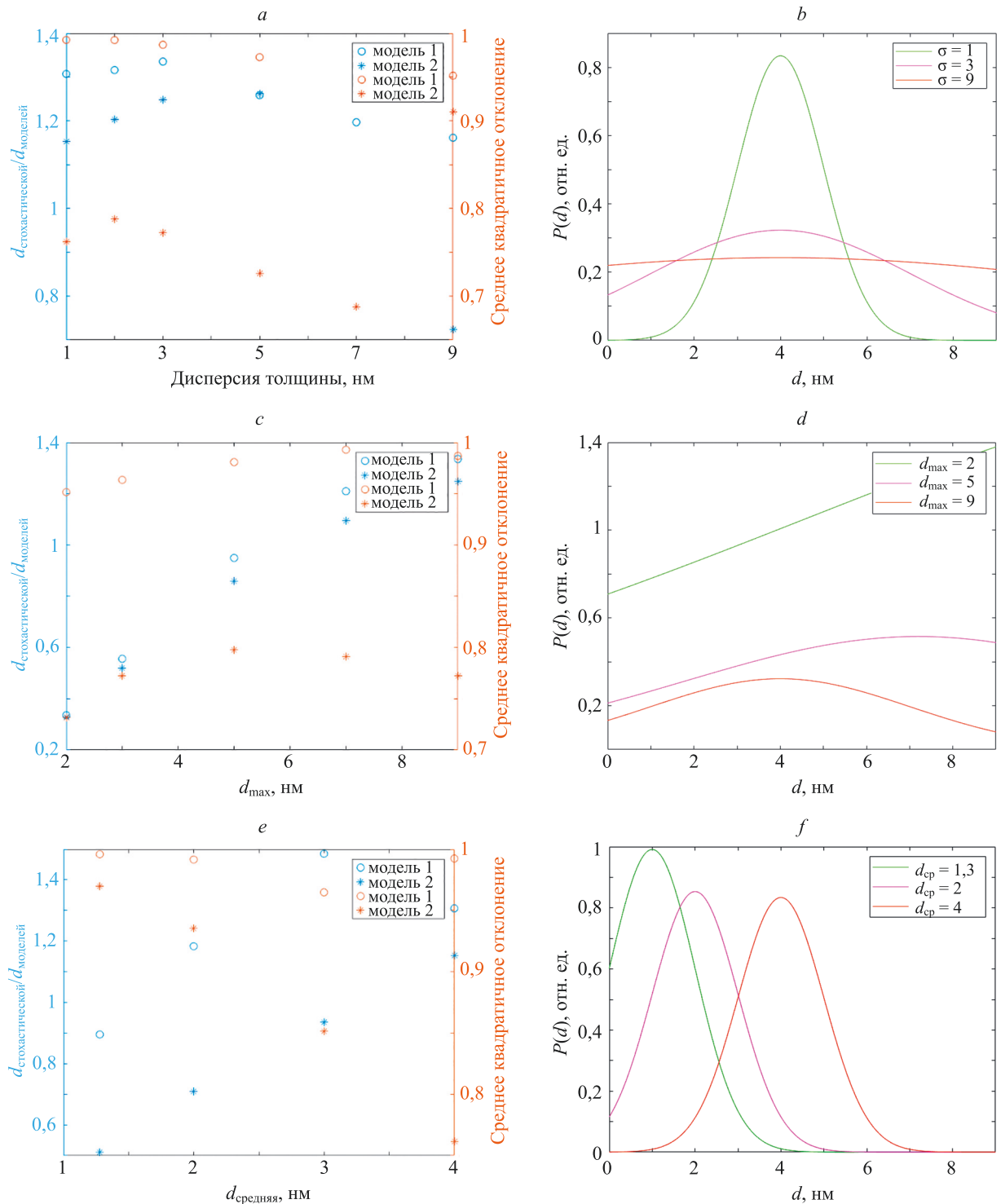


Рис. 5. Сравнение результатов послойного анализа (a, c, e) для различных функций стохастической поверхности (b, d, f)

Fig. 5. Comparison of layer profiling results (a, c, e) for various stochastic surfaces (b, d, f)

ней толщине 1,3 нм модель 1 точно описывает вид действительного спектра, но значение рассчитанной толщины занижено на 10 % по сравнению со средней толщиной модели стохастического слоя. При появлении дырок (при увеличении средней толщины) на сильно неоднородной поверхности модель 2 не пригодна для анализа мишени, а модель 1 дает завышенные результаты на 30 %.

Заключение

В работе предложена модель стохастической наноразмерной поверхности ультратонкой пленки. Результаты послойного анализа поверхности разными моделями показали, что описание стохастической наноразмерной поверхности моделями плоскопараллельных слоев и островковых наноструктур является затрудни-

тельным. Но тем не менее, если поверхность стремится к однородной, т. е. когда разброс высот небольшой, то модель плоскопараллельных слоев показывает хорошие результаты послойного анализа. Модель островковых наноструктур неудачно описывает различные неоднородности на поверхности. Таким образом, для более

точного послойного анализа ультратонкой пленки необходимо учитывать неоднородность реальной поверхности, в противном случае полученные результаты будут не соответствовать фактическим.

Литература

1. Tougaard S. Improved XPS analysis by visual inspection of the survey spectrum // *Surface and Interface Analysis*. 2018. V. 50. N 6. P. 657–666. <https://doi.org/10.1002/sia.6456>
2. Lubenchenko A.V., Batrakov A.A., Pavolotsky A.B., Lubenchenko O.I., Ivanov D.A. XPS study of multilayer multicomponent films // *Applied Surface Science*. 2018. V. 427. P. 711–721. <https://doi.org/10.1016/j.apsusc.2017.07.256>
3. Lukiantsev D.S., Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Pavolotsky A.B., Iachuk V.A., Pavlov O.N. The Formation of nanosuboxide layers in the oxide of niobium in low-power ion beam of argon // *Proc. of the 3rd International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*. 2021. P. 1–4. <https://doi.org/10.1109/REEPE51337.2021.9388002>
4. Lubenchenko A.V., Batrakov A.A., Shurkaeva I.V., Pavolotsky A.B., Krause S., Ivanov D.A., Lubenchenko O.I. XPS study of niobium and niobium-nitride nanofilms // *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*. 2018. V. 12. N 4. P. 692–700. <https://doi.org/10.1134/S1027451018040134>
5. Лубенченко А.В., Иванов Д.А., Лубенченко О.И., Паволоцкий А.Б., Лукьянцев Д.С., Ячук В.А., Павлов О.Н. Формирование неоднородных оксидных и субоксидных слоев на ультратонкой металлической пленке при многократном окислении и ионном распылении // *Журнал технической физики*. 2022. Т. 92. № 8. С. 1172–1178. <https://doi.org/10.21883/JTF.2022.08.52779.68-22>
6. Martín-Concepción A.I., Yubero F., Espinós J.P., Tougaard S. Surface roughness and island formation effects in ARXPS quantification // *Surface and Interface Analysis*. 2004. V. 36. N 8. P. 788–792. <https://doi.org/10.1002/sia.1765>
7. Varsányi G., Réé K., Mink G., Mohai M. Consideration of two dimensional surface roughnesses in quantitative XPS analysis // *Periodica Polytechnica Chemical Engineering*. 1987. V. 31. N 1-2. P. 3–17.
8. Fadley C.S., Baird R.J., Siekhaus W., Novakov T., Bergström S.Å.L. Surface analysis and angular distributions in X-ray photoelectron spectroscopy // *Journal of Electron Spectroscopy and Related Phenomena*. 1974. V. 4. N 2. P. 93–137. [https://doi.org/10.1016/0368-2048\(74\)90001-2](https://doi.org/10.1016/0368-2048(74)90001-2)
9. Zemek J. Electron spectroscopy of corrugated solid surfaces // *Analytical Sciences*. 2010. V. 26. N 2. P. 177–186. <https://doi.org/10.2116/analsci.26.177>
10. Kataev E., Wechsler D., Williams F.J., Köbl J., Tsud N., Franchi S., Steinrück H.-P., Lytken O. Probing the roughness of porphyrin thin films with X-ray photoelectron spectroscopy // *ChemPhysChem*. 2020. V. 21. N 20. P. 2293–2300. <https://doi.org/10.1002/cphc.202000568>
11. Mohai M. XPS MultiQuant: multimodel XPS quantification software // *Surface and Interface Analysis*. 2004. V. 36. N 8. P. 828–832. <https://doi.org/10.1002/sia.1775>
12. Mohai M. Calculation of layer thickness on rough surfaces by polyhedral model // *Surface and Interface Analysis*. 2008. V. 40. N 3–4. P. 710–713. <https://doi.org/10.1002/sia.2751>
13. Olejnik K., Zemek J., Werner W.S.M. Angular-resolved photoelectron spectroscopy of corrugated surfaces // *Surface Science*. 2005. V. 595. N 1–3. P. 212–222. <https://doi.org/10.1016/j.susc.2005.08.014>
14. Leprince-Wang Y., Yu-Zhang K. Study of the growth morphology of TiO₂ thin films by AFM and TEM // *Surface and Coatings Technology*. 2001. V. 140. N 2. P. 155–160. [https://doi.org/10.1016/S0257-8972\(01\)01029-5](https://doi.org/10.1016/S0257-8972(01)01029-5)
15. Wu O.K.T., Peterson G.G., LaRocca W.J., Butler E.M. ESCA signal intensity dependence on surface area (roughness) // *Applications of Surface Science*. 1982. V. 11–12. P. 118–130. [https://doi.org/10.1016/0378-5963\(82\)90058-7](https://doi.org/10.1016/0378-5963(82)90058-7)
16. Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Ivanova I.V. Formation of inelastic scattered background photoelectrons, X-ray

References

1. Tougaard S. Improved XPS analysis by visual inspection of the survey spectrum. *Surface and Interface Analysis*, 2018, vol. 50, no. 6, pp. 657–666. <https://doi.org/10.1002/sia.6456>
2. Lubenchenko A.V., Batrakov A.A., Pavolotsky A.B., Lubenchenko O.I., Ivanov D.A. XPS study of multilayer multicomponent films. *Applied Surface Science*, 2018, vol. 427, pp. 711–721. <https://doi.org/10.1016/j.apsusc.2017.07.256>
3. Lukiantsev D.S., Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Pavolotsky A.B., Iachuk V.A., Pavlov O.N. The Formation of nanosuboxide layers in the oxide of niobium in low-power ion beam of argon. *Proc. of the 3rd International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)*, 2021, pp. 1–4. <https://doi.org/10.1109/REEPE51337.2021.9388002>
4. Lubenchenko A.V., Batrakov A.A., Shurkaeva I.V., Pavolotsky A.B., Krause S., Ivanov D.A., Lubenchenko O.I. XPS study of niobium and niobium-nitride nanofilms. *Journal of Surface Investigation: X-ray, Synchrotron and Neutron Techniques*, 2018, vol. 12, no. 4, pp. 692–700. <https://doi.org/10.1134/S1027451018040134>
5. Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Pavolotsky A.B., Lukiantsev D.S., Iachuk V.A., Pavlov O.N. Formation of inhomogeneous oxide and suboxide layers on an ultra-thin metal film by multiple oxidation and ion sputtering. *Zhurnal tekhnicheskoy fiziki*, 2022, vol. 92, no. 8, pp. 1172–1178. (in Russian). <https://doi.org/10.21883/JTF.2022.08.52779.68-22>
6. Martín-Concepción A.I., Yubero F., Espinós J.P., Tougaard S. Surface roughness and island formation effects in ARXPS quantification. *Surface and Interface Analysis*, 2004, vol. 36, no. 8, pp. 788–792. <https://doi.org/10.1002/sia.1765>
7. Varsányi G., Réé K., Mink G., Mohai M. Consideration of two dimensional surface roughnesses in quantitative XPS analysis. *Periodica Polytechnica Chemical Engineering*, 1987, vol. 31, no. 1-2, pp. 3–17.
8. Fadley C.S., Baird R.J., Siekhaus W., Novakov T., Bergström S.Å.L. Surface analysis and angular distributions in X-ray photoelectron spectroscopy. *Journal of Electron Spectroscopy and Related Phenomena*, 1974, vol. 4, no. 2, pp. 93–137. [https://doi.org/10.1016/0368-2048\(74\)90001-2](https://doi.org/10.1016/0368-2048(74)90001-2)
9. Zemek J. Electron spectroscopy of corrugated solid surfaces. *Analytical Sciences*, 2010, vol. 26, no. 2, pp. 177–186. <https://doi.org/10.2116/analsci.26.177>
10. Kataev E., Wechsler D., Williams F.J., Köbl J., Tsud N., Franchi S., Steinrück H.-P., Lytken O. probing the roughness of porphyrin thin films with X-ray photoelectron spectroscopy. *ChemPhysChem*, 2020, vol. 21, no. 20, pp. 2293–2300. <https://doi.org/10.1002/cphc.202000568>
11. Mohai M. XPS MultiQuant: multimodel XPS quantification software. *Surface and Interface Analysis*, 2004, vol. 36, no. 8, pp. 828–832. <https://doi.org/10.1002/sia.1775>
12. Mohai M. Calculation of layer thickness on rough surfaces by polyhedral model. *Surface and Interface Analysis*, 2008, vol. 40, no. 3–4, pp. 710–713. <https://doi.org/10.1002/sia.2751>
13. Olejnik K., Zemek J., Werner W.S.M. Angular-resolved photoelectron spectroscopy of corrugated surfaces. *Surface Science*, 2005, vol. 595, no. 1-3, pp. 212–222. <https://doi.org/10.1016/j.susc.2005.08.014>
14. Leprince-Wang Y., Yu-Zhang K. Study of the growth morphology of TiO₂ thin films by AFM and TEM. *Surface and Coatings Technology*, 2001, vol. 140, no. 2, pp. 155–160. [https://doi.org/10.1016/S0257-8972\(01\)01029-5](https://doi.org/10.1016/S0257-8972(01)01029-5)
15. Wu O.K.T., Peterson G.G., LaRocca W.J., Butler E.M. ESCA signal intensity dependence on surface area (roughness). *Applications of Surface Science*, 1982, vol. 11–12, pp. 118–130. [https://doi.org/10.1016/0378-5963\(82\)90058-7](https://doi.org/10.1016/0378-5963(82)90058-7)
16. Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Ivanova I.V. Formation of inelastic scattered background photoelectrons, X-ray

- photoelectron spectroscopy from multilayer inhomogeneous surface // *Journal of Physics: Conference Series*. 2019. V. 1370. N 1. P. 012049. <https://doi.org/10.1088/1742-6596/1370/1/012049>
17. Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Yachuk V.A., Pavlov O.N., Lashkov I.A., Lukyantsev D.S. Non-destructive chemical and phase layer profiling of multicomponent multilayer thin ultrathin films // *Journal of Physics: Conference Series*. 2019. V. 1370. N 1. P. 012048. <https://doi.org/10.1088/1742-6596/1370/1/012048>

- photoelectron spectroscopy from multilayer inhomogeneous surface. *Journal of Physics: Conference Series*, 2019, vol. 1370, no. 1, pp. 012049. <https://doi.org/10.1088/1742-6596/1370/1/012049>
17. Lubenchenko A.V., Ivanov D.A., Lubenchenko O.I., Yachuk V.A., Pavlov O.N., Lashkov I.A., Lukyantsev D.S. Non-destructive chemical and phase layer profiling of multicomponent multilayer thin ultrathin films. *Journal of Physics: Conference Series*, 2019, vol. 1370, no. 1, pp. 012048. <https://doi.org/10.1088/1742-6596/1370/1/012048>

Авторы

Лукьянцев Денис Сергеевич — аспирант, Национальный исследовательский университет «МЭИ», Москва, 111250, Российская Федерация, stavden1996@mail.ru, <https://orcid.org/0000-0003-4350-8034>

Лубенченко Александр Владимирович — доктор технических наук, профессор, Национальный исследовательский университет «МЭИ», Москва, 111250, Российская Федерация, LubenchenkoAV@mpei.ru, <https://orcid.org/0000-0002-5998-5495>

Иванов Дмитрий Александрович — кандидат технических наук, доцент, профессор, Национальный исследовательский университет «МЭИ», Москва, 111250, Российская Федерация, iv_prof@mail.ru, <https://orcid.org/0000-0003-3760-0598>

Лубенченко Ольга Игоревна — старший преподаватель, Национальный исследовательский университет «МЭИ», Москва, 111250, Российская Федерация, IvanovaOI@mpei.ru, <https://orcid.org/0000-0002-6317-2841>

Федотов Александр Сергеевич — студент, Национальный исследовательский университет «МЭИ», Москва, 111250, Российская Федерация, courage_man@mail.ru, <https://orcid.org/0000-0002-7641-3173>

Authors

Denis S. Lukyantsev — PhD Student, National Research University “Moscow Power Engineering Institute”, Moscow, 111250, Russian Federation, stavden1996@mail.ru, <https://orcid.org/0000-0003-4350-8034>

Alexander V. Lubenchenko — D. Sc., Professor, National Research University “Moscow Power Engineering Institute”, Moscow, 111250, Russian Federation, LubenchenkoAV@mpei.ru, <https://orcid.org/0000-0002-5998-5495>

Dmitriy A. Ivanov — PhD, Associate Professor, Professor, National Research University “Moscow Power Engineering Institute”, Moscow, 111250, Russian Federation, iv_prof@mail.ru, <https://orcid.org/0000-0003-3760-0598>

Olga I. Lubenchenko — Senior Lecturer, National Research University “Moscow Power Engineering Institute”, Moscow, 111250, Russian Federation, IvanovaOI@mpei.ru, <https://orcid.org/0000-0002-6317-2841>

Alexandr S. Fedotov — Student, National Research University “Moscow Power Engineering Institute”, Moscow, 111250, Russian Federation, courage_man@mail.ru, <https://orcid.org/0000-0002-7641-3173>

Статья поступила в редакцию 18.08.2022
Одобрена после рецензирования 10.10.2022
Принята к печати 20.11.2022

Received 18.08.2022
Approved after reviewing 10.10.2022
Accepted 20.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1112-1118

УДК 538.958

Органические светоизлучающие диоды с новыми красителями на основе кумарина

Анна Владимировна Осадченко¹✉, Андрей Александрович Ващенко²,
Иван Александрович Захарчук³, Даниил Саюзович Дайбаге⁴,
Сергей Александрович Амброзевич⁵, Никита Юрьевич Володин⁶,
Дмитрий Андреевич Чепцов⁷, Сергей Михайлович Долотов⁸,
Валерий Федорович Травень⁹, Антон Игоревич Авраменко¹⁰,
Светлана Леонидовна Семенова¹¹, Александр Сергеевич Селюков¹²

^{1,3,4,5,12} Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация

^{1,4,12} Московский политехнический университет, Москва, 107023, Российская Федерация

^{2,4,5} Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация

^{6,7,8,9} Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация

^{10,11} Всероссийский институт научной и технической информации Российской академии наук, Москва, 125190, Российская Федерация

¹ Anna.vl.osadchenko@gmail.com✉, <https://orcid.org/0000-0001-9556-4885>

² vashchenkoa@lebedev.ru, <https://orcid.org/0000-0003-2084-5900>

³ zakharchukia@yandex.ru, <https://orcid.org/0000-0002-1502-6460>

⁴ daibagya@mail.ru, <https://orcid.org/0000-0003-1944-1546>

⁵ s.ambrozevich@mail.ru, <https://orcid.org/0000-0002-3906-0735>

⁶ nikita9963@ya.ru, <https://orcid.org/0000-0002-7801-9399>

⁷ dchepcov@yandex.ru, <https://orcid.org/0000-0002-9774-7922>

⁸ dolism@mail.ru, <https://orcid.org/0000-0002-0022-0535>

⁹ valerii.traven@gmail.com, <https://orcid.org/0000-0002-2204-7438>

¹⁰ anton1905@internet.ru, <https://orcid.org/0000-0002-0374-6428>

¹¹ s.l.semenova@internet.ru, <https://orcid.org/0000-0001-6899-1347>

¹² selyukov@lebedev.ru, <https://orcid.org/0000-0002-4007-6291>

Аннотация

Предмет исследования. Представлены результаты исследования люминесцентных свойств органических светоизлучающих диодов на основе новых люминесцентных соединений, содержащих кумариновый фрагмент. **Метод.** Изготовление светоизлучающих диодов проведено методами центрифугирования и термического напыления в вакууме в условиях чистой комнаты. Измерение характеристик светодиодов выполнено методами оптической спектроскопии, а также электрическими методами. **Основные результаты.** Экспериментально показано, что формирование активного слоя светодиода на основе люминесцентных соединений, содержащих кумариновое ядро, приводит к образованию димеров, спектр люминесценции которых существенно отличается от спектра исходного соединения в растворе толуола. Преобразование структуры соединения привело к изменению вольтамперных характеристик результирующего устройства и спектров свечения. Данные измерения возникли из-за различий электронной структуры исследуемых соединений, а также разницы значений подвижностей носителей заряда и высот потенциальных барьеров, возникающих на гетерогранице с другими рабочими слоями светодиода. **Практическая значимость.** Полученные результаты могут служить основой для систематизации знаний о зависимости свойств новых люминесцентных соединений, в состав которых входит кумариновое ядро, от их структуры. Структуры, разработанные в рамках работы, могут стать прототипами для промышленно выпускаемых светоизлучающих устройств, излучающих в том числе белый свет.

Ключевые слова

кумариновые красители, органические светоизлучающие диоды, электролюминесценция, связь структурно-свойства, вольтамперные характеристики

Благодарности

Исследование проведено в рамках проекта РФФИ 20-02-00222 А.

Ссылка для цитирования: Осадченко А.В., Ващенко А.А., Захарчук И.А., Дайбаге Д.С., Амброзевич С.А., Володин Н.Ю., Чепцов Д.А., Долотов С.М., Травень В.Ф., Авраменко А.И., Семенова С.Л., Селюков А.С. Органические светоизлучающие диоды с новыми красителями на основе кумарина // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1112–1118. doi: 10.17586/2226-1494-2022-22-6-1112-1118

Organic light-emitting diodes with new dyes based on coumarin

Anna V. Osadchenko^{1✉}, Andrey A. Vashchenko², Ivan A. Zakharchuk³, Daniil S. Daibagya⁴, Sergey A. Ambrozevich⁵, Nikita Yu. Volodin⁶, Dmitry A. Cheptsov⁷, Sergey M. Dolotov⁸, Valery F. Traven⁹, Anton I. Avramenko¹⁰, Svetlana L. Semenova¹¹, Alexander S. Selyukov¹²

^{1,3,4,5,12} Bauman Moscow State Technical University, Moscow, 105005, Russian Federation

^{1,4,12} Moscow Polytechnic University, Moscow, 107023, Russian Federation

^{2,4,5} Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation

^{6,7,8,9} Mendelev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation

^{10,11} Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences, Moscow, 125190, Russian Federation

¹ Anna.vl.osadchenko@gmail.com✉, <https://orcid.org/0000-0001-9556-4885>

² vashchenkooa@lebedev.ru, <https://orcid.org/0000-0003-2084-5900>

³ zakharchukia@yandex.ru, <https://orcid.org/0000-0002-1502-6460>

⁴ daibagya@mail.ru, <https://orcid.org/0000-0003-1944-1546>

⁵ s.ambrozevich@mail.ru, <https://orcid.org/0000-0002-3906-0735>

⁶ nikita9963@ya.ru, <https://orcid.org/0000-0002-7801-9399>

⁷ dcheptsov@yandex.ru, <https://orcid.org/0000-0002-9774-7922>

⁸ dolsm@mail.ru, <https://orcid.org/0000-0002-0022-0535>

⁹ valerii.traven@gmail.com, <https://orcid.org/0000-0002-2204-7438>

¹⁰ anton1905@internet.ru, <https://orcid.org/0000-0002-0374-6428>

¹¹ s.l.semenova@internet.ru, <https://orcid.org/0000-0001-6899-1347>

¹² selyukov@lebedev.ru, <https://orcid.org/0000-0002-4007-6291>

Abstract

The results of studying the luminescent properties of organic light-emitting diodes based on new luminescent compounds containing a coumarin fragment are presented. The light-emitting diodes were fabricated by spin-coating and thermal evaporation in an argon atmosphere in a clean room. Measurements of the LED characteristics were carried out by optical spectroscopy, as well as by electrical methods. It has been experimentally shown that deposition of an OLED active layer based on luminescent compounds containing a coumarin core can lead to the formation of dimers, the luminescence spectra of which differ significantly from the corresponding spectra of the original materials in toluene. A variation in the structure of the compound leads to a change in both the current-voltage characteristics of the resulting device and the luminescence spectra. These changes appeared due to the difference in the electronic structure of these materials as well as due to different values of charge carrier mobilities and the potential barriers at the heterointerface with other OLED layers. The results obtained may serve as the basis for systematizing knowledge about the dependence of the properties of new luminescent materials, which include a coumarin core, on their structure. The developed structures can become prototypes for industrially produced light-emitting devices that specifically emit white light.

Keywords

coumarin dyes, organic light-emitting diodes, electroluminescence, structure-properties relationship, current-voltage characteristics

Acknowledgements

The study was carried out within the RFBR project 20-02-00222 А.

For citation: Osadchenko A.V., Vashchenko A.A., Zakharchuk I.A., Daibagya D.S., Ambrozevich S.A., Volodin N.Yu., Cheptsov D.A., Dolotov S.M., Traven V.F., Avramenko A.I., Semenova S.L., Selyukov A.S. Organic light-emitting diodes with new dyes based on coumarin. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1112–1118 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1112-1118

Введение

Одна из актуальных задач в последнее время — создание эффективных, стабильных и долговечных источников излучения для задач освещения и индикации. Одним из подходов к решению данной задачи

является изготовление органических светоизлучающих диодов на основе как органических [1, 2], так и неорганических [3–6] люминесцентных материалов. Использование подобных устройств позволяет добиваться высоких значений яркости свечения с поверхности большой площади. Кроме того, дисплеи на основе ор-

ганических светодиодов обладают существенным преимуществом перед жидкокристаллическими панелями, поскольку они имеют максимально возможный угол обзора, а также широкий цветовой охват.

Тем не менее, к настоящему времени не решена проблема подбора надежных и в то же время дешевых материалов, обеспечивающих высокие показатели эффективности свечения. Используемые в промышленных образцах люминесцентные материалы, как правило, построены на основе металлоорганических комплексов [7, 8], в свой состав включающих металлы платиновой группы [9–11]. Эти соединения оказываются дорогими не только в силу стоимости исходных прекурсоров, но и из-за большой стоимости синтеза. В настоящей работе исследованы органические светоизлучающие диоды, изготовленные на основе новых люминесцентных материалов, содержащих кумариновый фрагмент. Эти материалы дешевы и могут быть исключительно перспективными благодаря процессам термостимулированной задержанной флуоресценции [12].

Техника эксперимента

Для изготовления светодиодов в соответствии с методикой, представленной в работе [13], были синтезированы два новых соединения, содержащих кумариновое ядро. К кумариновому ядру в положение 3 были прикреплены массивные заместители на основе: антрацена ((E)-3-(3-(антрацен-9-ил)акрилоил)-2H-хромен-2-он (соединение А) и карбоксильной группы (этил 7-(диэтиламино)-2-оксо-2H-хромен-3-карбоксилат (соединение В) с добавлением в положение 7 диэтиламинной группы (рис. 1).

Спектры экстинкции получены для растворов исследованных соединений в толуоле с концентрацией $2,5 \cdot 10^{-3}$ мг/мл с помощью спектрофотометра Perkin-ElmerLambda 45, чувствительного в диапазоне длин волн 190–1100 нм. Измерения проведены в кварцевых кюветках с длиной оптического пути 10 мм и прозрачных — в диапазоне 200–2000 нм.

Измерения спектров люминесценции при электрическом или оптическом возбуждении выполнены с помощью волоконного мини-ПЗС-спектрометра OceanOpticsMaya 2000 Pro, с рабочим диапазоном длин волн 200–1100 нм. Для оптического возбуждения применялся лазер PicoQuantLDH-C 400, работающий в

импульсном режиме с длительностью импульса 75 пс, частотой следования импульсов 40 МГц, и излучающий на длине волны 405 нм. В качестве образцов использованы те же растворы, что и при измерении спектров экстинкции.

Изготовление органических светоизлучающих диодов произведено в условиях чистой комнаты, оборудованной перчаточными боксами с аргонной атмосферой, содержащей не более 100 ppm кислорода и 300 ppm воды. В боксах проведены операции нанесения тонких пленок из жидкой фазы методом центрифугирования (spin-coating). Методом центрифугирования на диоды нанесены слои органического дырочного инжектирующего слоя PEDOT:PSS из водного раствора, дырочного транспортного слоя poly-TPD из раствора в хлорбензоле, а также исследованных люминесцентных соединений из раствора в толуоле. Нанесение электродов, а также электронного транспортного слоя TPBi осуществлено с помощью встроенной в аргонный перчаточный бокс напылительной установки LeyboldUnivex 300, позволяющей нанести нескольких рабочих слоев различных материалов без нарушения вакуума. В качестве катода последовательно нанесены два слоя: LiF толщиной 1 нм и Al — 100 нм. В качестве анода использованы стеклянные подложки с предварительно нанесенным слоем оксида индия-олова (ITO) производства LumTec (Тайвань).

Измерение вольтамперных характеристик выполнено с помощью установки на основе вольтметра B7-78/1, амперметра Keithley 6485, а также линейного стабилизированного источника питания MotechPPS 2017, управляемого с помощью персонального компьютера.

Результаты и их обсуждение

Результаты измерения спектров экстинкции и флуоресценции соединений А и В представлены на рис. 2. У соединения А край поглощения наблюдается в области $25\,000\text{ см}^{-1}$, а в области $28\,000\text{--}30\,000\text{ см}^{-1}$ видно плато с характерным значением коэффициента экстинкции порядка $1,0 \cdot 10^4\text{ л} \cdot \text{моль}^{-1} \cdot \text{см}^{-1}$. С увеличением волнового числа заметно существенное увеличение оптического поглощения. При этом максимум флуоресценции у соединения А расположен в области $17\,700\text{ см}^{-1}$ и имеет ширину на полувысоте порядка 3000 см^{-1} , он оказывается существенно сдвинутым по отношению к краю поглощения на 7300 см^{-1} . Отметим,

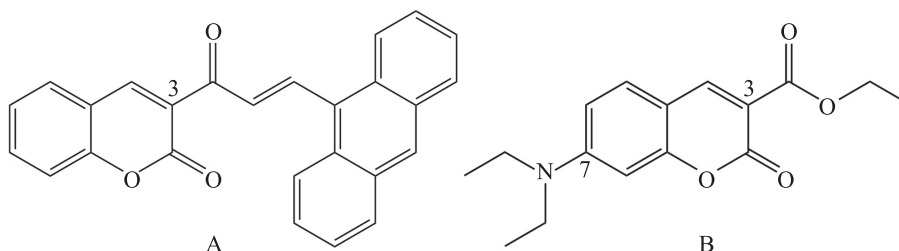


Рис. 1. Структурные формулы исследованных соединений: (E)-3-(3-(антрацен-9-ил)акрилоил)-2H-хромен-2-он (А); этил 7-(диэтиламино)-2-оксо-2H-хромен-3-карбоксилат (В)

Fig. 1. Structural formulas of the studied compounds: (E)-3-(3-(anthracen-9-yl)acryloyl)-2H-chromen-2-one (А), ethyl 7-(diethylamino)-2-oxo-2H-chromen-3-carboxylate (В)

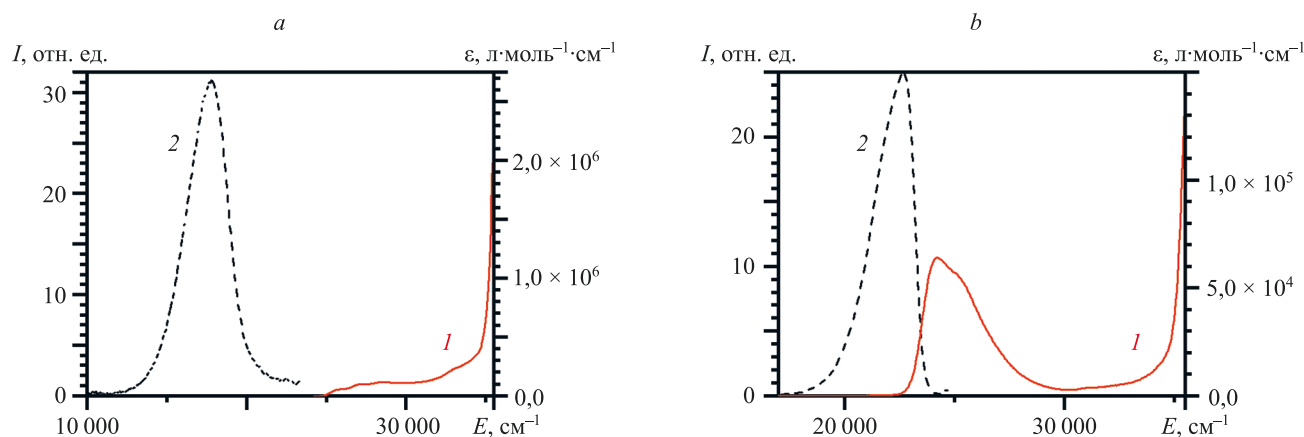


Рис. 2. Спектры экстинкции (кривая 1) и флуоресценции (кривая 2) для соединений А (а) и В (b). 1 — интенсивность, E — энергия, ϵ — коэффициент экстинкции

Fig. 2. Extinction (curve 1) and fluorescence (curve 2) spectra for compounds A and B

что это может быть связано с изменением конформации сложной молекулы, содержащей кумариновый и антраценовый фрагменты, при поглощении кванта возбуждения [14].

У соединения В максимум поглощения наблюдается в области $24\,300\text{ см}^{-1}$, при этом его ширина на полувысоте составляет 2900 см^{-1} . Полоса имеет сложную структуру и состоит как минимум из двух элементарных полос. Возникновение более высокоэнергетической полосы обусловлено, предположительно, образованием димеров в исследованном растворе [15]. Максимум флуоресценции расположен зеркально и находится в области $22\,700\text{ см}^{-1}$, а его ширина на полувысоте составляет 2100 см^{-1} . Положение бесфонной линии находится в области $23\,400\text{ см}^{-1}$. Величина коэффициента экстинкции в максимуме поглощения составляет $6,4 \cdot 10^4\text{ л} \cdot \text{моль}^{-1} \cdot \text{см}^{-1}$. В области $30\,000\text{ см}^{-1}$ коэффициент экстинкции падает до значений порядка $2,7 \cdot 10^3\text{ л} \cdot \text{моль}^{-1} \cdot \text{см}^{-1}$, а затем, с ростом волнового числа, увеличивается.

Исследование электролюминесценции проведено на образцах светоизлучающих диодов, имеющих

следующую структуру: ITO/PEDOT:PSS (50 нм)/poly-TPD (20 нм)/исследуемое соединение/TPBi (20 нм)/LiF (1 нм)/Al (100 нм). Спектры электролюминесценции изготовленных структур представлены на рис. 3, а.

В спектре электролюминесценции светодиода на основе соединения А, полученном при напряжении 6,5 В, наблюдается максимум в области $17\,200\text{ см}^{-1}$, что практически полностью соответствует положению максимума полосы флуоресценции соединения при оптическом возбуждении. Отметим, что видно небольшое увеличение ширины полосы на полувысоте, которая составила 3200 см^{-1} . В области $22\,000\text{--}28\,000\text{ см}^{-1}$ наблюдается слабое свечение, связанное с люминесценцией транспортного слоя TPBi, повторяющего спектр флуоресценции этого слоя [16].

В спектре электролюминесценции устройства, изготовленного на основе соединения В, при приложении напряжения 5 В наблюдаются два ярко выраженных максимума в областях $21\,600$ и $18\,300\text{ см}^{-1}$. Первый максимум соответствует максимуму флуоресценции при оптическом возбуждении, наблюдавшемуся в области $22\,700\text{ см}^{-1}$, и сдвинут в низкоэнергетическую

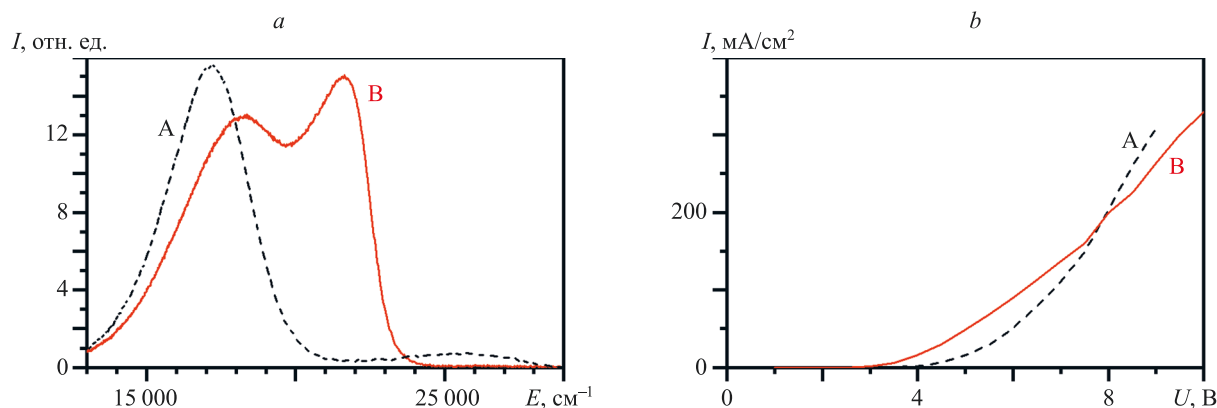


Рис. 3. Спектры электролюминесценции (а) и вольтамперные характеристики (b) светодиодных структур, изготовленных на основе соединений (А) и (В)

Fig. 3. Electroluminescence spectra (a) and current–voltage characteristics (b) of LED structures based on compounds (A) and (B)

область на 1100 см^{-1} . Второй максимум не наблюдался в спектре флуоресценции растворов и связан с образованием димеров [17] при изготовлении излучающего слоя светодиода.

Вольтамперные характеристики исследованных светодиодных структур представлены на рис. 3, б, и имеют характерный для органических светодиодов нелинейный вид. Видно, что напряжение включения светодиода, изготовленного на основе соединения А, равно примерно 3 В, а для соединения В — 4 В. Полученная разница предположительно связана с различием величин подвижностей носителей заряда в исследуемых соединениях [18]. Также разница может определяться различными высотами потенциальных барьеров [19], возникающих на границах с другими слоями светодиодов.

Заключение

В работе изготовлены источники света нового поколения — органические светоизлучающие диоды на основе новых люминесцентных соединений, содержащих кумариновое ядро. Экспериментально показано, что в ряде случаев в активном слое светодиода могут формироваться димеры, которые существенным образом изменяют спектр электролюминесценции по сравнению с люминесценцией при оптическом возбуждении. Это может оказаться преимуществом, особенно в случае, когда необходимо создать источники белого свечения. Полученные результаты являются основой для установления связи между структурой и люминесцентными свойствами новых кумариновых соединений. Существенное отличие спектров электролюминесценции свидетельствует о важности изготовления светодиодных структур, поскольку процессы люминесценции в них могут существенным образом отличаться от тех, что наблюдаются при оптическом возбуждении.

Литература

1. Luo J., Rong X.F., Ye Y.Y., Li W.Z., Wang X.-Q., Wang W. Research progress on triarylmethyl radical-based high-efficiency OLED // *Molecules*. 2022. V. 27. N 5. P. 1632. <https://doi.org/10.3390/molecules27051632>
2. Corrêa Santos D., Vieira Marques M.F. Blue light polymeric emitters for the development of OLED devices // *Journal of Materials Science: Materials in Electronics*. 2022. V. 33. N 16. P. 12529–12565. <https://doi.org/10.1007/s10854-022-08333-3>
3. Palomaki P. Quantum Dots + OLED = Your Next TV: Formerly rival technologies will come together in new Samsung displays // *IEEE Spectrum*. 2022. V. 59. N 1. P. 52–53. <https://doi.org/10.1109/MSPEC.2022.9676350>
4. Lee S., Hahm D., Yoon S.Y., Yang H., Bae W.K., Kwak J. Quantum-dot and organic hybrid light-emitting diodes employing a blue common layer for simple fabrication of full-color displays // *Nano Research*. 2022. V. 15. N 7. P. 6477–6482. <https://doi.org/10.1007/s12274-022-4204-y>
5. Ващенко А.А., Витухновский А.Г., Лебедев В.С., Селюков А.С., Васильев Р.Б., Соколик М.С. Органический светоизлучающий диод на основе плоского слоя полупроводниковых нанопластинок CdSe в качестве эмиттера // Письма в Журнал экспериментальной и теоретической физики. 2014. Т. 100. № 2. С. 94–98.
6. Селюков А.С., Витухновский А.Г., Лебедев В.С., Ващенко А.А., Васильев Р.Б., Соколик М.С. Электролюминесценция коллоидных квазидвумерных полупроводниковых наноструктур CdSe в гибридном светоизлучающем диоде // Журнал экспериментальной и теоретической физики. 2015. Т. 147. № 4. С. 687–701.
7. Eliseeva S.V., Bünzli J.C.G. Rare earths: jewels for functional materials of the future // *New Journal of Chemistry*. 2011. V. 35. N 6. P. 1165–1176. <https://doi.org/10.1039/C0NJ00969E>
8. Eliseeva S.V., Bünzli J.C.G. Lanthanide luminescence for functional materials and bio-sciences // *Chemical Society Reviews*. 2010. V. 39. N 1. P. 189–227. <https://doi.org/10.1039/B905604C>
9. Liang A., Ying L., Huang F. Recent progresses of iridium complex-containing macromolecules for solution-processed organic light-emitting diodes // *Journal of Inorganic and Organometallic Polymers and Materials*. 2014. V. 24. N 6. P. 905–926. <https://doi.org/10.1007/s10904-014-0099-8>
10. Wei Q., Fei N., Islam A., Lei T., Hong L., Peng R., Fan X., Chen L., Gao P., Ge Z. Small-molecule emitters with high quantum efficiency: mechanisms, structures, and applications in OLED devices // *Advanced Optical Materials*. 2018. V. 6. N 20. P. 1800512. <https://doi.org/10.1002/adom.201800512>
11. Jin J., Zhu Z., Yan J., Zhou X., Cao C., Chou P.-T., Zhang Y.-X., Zheng Z., Lee C.-S., Chi Y. Iridium (III) phosphors-bearing functional 9-phenyl-7, 9-dihydro-8h-purin-8-ylidene chelates and blue

References

1. Luo J., Rong X.F., Ye Y.Y., Li W.Z., Wang X.-Q., Wang W. Research progress on triarylmethyl radical-based high-efficiency OLED. *Molecules*, 2022, vol. 27, no. 5, pp. 1632. <https://doi.org/10.3390/molecules27051632>
2. Corrêa Santos D., Vieira Marques M.F. Blue light polymeric emitters for the development of OLED devices. *Journal of Materials Science: Materials in Electronics*, 2022, vol. 33, no. 16, pp. 12529–12565. <https://doi.org/10.1007/s10854-022-08333-3>
3. Palomaki P. Quantum Dots + OLED = Your Next TV: Formerly rival technologies will come together in new Samsung displays. *IEEE Spectrum*, 2022, vol. 59, no. 1, pp. 52–53. <https://doi.org/10.1109/MSPEC.2022.9676350>
4. Lee S., Hahm D., Yoon S.Y., Yang H., Bae W.K., Kwak J. Quantum-dot and organic hybrid light-emitting diodes employing a blue common layer for simple fabrication of full-color displays. *Nano Research*, 2022, vol. 15, no. 7, pp. 6477–6482. <https://doi.org/10.1007/s12274-022-4204-y>
5. Vashchenko A.A., Vitukhnovskii A.G., Lebedev V.S., Selyukov A.S., Vasiliev R.B., Sokolikova M.S. Organic light-emitting diode with an emitter based on a planar layer of CdSe semiconductor nanoplatelets. *JETP Letters*, 2014, vol. 100, no. 2, pp. 86–90. <https://doi.org/10.1134/S0021364014140124>
6. Selyukov A.S., Vitukhnovskii A.G., Lebedev V.S., Vashchenko A.A., Vasiliev R.B., Sokolikova M.S. Electroluminescence of colloidal quasi-two-dimensional semiconducting CdSe nanostructures in a hybrid light-emitting diode. *Journal of Experimental and Theoretical Physics*, 2015, vol. 120, no. 4, pp. 595–606. <https://doi.org/10.1134/S1063776115040238>
7. Eliseeva S.V., Bünzli J.C.G. Rare earths: jewels for functional materials of the future. *New Journal of Chemistry*, 2011, vol. 35, no. 6, pp. 1165–1176. <https://doi.org/10.1039/C0NJ00969E>
8. Eliseeva S.V., Bünzli J.C.G. Lanthanide luminescence for functional materials and bio-sciences. *Chemical Society Reviews*, 2010, vol. 39, no. 1, pp. 189–227. <https://doi.org/10.1039/B905604C>
9. Liang A., Ying L., Huang F. Recent progresses of iridium complex-containing macromolecules for solution-processed organic light-emitting diodes. *Journal of Inorganic and Organometallic Polymers and Materials*, 2014, vol. 24, no. 6, pp. 905–926. <https://doi.org/10.1007/s10904-014-0099-8>
10. Wei Q., Fei N., Islam A., Lei T., Hong L., Peng R., Fan X., Chen L., Gao P., Ge Z. Small-molecule emitters with high quantum efficiency: mechanisms, structures, and applications in OLED devices. *Advanced Optical Materials*, 2018, vol. 6, no. 20, pp. 1800512. <https://doi.org/10.1002/adom.201800512>
11. Jin J., Zhu Z., Yan J., Zhou X., Cao C., Chou P.-T., Zhang Y.-X., Zheng Z., Lee C.-S., Chi Y. Iridium (III) phosphors-bearing functional 9-phenyl-7, 9-dihydro-8h-purin-8-ylidene chelates and blue

- hyperphosphorescent OLED devices // *Advanced Photonics Research*, 2022, V. 3, N 7, P. 2100381. <https://doi.org/10.1002/adpr.202100381>
12. Chen J.X., Liu W., Zheng C.J., Wang K., Liang K., Shi Y.-Z., Ou X.-M., Zhang X.-H. Coumarin-based thermally activated delayed fluorescence emitters with high external quantum efficiency and low efficiency roll-off in the devices // *ACS Applied Materials & Interfaces*, 2017, V. 9, N 10, P. 8848–8854. <https://doi.org/10.1021/acsami.6b15816>
13. Traven V.F., Cheptsov D.A. Sensory effects of fluorescent organic dyes // *Russian Chemical Reviews*, 2020, V. 89, N 7, P. 713. <https://doi.org/10.1070/RCR4909>
14. Huang D., Chen Y., Zhao J. Access to a large stokes shift in functionalized fused coumarin derivatives by increasing the geometry relaxation upon photoexcitation: An experimental and theoretical study // *Dyes and Pigments*, 2012, V. 95, N 3, P. 732–742. <https://doi.org/10.1016/j.dyepig.2012.04.024>
15. Drexhage K.H. *Dye Lasers*. Springer-Verlag, 1973, 285 p.
16. Yin Y., Lü Z., Deng Z., Liu B., Mamytbekov Z.K., Hu B. White-light-emitting organic electroluminescent devices with poly-TPD as emitting layer // *Journal of Materials Science: Materials in Electronics*, 2017, V. 28, N 24, P. 19148–19154. <https://doi.org/10.1007/s10854-017-7871-9>
17. Kotchapradist P., Prachumrak N., Sunonnam T., Tarsang R., Namuangruk S., Sudyoasuk T., Keawin T., Jungsuttiwong S., Promarak V. N-coumarin derivatives as hole-transporting emitters for high efficiency solution-processed pure green electroluminescent devices // *Dyes and Pigments*, 2015, V. 112, P. 227–235. <https://doi.org/10.1016/j.dyepig.2014.06.032>
18. Sahoo R.K., Atta S., Singh N.D., Jacob C. Influence of functional derivatives of an amino-coumarin/MWCNT composite organic hetero-junction on the photovoltaic characteristics // *Materials Science in Semiconductor Processing*, 2014, V. 25, P. 279–285. <https://doi.org/10.1016/j.mssp.2014.01.001>
19. Yadav R.A.K., Dubey D.K., Chen S.Z., Liang T.W., Jou J.-H. Role of molecular orbital energy levels in OLED performance // *Scientific Reports*, 2020, V. 10, N 1, P. 9915. <https://doi.org/10.1038/s41598-020-66946-2>
- hyperphosphorescent OLED devices. *Advanced Photonics Research*, 2022, vol. 3, no. 7, pp. 2100381. <https://doi.org/10.1002/adpr.202100381>
12. Chen J.X., Liu W., Zheng C.J., Wang K., Liang K., Shi Y.-Z., Ou X.-M., Zhang X.-H. Coumarin-based thermally activated delayed fluorescence emitters with high external quantum efficiency and low efficiency roll-off in the devices. *ACS Applied Materials & Interfaces*, 2017, vol. 9, no. 10, pp. 8848–8854. <https://doi.org/10.1021/acsami.6b15816>
13. Traven V.F., Cheptsov D.A. Sensory effects of fluorescent organic dyes. *Russian Chemical Reviews*, 2020, vol. 89, no. 7, pp. 713. <https://doi.org/10.1070/RCR4909>
14. Huang D., Chen Y., Zhao J. Access to a large stokes shift in functionalized fused coumarin derivatives by increasing the geometry relaxation upon photoexcitation: An experimental and theoretical study. *Dyes and Pigments*, 2012, vol. 95, no. 3, pp. 732–742. <https://doi.org/10.1016/j.dyepig.2012.04.024>
15. Drexhage K.H. *Dye Lasers*. Springer-Verlag, 1973, 285 p.
16. Yin Y., Lü Z., Deng Z., Liu B., Mamytbekov Z.K., Hu B. White-light-emitting organic electroluminescent devices with poly-TPD as emitting layer. *Journal of Materials Science: Materials in Electronics*, 2017, vol. 28, no. 24, pp. 19148–19154. <https://doi.org/10.1007/s10854-017-7871-9>
17. Kotchapradist P., Prachumrak N., Sunonnam T., Tarsang R., Namuangruk S., Sudyoasuk T., Keawin T., Jungsuttiwong S., Promarak V. N-coumarin derivatives as hole-transporting emitters for high efficiency solution-processed pure green electroluminescent devices. *Dyes and Pigments*, 2015, vol. 112, pp. 227–235. <https://doi.org/10.1016/j.dyepig.2014.06.032>
18. Sahoo R.K., Atta S., Singh N.D., Jacob C. Influence of functional derivatives of an amino-coumarin/MWCNT composite organic hetero-junction on the photovoltaic characteristics. *Materials Science in Semiconductor Processing*, 2014, vol. 25, pp. 279–285. <https://doi.org/10.1016/j.mssp.2014.01.001>
19. Yadav R.A.K., Dubey D.K., Chen S.Z., Liang T.W., Jou J.-H. Role of molecular orbital energy levels in OLED performance. *Scientific Reports*, 2020, vol. 10, no. 1, pp. 9915. <https://doi.org/10.1038/s41598-020-66946-2>

Авторы

Осадченко Анна Владимировна — студент, Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация; ассистент, Московский политехнический университет, Москва, 107023, Российская Федерация, [sc 57439684100](https://orcid.org/0000-0001-9556-4885), <https://orcid.org/0000-0001-9556-4885>, anna.vl.osadchenko@gmail.com

Ващенко Андрей Александрович — кандидат физико-математических наук, заведующий отделом, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, [sc 35800121600](https://orcid.org/0000-0003-2084-5900), <https://orcid.org/0000-0003-2084-5900>, vashchenkoa@lebedev.ru

Захарчук Иван Александрович — студент, Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация, [sc 57672815700](https://orcid.org/0000-0002-1502-6460), <https://orcid.org/0000-0002-1502-6460>, zakharchukia@yandex.ru

Дайбаге Даниил Саюзович — студент, Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация; ассистент, Московский политехнический университет, Москва, 107023, Российская Федерация; младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация, [sc 57673090900](https://orcid.org/0000-0003-1944-1546), <https://orcid.org/0000-0003-1944-1546>, daibagya@mail.ru

Амброзевич Сергей Александрович — кандидат физико-математических наук, старший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация; доцент, Московский государственный технический университет им. Н.Э. Баумана, Москва, 105005, Российская Федерация, [sc 12789274600](https://orcid.org/0000-0002-3906-0735), <https://orcid.org/0000-0002-3906-0735>, s.ambrozevich@mail.ru

Володин Никита Юрьевич — студент, Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация, [sc 57672535600](https://orcid.org/0000-0002-7801-9399), <https://orcid.org/0000-0002-7801-9399>, nikita9963@ya.ru

Чепцов Дмитрий Андреевич — кандидат химических наук, старший преподаватель, Российский химико-технологический универ-

Authors

Anna V. Osadchenko — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation; Assistant, Moscow Polytechnic University, Moscow, 107023, Russian Federation, [sc 57439684100](https://orcid.org/0000-0001-9556-4885), <https://orcid.org/0000-0001-9556-4885>, anna.vl.osadchenko@gmail.com

Andrey A. Vashchenko — PhD (Physics & Mathematics), Head of Department, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, [sc 35800121600](https://orcid.org/0000-0003-2084-5900), <https://orcid.org/0000-0003-2084-5900>, vashchenkoa@lebedev.ru

Ivan A. Zakharchuk — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation, [sc 57672815700](https://orcid.org/0000-0002-1502-6460), <https://orcid.org/0000-0002-1502-6460>, zakharchukia@yandex.ru

Daniil S. Daibagya — Student, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation; Assistant, Moscow Polytechnic University, Moscow, 107023, Russian Federation; Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation, [sc 57673090900](https://orcid.org/0000-0003-1944-1546), <https://orcid.org/0000-0003-1944-1546>, daibagya@mail.ru

Sergey A. Ambrozevich — PhD (Physics & Mathematics), Senior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation; Associate Professor, Bauman Moscow State Technical University, Moscow, 105005, Russian Federation, [sc 12789274600](https://orcid.org/0000-0002-3906-0735), <https://orcid.org/0000-0002-3906-0735>, s.ambrozevich@mail.ru

Nikita Yu. Volodin — Student, Mendelev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation, [sc 57672535600](https://orcid.org/0000-0002-7801-9399), <https://orcid.org/0000-0002-7801-9399>, nikita9963@ya.ru

Dmitry A. Cheptsov — PhD (Chemistry), Senior Lecturer, Mendelev University of Chemical Technology of Russia, Moscow, 125047, Russian

ситет им. Д.И. Менделеева, Москва, 125047, Российская Федерация, [sc 56711985600](https://orcid.org/0000-0002-9774-7922), <https://orcid.org/0000-0002-9774-7922>, dchepcov@yandex.ru

Долотов Сергей Михайлович — кандидат химических наук, учебный мастер, Российский химико-технологический университет им. Д.И. Менделеева, Москва, 125047, Российская Федерация, [sc 6602502141](https://orcid.org/0000-0002-0022-0535), <https://orcid.org/0000-0002-0022-0535>, dolsm@mail.ru

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

Авраменко Антон Игоревич — научный редактор, Всероссийский институт научной и технической информации Российской академии наук, Москва, 125190, Российская Федерация, [sc 57208522812](https://orcid.org/0000-0002-0374-6428), <https://orcid.org/0000-0002-0374-6428>, anton1905@internet.ru

Семенова Светлана Леонидовна — кандидат физико-математических наук, ведущий научный сотрудник, Всероссийский институт научной и технической информации Российской академии наук, Москва, 125190, Российская Федерация, [sc 55787344500](https://orcid.org/0000-0001-6899-1347), <https://orcid.org/0000-0001-6899-1347>, s.l.semenova@internet.ru

Селюков Александр Сергеевич — кандидат физико-математических наук, младший научный сотрудник, Физический институт им. П.Н. Лебедева РАН, Москва, 119991, Российская Федерация; старший преподаватель, Московский политехнический университет, Москва, 107023, Российская Федерация, [sc 55787344500](https://orcid.org/0000-0002-4007-6291), <https://orcid.org/0000-0002-4007-6291>, selyukov@lebedev.ru

Federation, [sc 56711985600](https://orcid.org/0000-0002-9774-7922), <https://orcid.org/0000-0002-9774-7922>, dchepcov@yandex.ru

Sergey M. Dolotov — PhD (Chemistry), Training Master, Mendeleev University of Chemical Technology of Russia, Moscow, 125047, Russian Federation, [sc 6602502141](https://orcid.org/0000-0002-0022-0535), <https://orcid.org/0000-0002-0022-0535>, dolsm@mail.ru

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

Anton I. Avramenko — Scientific Editor, Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences, Moscow, 125190, Russian Federation, [sc 57208522812](https://orcid.org/0000-0002-0374-6428), <https://orcid.org/0000-0002-0374-6428>, anton1905@internet.ru

Svetlana L. Semenova — PhD (Physics & Mathematics), Leading Researcher, Russian Institute for Scientific and Technical Information of the Russian Academy of Sciences, Moscow, 125190, Russian Federation, [sc 55787344500](https://orcid.org/0000-0001-6899-1347), <https://orcid.org/0000-0001-6899-1347>, s.l.semenova@internet.ru

Alexander S. Selyukov — PhD (Physics & Mathematics), Junior Researcher, Lebedev Physical Institute of the Russian Academy of Sciences, Moscow, 119991, Russian Federation; Senior Lecturer, Moscow Polytechnic University, Moscow, 107023, Russian Federation, [sc 55787344500](https://orcid.org/0000-0002-4007-6291), <https://orcid.org/0000-0002-4007-6291>, selyukov@lebedev.ru

Статья поступила в редакцию 18.08.2022

Одобрена после рецензирования 12.10.2022

Принята к печати 30.11.2022

Received 18.08.2022

Approved after reviewing 12.10.2022

Accepted 30.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1119-1126

Fabrication and characterization of hybrid composite of Al6082/SiC/rice husk powder using friction stir processing

Nitesh Kumar¹✉, Prateek Gupta², Rakesh Kumar Singh³

^{1,2,3} Noida Institute of Engineering and Technology, Greater Noida, 201310, India

¹ niteshsah214@gmail.com✉, <https://orcid.org/0000-0003-1302-1874>

² prateekgupta1911@gmail.com, <https://orcid.org/0000-0002-7021-2821>

³ rakesh.singh@niet.co.in, <https://orcid.org/0000-0002-9646-8147>

Abstract

Friction Stir Processing (FSP) is a technology in which the microstructure of a material, as well as the mechanical properties of this material, is enhanced by the action of friction with a special tool. There are many applications for FSP materials which are in high demand in various industries, including aerospace, shipbuilding, instrumentation, and many others. The paper presents a study of a hybrid composite made from silicon carbide powder and rice husk (RHP) as reinforcement, included in the Al6082 aluminum alloy. The study of the characteristics of the fabricated composite was performed using an optical microscope, field emission scanning electron microscope (FESEM), X-ray diffraction (XRD) method. Tensile strength (UTS) and hardness were determined. It was revealed by microscopy that a modification of the microstructure occurs on the surface of the composite. X-ray diffraction analysis showed the presence in the spectrum of elements similar to the Al/SiC/RHP hybrid composition. It is shown that UTS tensile strength and Brinell hardness is being increased by factors of 1.36 and 1.75, respectively, compared to base aluminum material.

Keywords

friction stir processing, hybrid composite, SiC, rice husk powder, tensile strength, hardness

For citation: Kumar N., Gupta P., Singh R.K. Fabrication and characterization of hybrid composite of Al6082/SiC/rice husk powder using friction stir processing. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1119–1126. doi: 10.17586/2226-1494-2022-22-6-1119-1126

УДК 66.011

Изготовление и характеристика гибридного композита Al6082/SiC/порошок рисовой шелухи, получаемого методом фрикционного перемешивания

Нитеш Кумар¹✉, Праatik Гупта², Ракеш Кумар Сингх³

^{1,2,3} Инженерно-технологический институт Нойды, Большая Нойда, 201310, Индия

¹ niteshsah214@gmail.com✉, <https://orcid.org/0000-0003-1302-1874>

² prateekgupta1911@gmail.com, <https://orcid.org/0000-0002-7021-2821>

³ rakesh.singh@niet.co.in, <https://orcid.org/0000-0002-9646-8147>

Аннотация

Обработка трением с перемешиванием (Friction Stir Processing, FSP) — технология, в которой микроструктура материала и его механические свойства усиливаются за счет воздействия трения специальным инструментом. Существует множество применений материалов, получаемых методом FSP, которые используются в различных отраслях промышленности, в том числе аэрокосмической, судостроении, приборостроении и многих других. В работе представлено исследование гибридного композита, изготавливаемого из порошков карбида кремния (SiC) и Rice Husk Powder (RHP) в качестве армирующего элемента, включаемого в алюминиевый сплав Al6082. Исследование характеристик изготовленного композита выполнено с применением оптического микроскопа, автоэмиссионного сканирующего электронного микроскопа рентгеновским дифракционным методом. Определены прочность на растяжение и твердость. Методом микроскопии выявлено, что на поверхности композита происходит модификация микроструктуры. Рентгеноструктурный анализ показал наличие в спектре

© Kumar N., Gupta P., Singh R.K., 2022

элементов, аналогичных гибридной композиции Al/SiC/RHP. Показано увеличение прочности на растяжение и твердости по Бриннелю в 1,36 и 1,75 раза соответственно по сравнению с основным алюминиевым материалом.

Ключевые слова

обработка трением с перемешиванием, гибридный композит, карбид кремния, порошок рисовой шелухи, предел прочности на растяжение, твердость

Ссылка для цитирования: Кумар Н., Гупта П., Сингх Р.К. Изготовление и характеристика гибридного композита Al6082/SiC/порошок рисовой шелухи, получаемого методом фрикционного перемешивания // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1119–1126 (на англ. яз.). doi: 10.17586/2226-1494-2022-22-6-1119-1126

Introduction

In the current scenario of the industry, every industrialist wants to use good material getting some superior quality mechanical properties on that material. There are various processes from which composite material is fabricated. For example, that processes are used in powder metallurgy, casting, diffusion bonding and much more. After fabrication, there is lot of waste as well as defects have also been produced, but out of these, Friction Stir Processing (FSP) has lower drawbacks while fabrication as compared to other processes. For this reason, FSP has been used to improve or modify their microstructure by solid-state plastic deformation which will help utilize them in different structural domains as well as in metallurgical industries. FSP came in 1991 after a few years of friction stir welding [1] which formed the basis of the manufacturing process. In FSP, pin of tool is inserted on the groove by applying axial force at a particular rpm. Plastic deformation has taken place when sufficient heat is generated and there is micro structural modification occurred by using different reinforcements on it. The schematic diagram of the FSP is shown in Fig. 1.

FSP can be done on various materials like magnesium, aluminum, copper, zinc, etc. At the same time, aluminum has a low weight, as well as good strength, that is why it is preferable for various applications due to having these properties [2]. For this research aluminum 6082 was chosen. To enhance the microstructure of the material by using FSP, some additive (called reinforcement) is added. Silicon Carbide (SiC) and Rice Husk Powder (RHP) are used as reinforcement.

SiC is ceramic nature type material which is a quite common material used in industry. There are various useful

properties of SiC material. For example, it has high thermal shock resistance. It has also high strength as well as high hardness and high wear resistance properties. For this reason, SiC was chosen as reinforcement for our research. Hybrid composites use many types of reinforcement in a single matrix intending to achieve a synergistic impact between the qualities of the reinforcements and the overall properties of the composite [3]. RHP is an agricultural waste material that is produced from rice mills. It is also burned after the crop grows and mixed in the air due to having tiny particles and forms various unwanted gases. These gases produce unwanted containment which is very dangerous for the lungs. This waste has created many health-related issues while mixing these particles in the air. But due to showing its mechanical properties, it can be the best while using a hybrid composite. It has enormous potential to increase strength as well as hardness [4]. It has also a wear resistance nature. So, by using these reinforcements, it will not only reduce the waste but also make an eco-friendly environment.

Literature Review

Fuse et al., [5], examined the effect of process parameters, i.e., number of passes, bobbin tool and rotational speed on Al 6064 when using FSP. The result revealed that with the use of bobbin tool at the third passes of FSP maximum hardness is achieved, i.e., 95.11 HV which is 28 times more than in the base material.

Maji et al., [6], examined the tool traverse speed, no. of passes, rotational speed, and plunge depth on Al 7075-T6 when using FSP. Results revealed that grain size is reduced to 5.69 from 32.34 μm at a lower traverse speed in the 2nd pass of processing.

Saravanakumar et al., [7], analyzed the effect of process parameters, i.e., rotational speed, transverse speed and axial force on Cu when using FSP. They revealed that 18 % volume reinforced rice husk ash has maximum wear resistance of $5.3 \cdot 10^{-5} \text{ mm}^3/\text{Nm}$ due to having an increase in hardness.

Parbhu et al., [8], reported the effect of the process parameters, i.e., tool material, tool rotation direction, tool rotation speed and feed rate on Al 6082 when using FSP. Results show that a composite abrasion resistance is increased twice compared to the base material while incorporating CaCO_3 .

Patil et al., [9], investigated the effect of the process parameters, i.e., downward axial force, rotational speed, tool traverse speed, and tilt angle on Al 7075 when using FSP. Authors concluded that higher micro hardness of 604 HV was achieved in linear lined holes pattern which

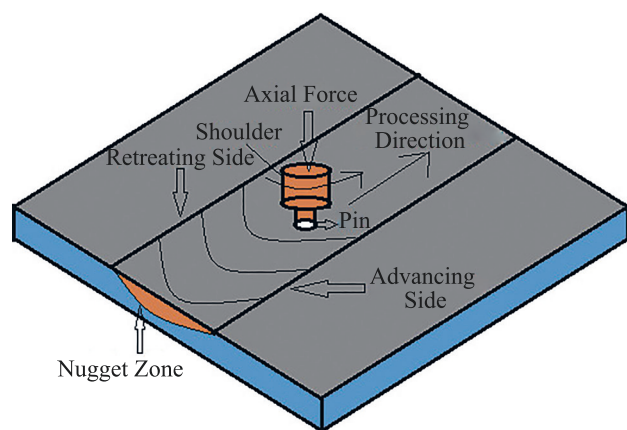


Fig. 1. Schematic diagram of friction stir processing

means it has superior hardness than the zigzag lined holes pattern and grooves pattern.

Liu et al., [10], examined the effect of process parameters, i.e., tool rotation speed, traverse speed, and tilt angle on Al 7075 when using FSP. Results revealed that with the addition of scandium its strength drops in comparison to the T6 original sample whereas the damping force increases with the addition of the scandium.

Luo et al., [11], investigated the effect of the process parameters, i.e., tool rotating direction, transverse direction, and multi-pass on AZ61 when using FSP. Results concluded that, when using a multipass FSP, grain size is refined to $7.8\ \mu\text{m}$ due to having dynamic recrystallization.

Nadammal et al., [12], reported the effect of process parameters, i.e., tool plunge depth, tool rotation rate, and traverse speed on Al 5086 when using FSP. Results revealed that because of the comparatively reduced heat input received and the resistance supplied by some of the shear texture components, the creation of such a weak texture can occur towards the bottom of the nugget zone.

By analysing these research papers, we can say that FSP finds application in various domains. These authors are generally used aluminum as a base material due to having a lower weight with good strength. So, in this research, we also use aluminum as a base material. But as a hybrid reinforcement, SiC (2 wt.%) and RHP (2, 4 and 6 wt.%) were used. This mixed hybrid composite will not only reduce the cost of material by the utilization of waste material RHP but also reduce environmental pollution. After fabrication, their different characterisation, like optical microscopic, X-ray Diffraction (XRD), tensile and hardness, were also examined. So, it will be beneficial for industrialists as well as for researchers to utilize the process in the best way.

Materials and Methods

For this fabrication, rectangular plate of Al 6082 with dimensions ($80 \times 95 \times 10\ \text{mm}$) is taken as a base material which has the following chemical composition: Aluminum (Al) — 95.2 wt.%; Silicon (Si) — 1.3 wt.%; Magnesium (Mg) — 1.2 wt.%; Manganese (Mn) — 1.0 wt.%; Iron (Fe) — 0.5 wt.%; Chromium (Cr) — 0.25 wt.%; Zinc (Zn) — 0.2 wt.%; Titanium (Ti) — 0.1 wt.%; Copper (Cu) — 0.1 wt.%.

For this reinforcement, SiC is taken which has extremely high thermal shock resistance. Used SiC reinforcement is shown in Fig. 2.

For the hybrid composite, another reinforcement is used, i.e., RHP. The Scanning Electron Microscope (SEM) image of RHP is illustrated in Fig. 3. **The average size of RHP is $58\ \mu\text{m}$.** From SEM image, the structure of the RHP is observed.

The chemical composition with their wt.% of RHP is identified with the help of the Energy dispersive X-ray (EDX) spectrometer. Its spectrum is illustrated in the Fig. 4. From EDX spectrum, maximum amount of elements observed in RHP are C and Ni followed by Cu, Mg, Ca and Zn.

Groove is prepared on the plate with the shaper machine, which is shown in Fig. 5. Dimension of the



Fig. 2. SiC reinforcement. The average size of the SiC particles is $150\ \mu\text{m}$

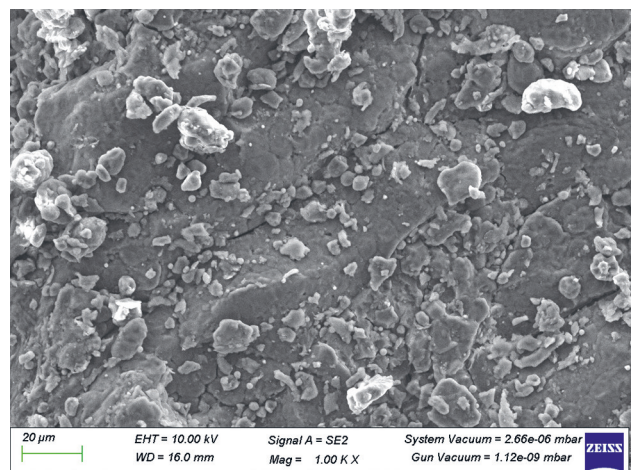


Fig. 3. SEM image of RHP. The average size of RHP is $58\ \mu\text{m}$

groove is ($6 \times 6\ \text{mm}$). Reinforcements (SiC and RHP) are inserted into the groove of the plate.

For FSP, threaded tool is taken. It is made up of high carbon and high chromium steel which was analyzed in the study of various literature reviews; there the information was found that tools made of these materials are suitable for the fabrication of aluminum alloys [13]. The diameter

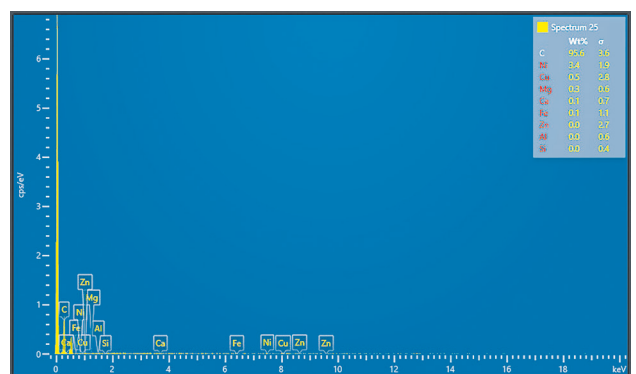


Fig. 4. EDX spectrum of the RHP



Fig. 5. Grooving on the sample

of the shoulder is 16 mm whereas the length of the pin is 8 mm and the diameter of the pin is 6 mm (it is shown in Fig. 6).

Process parameters have a vital role in the FSP. There are various process parameters like rotation speed, rotation direction, traverse speed, tilt angle, pin profile, shoulder

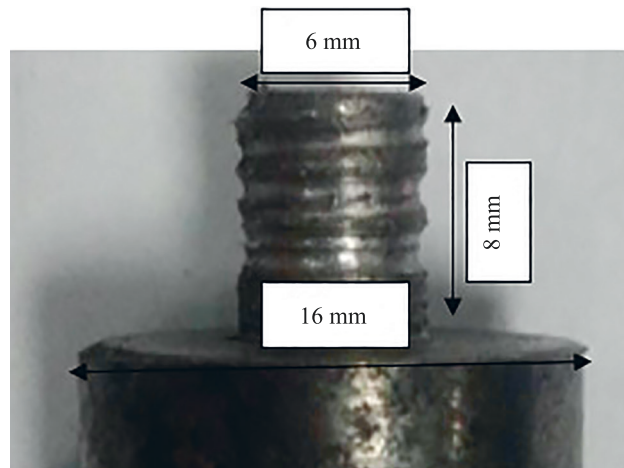


Fig. 6. Threaded tool pin profile

diameter, and much more. Out of these, some process parameters have been adopted, for example, the rotation speed is taken equal to 1500 rpm, the traverse speed is taken equal to 50 mm/min and the tilt angle is taken equal to 3° to show the effect on the material.

Experimentation

FSP was implemented on the vertical milling machine. The set-up of the FSP for fabrication of hybrid composite AA6082/SiC/RHP is illustrated in Fig. 7. Using a threaded tool, the fabrication was completed. The tool has been inserted into the collet. The tool shoulder contacts the workpiece when an axial downward force is applied, while the pin is put into the sample groove **on which reinforcements are inserted on it**. While the shoulder comes into contact, workpiece with revolving at 1500 rpm speed creates friction, and then plastic deformation takes



Fig. 7. FSP set-up machine

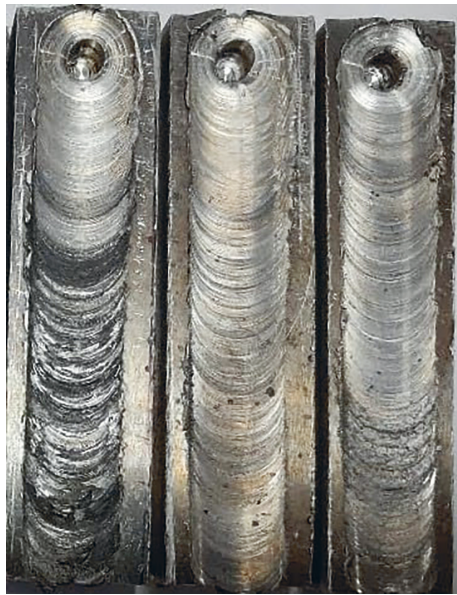


Fig. 8. Fabricated hybrid composite AA6082/SiC/RHP

place on **FSP Zone**. The traverse speed, on the other hand, aids in the movement of the workpiece in the processing direction as well as the deposition of material from back to forth. This process is continuously taking place after the processing tool is removed by the removal of axial force, and the power supply is off.

A fabricated sample of hybrid composite AA6082/SiC/RHP is shown in Fig. 8.

Results and Discussions

Microstructural Analysis

An optical microscope is used to show the macrostructural change in the material. All the fabricated

hybrid composite of AA6082/SiC/RHP is examined under the microscope on the composite sample with the size 10×10 mm (Fig. 9). Before examining, the sample is rubbed on the emery paper of the grit no. 400, 600, 800, 1200 and 1500. Then the rubbed sample is polished with the chemical powder Al_2O_3 . This chemical powder is used to better surface finish as well as to clear structure. The polished sample is also etched with HNO_3 solution to enhance the contrast on the surface so that the microstructure can be seen easily.

Fig. 9, *a* shows the microscopic view of base material without fabrication in which no defect on the material found. This is confirmed by the examining different sides under an optical microscope and it was found that there is not any porosity as well as available cracks on unprocessed workplace. Fig. 9, *b* shows the fabricated composite microscopic image in which 2 wt.% of SiC as well as 2 wt.% of powder of rice husk reinforcement is mixed. Fig. 9, *c* shows the microscopic image of fabricated composite on which 2 % SiC and 4 % powder of rice husk is mixed as a wt. percentage of the base material. Similarly, Fig. 9, *d* shows the fabricated composite microscopic image on which reinforcement 2 % SiC and 6 % RHP are incorporated with wt. percentage of the base material. It can be seen from these analyzed microscopic images that a superior microstructure is achieved in the sample on which 2 % SiC and 4 % powder of rice husk are mixed. There takes place a proper mixture between the reinforcement and the base material. Due to plastic deformation, good dynamic recrystallization takes place with threaded tools creating a better microstructure than in the base material. Due to medium rpm, i.e., 1500 rpm, there is sufficient heat generated which helps microns reinforcements to mix properly. Similarly, tilt angle of 3° also helps to get sufficient friction produced between shoulder of tool and surface of workpiece. As a result, process parameters are also affected the microstructure of the composite.

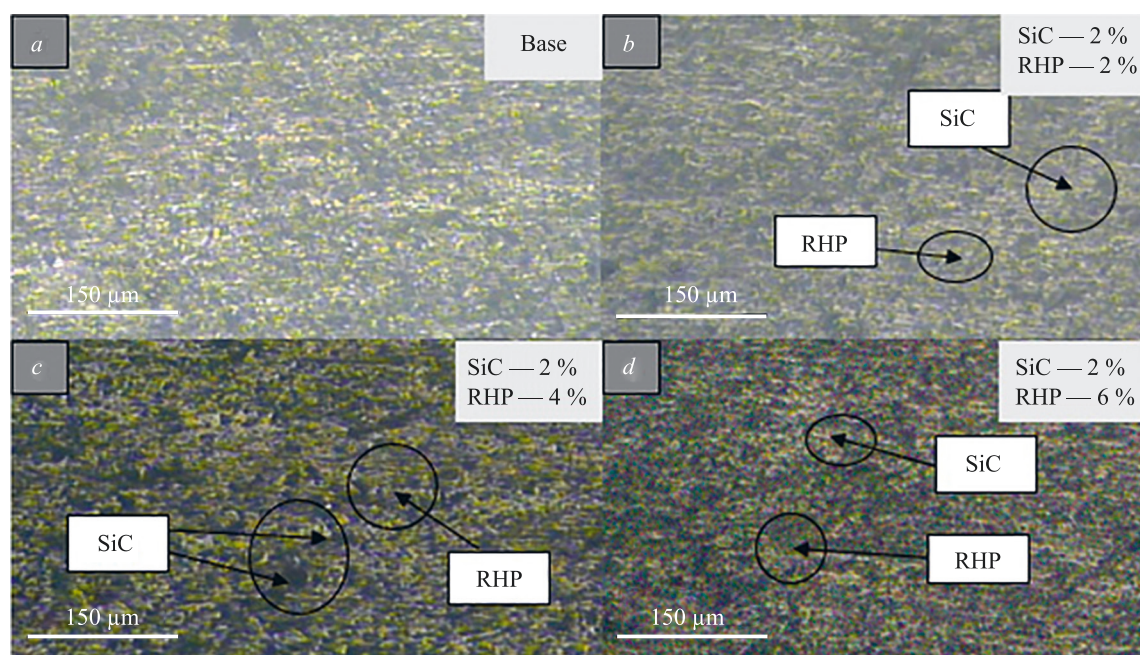


Fig. 9. Microstructure of: base material AA6082 (*a*); AA6082/SiC/2 % RHP (*b*); AA6082/SiC/4 % RHP (*c*); AA6082/SiC/6 % RHP (*d*)

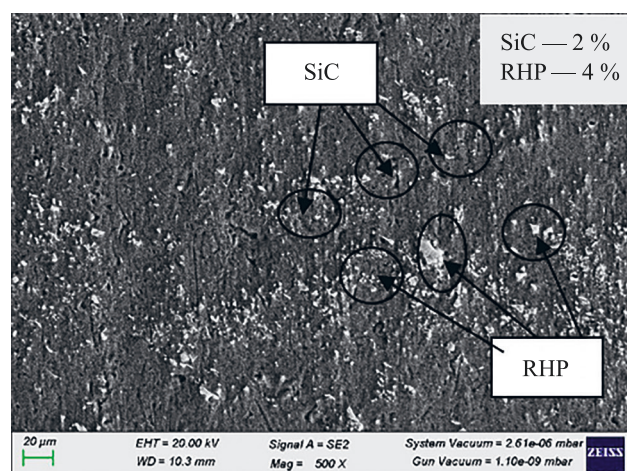


Fig. 10. FESEM of fabricated AA6082/SiC/RHP hybrid composite

To get further analysis, Field Emission Scanning Electron Microscope (FESEM) is also used for that fabricated composite to get more information on the microstructure (Fig. 10).

From FESEM analysis, it was observed that the correct distribution of the reinforcement, i.e., SiC, as well as the RHP on the fabricated composite, was received. These microns reinforcements are identified by the showing brighter particles of reinforcements distributed over the aluminium metal matrix. Sufficient plastic deformation occurred at a proper tilt angle which helps the better distribution of particles due to sufficient heat generated between the tool shoulder and workpiece. As a result, a fabricated hybrid composite with modified microstructure is obtained. To get elemental analysis information an EDX spectrum is also used which is illustrated in Fig. 11.

EDX spectrum of the fabricated AA6082/SiC/RHP hybrid composite gives the detailed information about elements used with their wt. percentage on that composite. Maximum amount of C, Si, and Mg is found followed by the Fe, Zn, and Ni on the aluminum matrix. So, the EDX spectrum can confirm that there is a presence of elements on the aluminum matrix which are also observed on the EDX spectrum of RHP. These elements are very useful for strengthening the material as well as to increase the hardness of a material. Also EDX spectrum confirms that

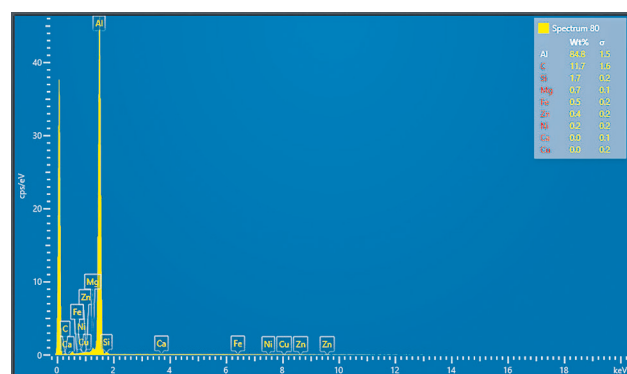


Fig. 11. EDX spectrum of fabricated AA6082/SiC/RHP hybrid composite

there is proper utilization of reinforcement elements on the metal matrix.

XRD analysis

XRD is used to detect the phase of crystalline structure and also detect the various compound formations from the presence of elements. XRD characterization is done from the IIT Roorkee Lab, India. XRD is done by using Bruker D-8 advanced diffractometer with CuK radiation. The samples were scanned at 2 kcps with a scanning speed of 12 °/min while rotating the MiniFlex (300/600) goniometer in the 2θ range of 5–90°.

The XRD result in Fig. 12 shows that the compound formation has taken place from the same element as the elemental composition of base material AA6082. There also got information that no harmful effects present for machining compound formation.

Fig. 13 shows the XRD result of the fabricated hybrid composite of AA6082/SiC/RHP. The result of the XRD revealed that different compounds and elements are observed on the different peaks. A compound like SiO₂ is hard in nature as well as it has also beneficial to improve the hardness of the composite. Similarly, there is also observed MgO which is physically and chemically stable at high temperatures. The compound, as well as elements which are observed on the peak of the fabricated hybrid composite AA6082/SiC/RHP, were very useful in enhancing the mechanical properties.

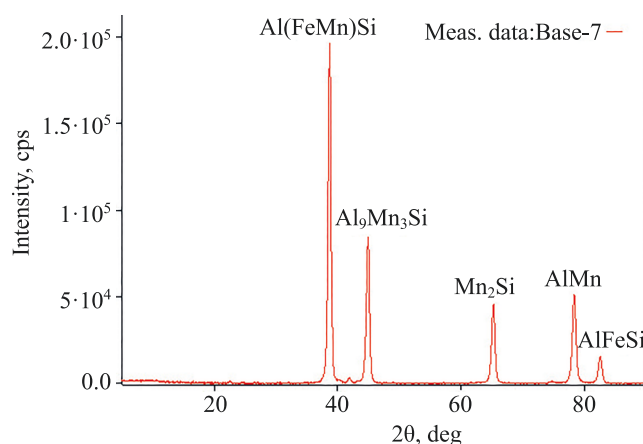


Fig. 12. XRD of base material AA6082

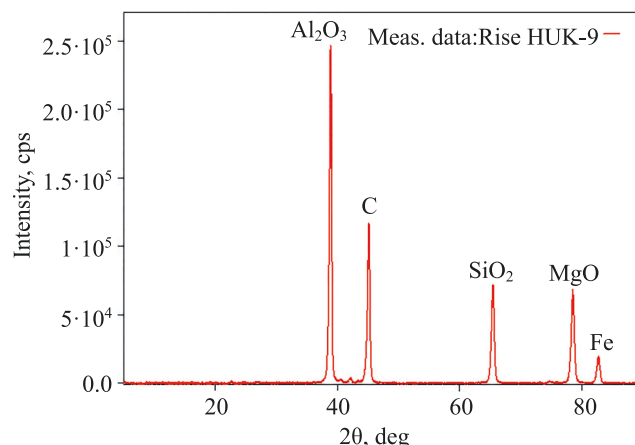


Fig. 13. XRD of AA6082/SiC/RHP

Ultimate Tensile Strength

After fabrication, ultimate tensile strength is also measured with the help of a tensometer at NIET Greater Noida, India material science lab. From the tensometer it was found that the fabricated composite has the following ultimate tensile strength which is shown in Fig. 14.

The ultimate tensile strength of the fabricated composite shows different data as the varying reinforcement wt. %. The maximum ultimate tensile strength is achieved in the sample on which reinforced 2 % SiC whereas the 4 % powder of rice husk. The used reinforcement increases the material strength up to a point of 4 wt. % of RHP, but beyond that, it decreases its value when utilising 6 wt. % of RHP.

Due to the agglomeration of **reinforcement**, the tensile strength has decreased at 6 wt. % but, while increasing reinforcement up to 4 %, the proper mixing happens and sufficient plastic deformation has occurred as a resulting tensile strength is increased. The tensile strength is found to be higher than the base material, also due to the used suitable process parameter. This contributes to the flow of the material, and mixing with reinforcement particles at 1500 rpm is very appropriate. Dinaharan et al., [14], showed a similar result, the tensile strength is increased, while mixing the rice husk as reinforcement in the Al 6061 alloy, because the composite has more grain refinement than the aluminium matrix, so the tensile strength is increased according to the Hall-Petch equation.

Hardness

Hardness is also measured on the Brinell Hardness testing machine at the material science lab NIET Greater Noida, India. Hardness is measured at five points on the fabricated composite and we take an average of them to achieve good accuracy. The hardness of that fabricated composite is shown in Fig. 15.

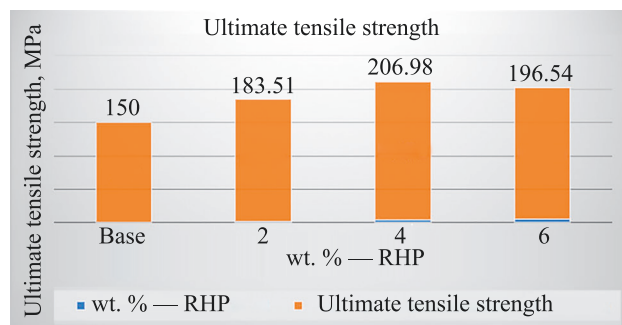


Fig. 14. Ultimate tensile strength of fabricated hybrid composite AA6082/SiC/RHP

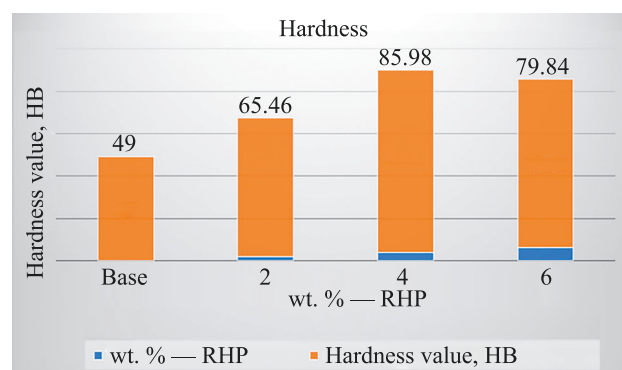


Fig. 15. Hardness of fabricated hybrid composite AA6082/SiC/RHP

The hardness of the fabricated composite shows increasing order up to a certain limit; after that it is decreased which is shown in Fig. 15. The maximum hardness achieved is 85.98 HB on which reinforcement of 2 % SiC as well as 4 % RHP is mixed as a wt. % on Al 6082 alloy. The hardness is drastically increased due to the threaded tool creating proper mixing with of powder of rice husk and SiC, both have good bond formation while mixing with aluminum. Fatchurrohman et al., [15], also showed similar result due to the inclusion of tiny eutectics particles by the mixing of RHP in the main aluminium matrix which also has resulted in increased hardness.

Conclusions

A hybrid composite of AA6082/SiC/RHP is successfully fabricated with help of FSP. Fabricated composite has also been tested by different characterization to show their bonding as well as their mechanical properties. The optical microscope and FESEM show that a modified microstructure is obtained compared to the base material, whereas the EDX spectrum shows the presence of elements that are similar to those found in EDX spectrum of RHP on the aluminium matrix. Similarly, from the XRD result it was found that the observed compounds as well as elements on the peak are helpful in nature and also improve the nature of composites. There is also ultimate tensile strength as well as the hardness are increased 1.36 times and 1.75 times, respectively, compared to the base material. So the fabricated hybrid composite has superior microstructure as well as improved mechanical strength and is also useful for various applications.

References

1. Thomas W.M., Nicholas E.D., Needham J.C., Murch M.G., Temple-Smith P., Dawes C.J. Friction stir butt welding. *International Patent Application PCT/GB92/02203*, 1991.
2. Dhayalan R., Kalaiselvan K., Sathiskumar R. Characterization of AA6063/SiC-Gr surface composites produced by FSP technique. *Procedia Engineering*, 2014, vol. 97, pp. 625–631. <https://doi.org/10.1016/j.proeng.2014.12.291>
3. Shahzad A. Investigation into fatigue strength of natural/synthetic fiber-based composite materials. *Mechanical and Physical Testing of Biocomposites, Fibre-Reinforced Composites and Hybrid Composites*, 2019, pp. 215–239. <https://doi.org/10.1016/B978-0-08-102292-4.00012-6>

Литература

1. Thomas W.M., Nicholas E.D., Needham J.C., Murch M.G., Temple-Smith P., Dawes C.J. Friction stir butt welding. *International Patent Application PCT/GB92/02203*, 1991.
2. Dhayalan R., Kalaiselvan K., Sathiskumar R. Characterization of AA6063/SiC-Gr surface composites produced by FSP technique // *Procedia Engineering*. 2014. V. 97. P. 625–631. <https://doi.org/10.1016/j.proeng.2014.12.291>
3. Shahzad A. Investigation into fatigue strength of natural/synthetic fiber-based composite materials // *Mechanical and Physical Testing of Biocomposites, Fibre-Reinforced Composites and Hybrid Composites*. 2019. P. 215–239. <https://doi.org/10.1016/B978-0-08-102292-4.00012-6>

4. Yadav R., Dwivedi V.K., Dwivedi S.P. Eggshell and rice husk ash utilization as reinforcement in development of composite material: A review. *Materials Today: Proceedings*, 2021, vol. 43, pp. 426–433. <https://doi.org/10.1016/j.matpr.2020.11.717>
5. Fuse K., Badheka V., Patel V., Andersson J. Dual sided composite formation in Al 6061/B4C using novel bobbin tool friction stir processing. *Journal of Materials Research and Technology*, 2021, vol. 13, pp. 1709–1721. <https://doi.org/10.1016/j.jmrt.2021.05.079>
6. Maji P., Nath R.K., Paul P., Bhogendra Meitei R.K., Ghosh S.K. Effect of processing speed on wear and corrosion behavior of novel MoS₂ and CeO₂ reinforced hybrid aluminum matrix composites fabricated by friction stir processing. *Journal of Manufacturing Processes*, 2021, vol. 69, pp. 1–11. <https://doi.org/10.1016/j.jmapro.2021.07.032>
7. Saravanakumar S., Gopalakrishnan S., Kalaiselvan K., Prakash K.B. Microstructure and mechanical properties of Cu/RHA composites fabricated by friction stir processing. *Materials Today: Proceedings*, 2021, vol. 45, pp. 879–883. <https://doi.org/10.1016/j.matpr.2020.02.935>
8. Sivanesh Prabhu M., Elaya Perumal A., Arulvel S., Franklin Issac R. Friction and wear measurements of friction stir processed aluminium alloy 6082/CaCO₃ composite. *Measurement*, 2019, vol. 142, pp. 10–20. <https://doi.org/10.1016/j.measurement.2019.04.061>
9. Patil N.A., Safwan A., Pedapati S.R. Effect of deposition methods on microstructure and mechanical properties of Al 7075 alloy-rice husk ash surface composites using friction stir processing. *Materials Today: Proceedings*, 2020, vol. 29, pp. 143–148. <https://doi.org/10.1016/j.matpr.2020.05.639>
10. Liu C.Y., Zhang B., Ma Z.Y., Jiang H.J., Zhou W.B. Effect of Sc addition, friction stir processing, and T6 treatment on the damping and mechanical properties of 7055 Al alloy. *Journal of Alloys and Compounds*, 2019, vol. 772, pp. 775–781. <https://doi.org/10.1016/j.jallcom.2018.09.109>
11. Luo X.C., Zhang D.T., Zhang W.W., Qiu C., Chen D.L. Tensile properties of AZ61 magnesium alloy produced by multi-pass friction stir processing: Effect of sample orientation. *Materials Science and Engineering: A*, 2018, vol. 725, pp. 398–405. <https://doi.org/10.1016/j.msea.2018.04.017>
12. Nadammal N., Kailas S.V., Szpunar J., Suwas S. Development of microstructure and texture during single and multiple pass friction stir processing of a strain hardenable aluminium alloy. *Materials Characterization*, 2018, vol. 140, pp. 134–146. <https://doi.org/10.1016/j.matchar.2018.03.044>
13. Elangovan K., Balasubramanian V. Influences of pin profile and rotational speed of the tool on the formation of friction stir processing zone in AA2219 aluminium alloy. *Materials Science and Engineering: A*, 2007, vol. 459, no. 1–2, pp. 7–18. <https://doi.org/10.1016/j.msea.2006.12.124>
14. Dinaharan I., Kalaiselvan K., Murugan N. Influence of rice husk ash particles on microstructure and tensile behavior of AA6061 aluminum matrix composites produced using friction stir processing. *Composites Communications*, 2017, vol. 3, pp. 42–46. <https://doi.org/10.1016/j.coco.2017.02.001>
15. Fatchurrohman N., Farhana N., Marini C.D. Investigation on the effect of friction stir processing parameters on micro-structure and micro-hardness of rice husk ash reinforced Al6061 metal matrix composites. *IOP Conference Series: Materials Science and Engineering*, 2018, vol. 319, no. 1, pp. 012032. <https://doi.org/10.1088/1757-899X/319/1/012032>
4. Yadav R., Dwivedi V.K., Dwivedi S.P. Eggshell and rice husk ash utilization as reinforcement in development of composite material: A review // *Materials Today: Proceedings*. 2021. V. 43. P. 426–433. <https://doi.org/10.1016/j.matpr.2020.11.717>
5. Fuse K., Badheka V., Patel V., Andersson J. Dual sided composite formation in Al 6061/B4C using novel bobbin tool friction stir processing // *Journal of Materials Research and Technology*. 2021. V. 13. P. 1709–1721. <https://doi.org/10.1016/j.jmrt.2021.05.079>
6. Maji P., Nath R.K., Paul P., Bhogendra Meitei R.K., Ghosh S.K. Effect of processing speed on wear and corrosion behavior of novel MoS₂ and CeO₂ reinforced hybrid aluminum matrix composites fabricated by friction stir processing // *Journal of Manufacturing Processes*. 2021. V. 69. P. 1–11. <https://doi.org/10.1016/j.jmapro.2021.07.032>
7. Saravanakumar S., Gopalakrishnan S., Kalaiselvan K., Prakash K.B. Microstructure and mechanical properties of Cu/RHA composites fabricated by friction stir processing // *Materials Today: Proceedings*. 2021. V. 45. P. 879–883. <https://doi.org/10.1016/j.matpr.2020.02.935>
8. Sivanesh Prabhu M., Elaya Perumal A., Arulvel S., Franklin Issac R. Friction and wear measurements of friction stir processed aluminium alloy 6082/CaCO₃ composite // *Measurement*. 2019. V. 142. P. 10–20. <https://doi.org/10.1016/j.measurement.2019.04.061>
9. Patil N.A., Safwan A., Pedapati S.R. Effect of deposition methods on microstructure and mechanical properties of Al 7075 alloy-rice husk ash surface composites using friction stir processing // *Materials Today: Proceedings*. 2020. V. 29. P. 143–148. <https://doi.org/10.1016/j.matpr.2020.05.639>
10. Liu C.Y., Zhang B., Ma Z.Y., Jiang H.J., Zhou W.B. Effect of Sc addition, friction stir processing, and T6 treatment on the damping and mechanical properties of 7055 Al alloy // *Journal of Alloys and Compounds*. 2019. V. 772. P. 775–781. <https://doi.org/10.1016/j.jallcom.2018.09.109>
11. Luo X.C., Zhang D.T., Zhang W.W., Qiu C., Chen D.L. Tensile properties of AZ61 magnesium alloy produced by multi-pass friction stir processing: Effect of sample orientation // *Materials Science and Engineering: A*. 2018. V. 725. P. 398–405. <https://doi.org/10.1016/j.msea.2018.04.017>
12. Nadammal N., Kailas S.V., Szpunar J., Suwas S. Development of microstructure and texture during single and multiple pass friction stir processing of a strain hardenable aluminium alloy // *Materials Characterization*. 2018. V. 140. P. 134–146. <https://doi.org/10.1016/j.matchar.2018.03.044>
13. Elangovan K., Balasubramanian V. Influences of pin profile and rotational speed of the tool on the formation of friction stir processing zone in AA2219 aluminium alloy // *Materials Science and Engineering: A*. 2007. V. 459. N 1–2. P. 7–18. <https://doi.org/10.1016/j.msea.2006.12.124>
14. Dinaharan I., Kalaiselvan K., Murugan N. Influence of rice husk ash particles on microstructure and tensile behavior of AA6061 aluminum matrix composites produced using friction stir processing // *Composites Communications*. 2017. V. 3. P. 42–46. <https://doi.org/10.1016/j.coco.2017.02.001>
15. Fatchurrohman N., Farhana N., Marini C.D. Investigation on the effect of friction stir processing parameters on micro-structure and micro-hardness of rice husk ash reinforced Al6061 metal matrix composites // *IOP Conference Series: Materials Science and Engineering*. 2018. V. 319. N 1. P. 012032. <https://doi.org/10.1088/1757-899X/319/1/012032>

Authors

Nitesh Kumar — M. Tech., Student, Noida Institute of Engineering and Technology, Greater Noida, 201310, India, <https://orcid.org/0000-0003-1302-1874>, niteshsah214@gmail.com

Prateek Gupta — PhD, Assistant Professor, Noida Institute of Engineering and Technology, Greater Noida, 201310, India, <https://orcid.org/0000-0002-7021-2821>, prateekgupta1911@gmail.com

Rakesh Kumar Singh — PhD, Assistant Professor, Noida Institute of Engineering and Technology, Greater Noida, 201310, India, <https://orcid.org/0000-0002-9646-8147>, rakesh.singh@niet.co.in

Авторы

Кумар Нитеш — магистр, студент, Инженерно-технологический институт Нойды, Большая Нойда, 201310, Индия, <https://orcid.org/0000-0003-1302-1874>, niteshsah214@gmail.com

Гупта Пратик — PhD, доцент, Инженерно-технологический институт Нойды, Большая Нойда, 201310, Индия, <https://orcid.org/0000-0002-7021-2821>, prateekgupta1911@gmail.com

Сингх Ракеш Кумар — PhD, доцент, Инженерно-технологический институт Нойды, Большая Нойда, 201310, Индия, <https://orcid.org/0000-0002-9646-8147>, rakesh.singh@niet.co.in

КОМПЬЮТЕРНЫЕ СИСТЕМЫ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ COMPUTER SCIENCE

doi: 10.17586/2226-1494-2022-22-6-1127-1135

A multi-path secure routing for the detection of node capturing attack in wireless sensor network

Jayaraman Kolangiappan^{1✉}, Angamuthu Senthil Kumar²

¹ Department of Computer Science, Periyar University, Salem, 636011, India

² Department of Computer Science, Arignar Anna Government Arts College, Namakkal, 637002, India

¹ jkakshiya@gmail.com✉, <https://orcid.org/0000-0001-5093-5822>

² senthilkumarmca76@gmail.com, <https://orcid.org/0000-0001-5131-7428>

Abstract

Over the past few years, the devices in Wireless Sensor Networks (WSN) are growing exponentially due to the emergence of many sophisticated applications. This tremendous growth leads to serious security challenges, and the devices of WSN should be protected from various attacks. WSN can be configured dynamically without fixed infrastructure and the devices can be talked with one another in an ad-hoc manner. Due to the dynamic nature of WSN, routing is considered as the challenging task that should be performed efficiently with robust routing mechanism. Even though many routing schemes have been emerged for WSN, they are not well scalable in very large-scale environment. This work introduces multi path routing strategy, and the routing will be selected based on trusted nodes. First, the trusted nodes are identified using trusted metrics of each node in the network. These metrics are calculated based on the threshold value of nodes. Then, secure routing is established by isolating node capturing attacks from the path. The performance of the work is analyzed in terms of packet loss, computational time and throughput. The paper compares the performance with the state-of-the-art routing schemes such as EMBTR (Enhanced Multi Attribute Based Attack Resistance), TSRM (Trust based secure routing model), and TARF (Trust-aware routing framework for WSNs). The outcome of the simulation shows that the proposed scheme outperforms the other state-of-the-work in terms of computational cost, throughput, and delay.

Keywords

routing, security, node capturing, WSN

For citation: Kolangiappan J., Senthil Kumar A. A multi-path secure routing for the detection of node capturing attack in wireless sensor network. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1127–1135. doi: 10.17586/2226-1494-2022-22-6-1127-1135

УДК 004.77

Многопутевая безопасная маршрутизация для обнаружения атаки с захватом узла в беспроводной сенсорной сети

Джаяраман Колангиаппан^{1✉}, А. Сентил Кумар²

¹ Университет Перияр, Салем, 636011, Индия

² Аригнар Анна Говермент Артс Колледж, Намаккал, 637002, Индия

¹ jkakshiya@gmail.com✉, <https://orcid.org/0000-0001-5093-5822>

² senthilkumarmca76@gmail.com, <https://orcid.org/0000-0001-5131-7428>

Аннотация

За последние несколько лет количество устройств, используемых в беспроводных сенсорных сетях (Wireless Sensor Networks, WSN), растет в геометрической прогрессии. Данный рост связан с появлением множества сложных приложений, что приводит к серьезным проблемам безопасности. Устройства в WSN должны быть защищены от воздействия различных атак. Сети можно настроить динамически без фиксированной инфраструктуры, а их устройства могут обмениваться данными друг с другом в режиме ad-hoc. Из-за динамической природы WSN-маршрутизация считается сложной задачей, которая должна эффективно

выполняться с помощью надежного механизма. Несмотря на то, что для таких сетей разработано множество схем маршрутизации, они плохо масштабируются. В работе представлена стратегия многопутевой маршрутизации. Маршрутизация выбрана на основе доверенных узлов, которые идентифицированы с использованием доверенных показателей каждого узла в сети. Метрики рассчитаны на основе порогового значения узлов и далее установлена безопасная маршрутизация с помощью изоляции от пути узла, перехватывающего атаки. Выполнен анализ производительности работы сети с точки зрения потери пакетов, времени вычислений и пропускной способности. Приведено сравнение производительности с современными схемами маршрутизации, такими как расширенная устойчивость к атакам на основе нескольких атрибутов (Enhanced Multi Attribute Based Attack Resistance, EMBTR), модель безопасной маршрутизации на основе доверия (Trust based Secure Routing Model, TSRM) и инфраструктура маршрутизации с поддержкой доверия для WSN (Trust-Aware Routing Framework for WSNs, TARF). Результат моделирования показал, что предложенная схема превосходит другие варианты с точки зрения вычислительных затрат, пропускной способности и задержки.

Ключевые слова

маршрутизация, безопасность, захват узла, WSN

Ссылка для цитирования: Колангианпан Д., Сентил Кумар А. Многопутевая безопасная маршрутизация для обнаружения атаки с захватом узла в беспроводной сенсорной сети // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1127–1135 (на англ. яз.). doi: 10.17586/2226-1494-2022-22-6-1127-1135

Introduction

Wireless Sensor Network (WSN) is a type of wireless network where group of sensors are deployed in any environment to sense the surroundings. The sensed data will be transferred to the recipient through intermediate nodes or Base Station (BS). Since the sensors are deployed in any environment, many security violations are possible including physical attacks. The applications of WSN are emerged in many real time applications such as surveillance, health care, agriculture and many areas. Most of the applications of WSN carry user's private data that needs to be protected from the attackers. Therefore, security is crucial for WSN application to face the challenges in current and future applications. The main focus of the paper is secure routing because multiple hops are possible for the WSN data to reach its destination. If any of the nodes in the path is compromised, then the security of the entire network is questionable [1–3]. Many routing attacks are possible during the routing process. Some of the attacks are DoS, packet forwarding, selective packet, wormhole, black hole, node capturing attacks, and much more [4–6]. DoS attack will shut the service of the whole system by sending frequent number of packets in the intention of flooding the bandwidth. Selective forwarding attack will not let the specific sender to reach the destination just by drop the packets selectively. In general, there are two classifications of attacks: active attack and passive attack. During the execution of active attack, the attacker will put some efforts to harm the confidential data of the user. In case of passive attack, attacker will not harm the system; however they will just monitor the communication channel in order to steal the user's credentials. This work focuses on node capturing attack since this attack is the combination of active attack and passive attack that could harm the environment more than any other attacks. This attack is considered as the most catastrophic one because the node captured by the invader will steal the cryptographic information of the network and cause malicious function inside the network. Thus, the confidentiality and integrity of the system cannot be preserved. The invader can do variety of malfunction to damage the security of the system [7]. First, it turns a normal node into a malicious node.

Then it will gradually increase the number of malicious nodes in the network. The security mechanism for the system is categorized as low level and high-level security mechanisms. Low level attacks depend on the lower layers of networks. Network layer and other low-level layers should be protected against attack such as jamming attack and DoS attack. Session hijacking, caching, spoofing replay attacks are some of the attacks of higher layers. Despite WSN is challenging to variety of attacks, routing attack should be prevented to safeguard the private data of the user. The architecture of WSN consists of entity such as user, internet, BS, and sensor nodes. The deployment of sensor nodes has done as cluster-based technique. User can request the data from the sensor through internet. Request and response will be transferred to the intended node through many intermediate nodes or BS. Here each node can act as both router and data collector. The primary security requirement to safeguard the sensors in WSN is to protect the communication channel. Since the deployment of sensors can be in anywhere environment, many attacks in the network are possible during the communication of cluster nodes. This work considers node capturing attack to protect the WSN environment from this attack. Many works have been evolved for the prevention of node capturing from WSN. During the execution of node capturing, the adversary will modify the sensitive details of the captured node and redeploy it to the network. This captured node will act as a regular node and communicate with the other nodes in the cluster. All the nodes in the cluster will share the confidential information such as cryptographic key and other shared keys without the knowledge of malicious node present in the network. The further section of the paper is divided to four parts. In the first part, the work discusses the literature of various related works for the detection of node capturing attack. Second part introduces the proposed work and the discussion about the novelty of how it differs from other works. In the third part, the results of the proposed work are discussed. Finally, the work is concluded.

Related works

There are various trust related schemes were employed in the previous works for the security of routing in WSN.

In [8], authors introduced a new approach called program integrity verification for the detection of node captured attack. The proposed technique is embedded in the access point of the cluster in order to safeguard the content of program memory. Additionally, the content of the data to be communicated is encrypted with hash-based encryption scheme in order to avoid any modification during transit. Authors claim that the proposed work can efficiently elude the node capture attack. However, the proposed method depends on the strong encryption scheme which makes the key generation process inefficient that leads to increased time complexity. Authors in [9] proposed a replica detection technique where they compared the earlier schemes for replica detection. The performances of the schemes are compared in terms of time, cost, energy, and packet delivery rate. The authors conclude that RED detection method achieves better detection rate among all the schemes. However, location dependent schemes are not efficient because it incurs memory overhead to save the location related information. Lin & Wu, 2016 [10], proposed a novel method for enhancing the efficiency of node capturing detection. First, authors developed a model to identify the unique behavior of node capturing attack. Then, they designed a countermeasure to defense against this attack. The proposed Matrix based model is well efficient in terms of detection rate and time. However, the computational complexity degrades the performance of the work that leads to increased energy. Authors in [11] discuss the importance of eluding node capturing in WSN environment. The authors proposed a multi factor authentication scheme for the prevention of node capturing attack. The researchers first investigated the signature of the attack and the capabilities of adversaries performing node capturing attack. The work also classified the different types of node capturing technique executed by the adversaries and modeled each type to find its nature. Even though, the suggested framework can be useful for creating secure path, the work increases the time complexity in processing multi factor authentication. In [12], authors proposed trusted routing schemes based on trusted metrics. The work computed the trusted value for each node in the network. Then the nodes are selected in the communication path based on the trusted value stored in a node. This value is calculated from the behavior of each node during the communication in the network. The performance of the work was analyzed in terms of node's reliability rate, stability rate, and elapsed time. Authors claimed that the work outperforms the other related works with their Quality of Service metrics. However, there is no proper evident shown in the work about the practical complexity in selecting the best route among all the possible trusted routes. Authors in [13] proposed an efficient security scheme for the isolation of blackhole, wormhole, and gray-hole attack with a single solution. In this work, trustworthy path can be established based on the capability-based model introduced here. However, it is unclear how efficiently the proposed model is in protecting against all three attacks in a single solution.

From the literature studies, we came to understanding that most of the current model focused on eliminating the attacks without properly investigating about the

computational complexity in the detection process of their solutions. Each solution in the existing schemes is proportional to one another. We also learnt that creating trust relation between the nodes in the cluster make the routing process more secure. To make the detection process more efficient, we proposed a secure routing scheme with reduced energy, computational cost and delay.

The proposed secure routing scheme

The work proposed a trust based secure routing based on enhanced Ad-hoc On Demand Vector protocol. The main objective of the work is to elude node capturing with reduced time and cost. To achieve this, we use only few parameters for consideration during the computational work process. The main target of the attacker is network layer because this layer is responsible for addressing, and packet delivery. Many attacks are possible to redirect the packet into attacker's destination. To combat the node capturing efficiently and to provide a trusted route, the work introduces the detection methods such as taken based authentication, and threshold based trusted metrics.

Token based authentication

A token-based protocol is proposed for authentication of devices before they send data to other devices. A token is generated in a form of ticket which needs to be shown before accessing any device in the network. This token is created based on the information about the device such as device id, random key, and the access rights. If a device wants to communicate with other device in the cluster, both sender and receiver need to be mutually authenticated before start its transmission process. This protocol consists of two phases. In phase 1, token generation process takes place. In phase 2, application of AES_GCM based encryption is applied. The token will be generated based on the rights R_i , device identifier ID_i and Random number $RAND_i$. The random number used in the protocol is to protect the network from forgery. After the token is created, it is hashed and encrypted with public key.

Fig. 1 shows the architecture of the proposed token-based scheme. If device A wants to connect with device B, it will send the token to device B. The token consist of details such as device identifier ID_i , Rights that the device possess to access a resource ' r ' are hashed and encrypted with encryption key E_k using AES algorithm. After receiving the token from device A, device B will decrypt the token using public key which is made available publicly in the cluster. Then it will check the hash function for any modification. If the token is valid, then device A will be allowed to access device B. Similarly, device B will initiate the same process for verification. In this way both devices are mutually authenticated. For the encryption of token, AES_GCM based algorithm is used in the work. This algorithm supports both authentication and confidentiality.

The input of the algorithm is Starting Variable for generating pseudo random number, public key to be used for encryption, plain text and authentication field. The output of the algorithm is encrypted text. Starting Variable can be generated by a device during communication process to perform authenticated encryption. Authentication file is used to authenticate the address of the sender and receiver.

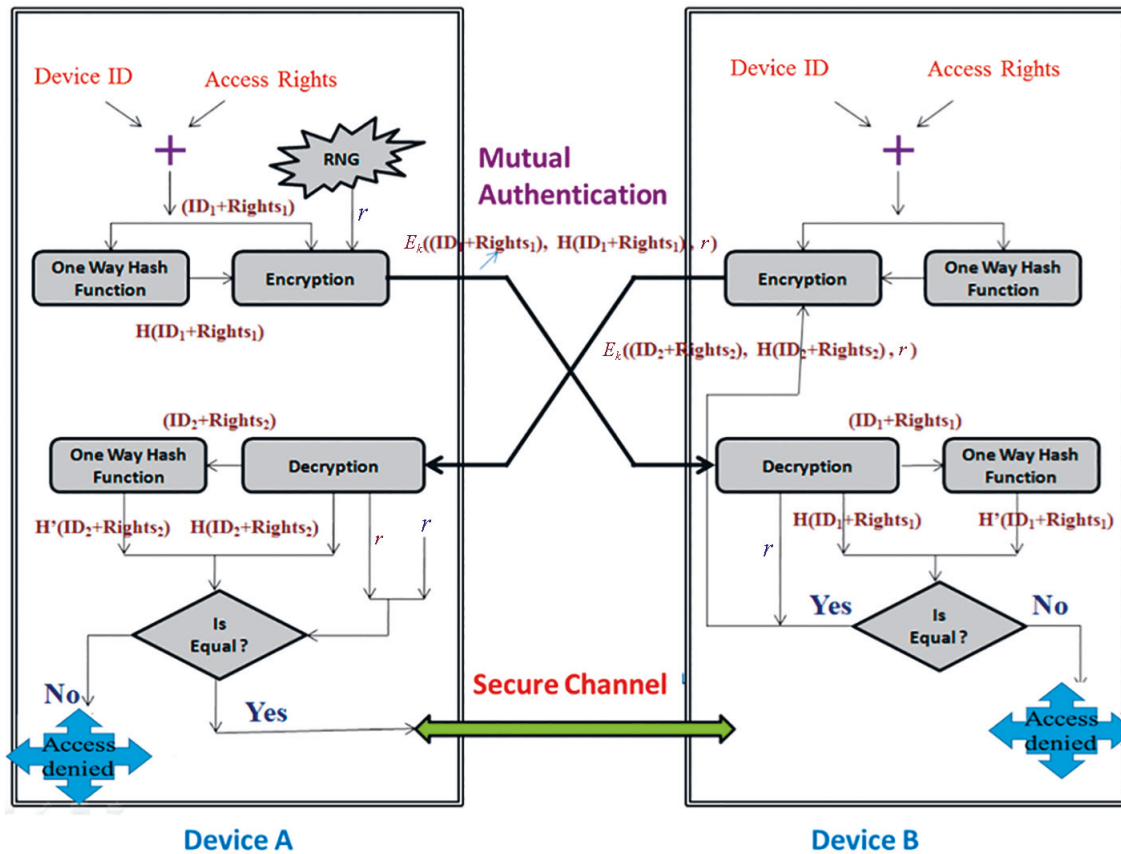


Fig. 1. The proposed Architecture

The token-based authentication protocol is designed based on the delegation model which consists of three devices such as Sensor device, Coordinator, and Gateway device. Coordinator node acts as a key distribution device, it establishes a secret key, then shares it with the other two devices such as Sensor node and the gateway node. There is an assumption that the Sensor and the Coordinator have a pre-established trust relationship and they share symmetric key K_S but the devices don't have direct communication. However, there is direct communication possible for the gateway device with both the Sensor and the Coordinator. The result of this protocol is a symmetric key to be shared between gateway and sensor.

The Sensor, the gateway, and the coordinator are denoted by S , G and C respectively in the following discussion.

The steps involved in the protocol specifications is given below:

1. $G \rightarrow S: \{\text{Nonce}_i, \text{ID}_i\}$

The protocol session is initiated by the gateway node by sending a message 1 to Sensor node S . The trust relationship is established by sharing a symmetric key K_S . G sends this message to S by generating Nonce_i with its identifier ID_i and send it to S in cleartext.

2. $S \rightarrow G: \{\text{Nonce}_i, \text{ID}_i, \text{Nonce}_j, \text{ID}_S, \text{ID}_C\}$

Sensor node S will not have any knowledge about gateway node G . However, it sends message 2 to Coordinator C to perform authentication and delegation.

This message consists of information about all the three entities, S , C , and G , and nonce values from S and G .

3. $G \rightarrow C: \{\text{Nonce}_i, \text{ID}_i, \text{Nonce}_j, \text{ID}_S, \text{ID}_C\}$

After the message 2 is received, G first verifies the message 1 with its Nonce_j and ID_G . It ensures duplicate identification has not occurred, i.e., ID_S , ID_G and ID_C , ID_G and ID_S , ID_C . After the successful verification, the gateway node G will forward message 3 to the coordinator node of S .

4. $C \rightarrow G: \{\text{Nonce}_i, \text{ID}_G, \text{Nonce}_j, \text{ID}_S, \text{ID}_C, \text{Nonce}_k, g^x\}$

When message 3 is received, C has to perform the following verification:

- Check the identification information received.
- Authenticate G .
- Sensor node S_i is fixed.
- G 's request for access on S can be granted.
- Authentication and authorization are based on G 's identity. Authorization can be done with additional information required in different situations. In that case, data can be validated by C and G through protocol exchange with the help of message 4, 5 and 6.

Later Nonce_3 can be established by S , D-H-key K_x and the data g^x to build message 4, which is similar to message 1 of SIGMA protocol.

The message generated here will comprise all the nonce Nonce_i in order to improve the three way communication. The gateway node G will prevent the chance of a DoS attack on C .

5. $G \rightarrow C: \{\text{Nonce}_1, \text{Nonce}_2, \text{Nonce}_3, g^y, S_G(g^x||g^y), [\text{ID}_S, \text{PUK}_G]K_S\}$

Here messages 4, 5 and 6 are used to authenticate C and G mutually. Initially G performs authentication by verifying nonce values and the value g_x received from message 3. Then message 5 is sent from the gateway node G , which is related to the SIGMA protocol in message 2, then MAC value is computed based on identity of G , nonce values, identifiers using its public key certificate PUK_G .

The key K_S can be obtained from the equations:

$$K_d = f(h(\text{Nonce}_1||\text{Nonce}_2||\text{Nonce}_3), g^{x,y}), \quad (1)$$

$$K_s = f(K_s, 1). \quad (2)$$

In the above equations (1) and (2), K_d and K_s are the decryption keys of any symmetric key, respectively. $f(h)$ represents the hash function to be generated. $g^{x,y}$ represent the data generated by gateway node for device X and device Y .

6. $C \rightarrow G: \{a, b, S_c(g^y||g^x), [\text{PUK}_C]K_S\}$

The following verifications are done by C

- All the nonce values, PUK_G , and ID_S verified against message 4.
- PUK_G is verified and authorize G for a key to be received by S .
- The signature in the PUK_G SG ($g^x||g^y$) is verified.
- The public value g_y is verified.
- K_S is used to verify the MAC on $[\text{ID}_S, \text{PUK}_G]K_S$.

In the above step, g^x, g^y are the data generated by gateway node for device X and device Y , K_x is the exchange of secret key and K_x is the symmetric key.

Threshold based resistance to node capturing

The threshold-based protocol is also introduced in the work to elude the node capturing from the route. The type of attack that we consider for evaluation is node capturing attack. It is assumed that all the nodes in the WSN are randomly distributed in the cluster, and communication can take place in both single hop and multi hop [14–16]. Behavior of each node is identified and a threshold value is fixed for each node in the network and the details are stored in BS based on their activities. This threshold value is fixed for a node from its sending behavior in a particular time gap. If any node exceeds the threshold, then it is identified as suspicious node. Then the algorithm will compute route analysis. During this analysis, the algorithm will check the distance of the suspicious node. If the distance of the node

is in one hop, then it is possible to detect the attack easily [17–20]. If the node is in multi hop distance, then it will check all the nodes in the network in route analysis process. If it is confirmed that the node sends more than allowed, then the node is considered as malicious node and it will be removed from the network. The main characteristics of the proposed work, its advantages and disadvantages are given in Table 1.

Results and discussion

The implementation of the proposed work is simulated in real time scenario. The simulation parameter is shown in Table 2. The performance of the work is evaluated in terms of energy, delay, and throughput under 2 different conditions.

- 1) WSN with node capturing under normal condition
- 2) WSN with node capturing under Mountain Safety Research (MSR), Trust-Aware Routing Framework for WSNs (TARF), Trust based Secure Routing Model (TSRM), and Enhanced Multi Attribute Based Attack Resistance (EMBTR).

There are two sets of simulation performed to evaluate the performance of the proposed work. In the first set, we did the evaluation under traffic intervals. The time of the interval is specified from 1s to 10s. Traffic interval under 1s is considered as fast and 10s is considered as slow.

Performance analysis under traffic interval in terms of energy, delay, and throughput

Fig. 2 shows the results of the work by considering energy, delay, and throughput under traffic interval from 1s to 10s. From the results, we observed that the proposed MSR enhances the energy, delay, and throughput by analyzing the network traffic and isolate the node capturing attempt. The important reason for the improved energy is that the MSR detects the node which performs node capturing and removes it from the network. Thereby saves the energy that may rise because of that malicious node. The reason for minimum delay is that the work can effectively detect the attack and isolate it from the network that gives the way for other nodes to continue their work without extra delay. Execution of this attack in one node can spread the same to many nodes in the network which will make the other nodes to wait for accessing the channel long time. After applying MSR, throughput is improved by reducing the waiting time of the nodes in the network under without attack condition.

Table 1. Summary of proposed work

Characteristics	Advantages	Disadvantages
Key pre distribution	Combination of secret key and public key	Sharing of key
Shared key discovery	Secured communication path through pairwise key	Broadcasting of messages related to shared key
Authenticated encryption	Supports both authentication and encryption	Software implementation is complex
Threshold based countermeasure	Easily detect the attacker node based on the threshold value stored in each node	Path analysis in multi hop network is difficult

Table 2. Simulation parameters

Parameter	Setting
Interface	Wireless channel
Antenna	Omni direction
Access control	MAC80.2.15.4
Routing	Ad-hoc on demand
Number of nodes	50
Power, mW	35
Energy level, J	100

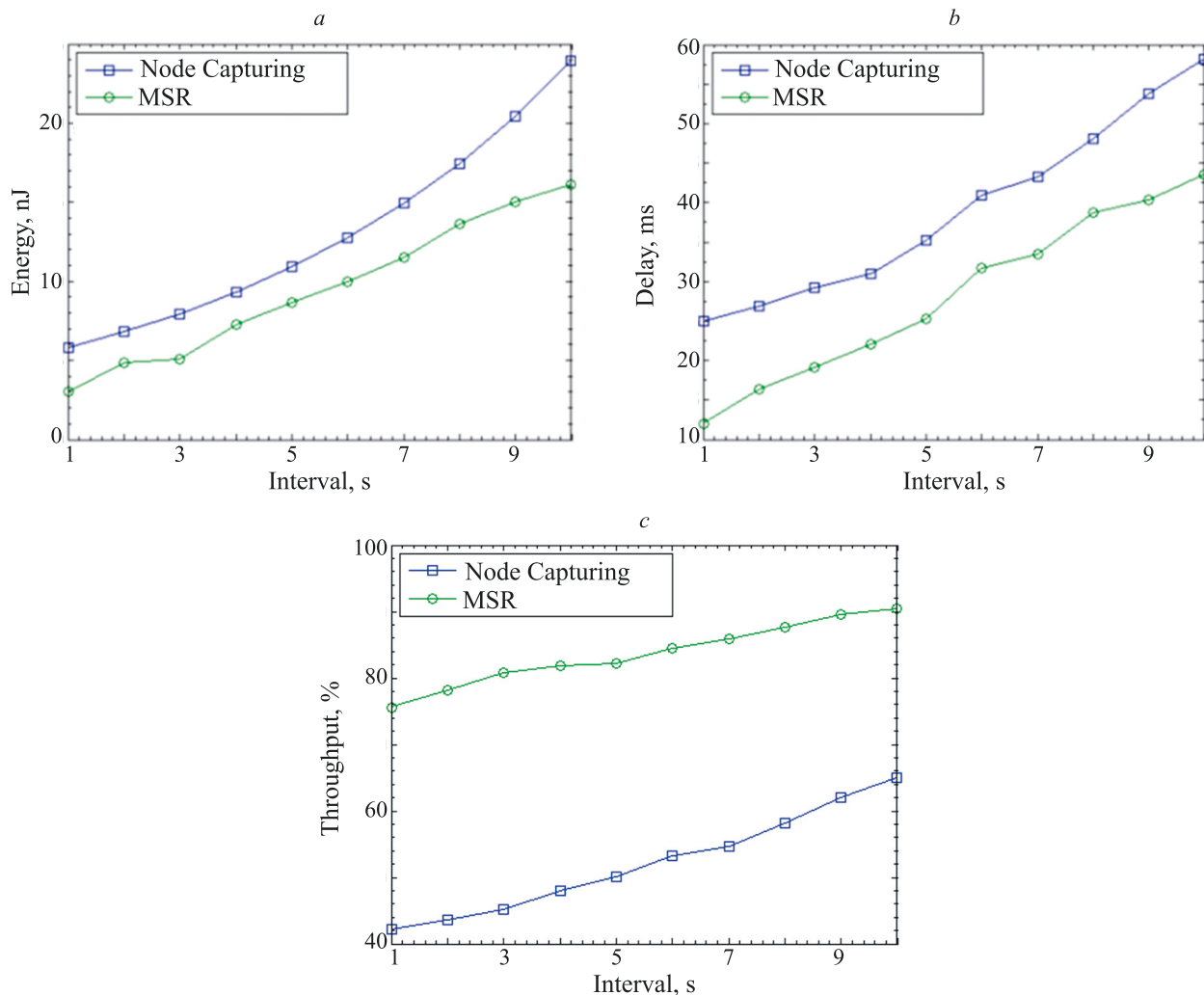


Fig. 2. Energy analysis (a), delay analysis (b) and throughput (c) in the interval of traffic with attack and MSR

Performance analysis with increased attacker nodes

Fig. 3 shows the outcome of the work by considering energy, delay, and throughput under increased malicious nodes. From the graph, we observed that the performance of the network has been significantly improved after applying MSR. This improvement is due to the efficient mechanism of the proposed work in defending the attack. The proposed work is effective against increased malicious nodes in the network by detecting multiple attacks that happened in the network. The work achieves reduced

energy than the other schemes such as TARF, TSRM, and EMBTR due to the removal of node capturing attacks by considering only trusted nodes in the network. The algorithm will select the trusted path based on the trusted metrics calculated for each node. All the trusted nodes in the network are selected as a route for data transmission which leads to reduction in energy. The work achieves better delay and throughput than the other schemes because of the reduction in waiting time and the response in sensing for node availability.

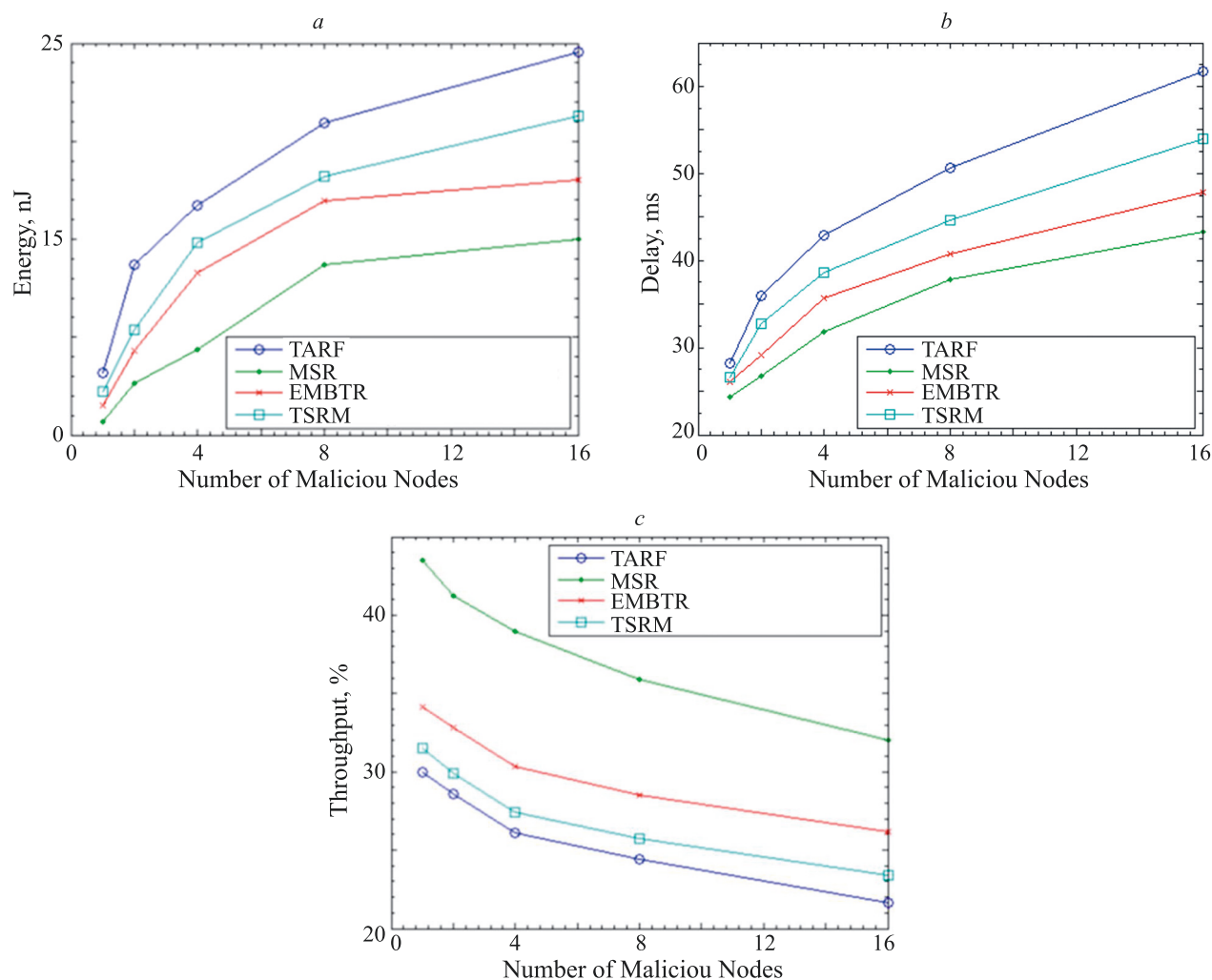


Fig. 3. Energy analysis (a), latency (b) and throughput (c) analyzes with increased number of malicious nodes

Conclusion

The work presents a multi path secure routing scheme MSR. Since the WSN nodes can be deployed in any environment such as military, hospital, banking and many more areas, there should be a strong security mechanism to protect the sensitive data. In our proposed work, a secure channel based on trusted nodes can be established for the transmission of data. Each node should possess a token for authentication before accessing any device. Only trusted nodes are participated during the communication process. This trusted node is identified based on the threshold value stored on each node in the network. Threshold value is the average number of data sent and received for a particular

node. The nodes can be in three states, such as normal state, suspicious state, and attack state, to identify the attack. If any of the nodes is confirmed as attacker node based on the trusted metrics, then it is removed from the network. The novelty of the work is its multi hop path analysis process which takes place once the suspicious node is confirmed. The performance of the work is compared with other works such as TARF, TSRM, and EMBTR. Simulation results prove the reliability of the proposed MSR scheme. According to the results obtained in the work, the model is proposed in terms of energy, delay, packet delivery rate, and false node detection rate. In future, we plan to investigate our work on different types of jamming attacks.

References

1. Albakri A., Harn L. Non-Interactive group key pre-distribution scheme (GKPS) for end-to-end routing in wireless sensor networks. *IEEE Access*, 2019, vol. 7, pp. 31615–31623. <https://doi.org/10.1109/ACCESS.2019.2900390>
2. Airehrour D., Gutierrez J.A., Ray S.K. SecTrust-RPL: A secure trust-aware RPL routing protocol for Internet of Things. *Future Generation Computer Systems*, 2019, vol. 93, pp. 860–876. <https://doi.org/10.1016/j.future.2018.03.021>

Литература

1. Albakri A., Harn L. Non-Interactive group key pre-distribution scheme (GKPS) for end-to-end routing in wireless sensor networks // *IEEE Access*. 2019. V. 7. P. 31615–31623. <https://doi.org/10.1109/ACCESS.2019.2900390>
2. Airehrour D., Gutierrez J.A., Ray S.K. SecTrust-RPL: A secure trust-aware RPL routing protocol for Internet of Things // *Future Generation Computer Systems*. 2019. V. 93. P. 860–876. <https://doi.org/10.1016/j.future.2018.03.021>

3. Ali R., Pal A.K., Kumari S., Karuppiiah M., Conti M. A secure user authentication and key-agreement scheme using wireless sensor networks for agriculture monitoring. *Future Generation Computer Systems*, 2018, vol. 84, pp. 200–215. <https://doi.org/10.1016/j.future.2017.06.018>
4. Han L., Zhou M., Jia W., Dalil Z., Xu X. Intrusion detection model of wireless sensor networks based on game theory and an autoregressive model. *Information Sciences*, 2019, vol. 476, pp. 491–504. <https://doi.org/10.1016/j.ins.2018.06.017>
5. Shajilin Loret J.B., Vijayalakshmi K. Security enrichment with trust multipath routing and key management approach in WMN. *IETE Journal of Research*, 2018, vol. 64, no. 5, pp. 709–721. <https://doi.org/10.1080/03772063.2017.1369365>
6. Usman A.B., Gutierrez J. DATM: A dynamic attribute trust model for efficient collaborative routing. *Annals of Operations Research*, 2019, vol. 277, no. 2, pp. 293–310. <https://doi.org/10.1007/s10479-018-2864-5>
7. Zawaideh F., Salamah M. An efficient weighted trust-based malicious node detection scheme for wireless sensor networks. *International Journal of Communication Systems*, 2019, vol. 32, no. 3, pp. e3878. <https://doi.org/10.1002/dac.3878>
8. Agrawal S., Das M.L., Lopez J. Detection of node capture attack in wireless sensor networks. *IEEE Systems Journal*, 2019, vol. 13, no. 1, pp. 238–247. <https://doi.org/10.1109/JSYST.2018.2863229>
9. Mishra A.K., Turuk A.K. A comparative analysis of node replica detection schemes in wireless sensor networks. *Journal of Network and Computer Applications*, 2016, vol. 61, pp. 21–32. <https://doi.org/10.1016/j.jnca.2015.12.001>
10. Lin C., Wu G. Enhancing the attacking efficiency of the node capture attack in WSN: a matrix approach. *Journal of Supercomputing*, 2013, vol. 66, no. 2, pp. 989–1007. <https://doi.org/10.1007/s11227-013-0965-0>
11. Wang C., Wang D., Tu Y., Xu G., Wang H. Understanding node capture attacks in user authentication schemes for wireless sensor networks. *IEEE Transactions on Dependable and Secure Computing*, 2022, vol. 19, no. 1, pp. 507–523. <https://doi.org/10.1109/TDSC.2020.2974220>
12. Khan F.A.B., Hannah L., Devi K.S., Rajalakshmi S. A multi-attribute based trusted routing for embedded devices in MANET-IoT. *Microprocessors and Microsystems*, 2022, vol. 89, pp. 10446. <https://doi.org/10.1016/j.micpro.2022.104446>
13. Vidhya Lakshmi G., Vaishnavi P. An efficient security framework for trusted and secure routing in MANET: A comprehensive solution. *Wireless Personal Communications*, 2022, vol. 124, no. 1, pp. 333–348. <https://doi.org/10.1007/s11277-021-09359-2>
14. Nikokheslat H.D., Ghaffari A. Protocol for controlling congestion in wireless sensor networks. *Wireless Personal Communications*, 2017, vol. 95, no. 3, pp. 3233–3251. <https://doi.org/10.1007/s11277-017-3992-y>
15. Jin X., Liang J., Tong W., Lu L., Li Z. Multi-agent trust-based intrusion detection scheme for wireless sensor networks. *Computers and Electrical Engineering*, 2017, vol. 59, pp. 262–273. <https://doi.org/10.1016/j.compeleceng.2017.04.013>
16. Wu F., Xu L., Kumari S., Li X. A new and secure authentication scheme for wireless sensor networks with formal proof. *Peer-to-Peer Networking and Applications*, 2017, vol. 10, no. 1, pp. 16–30. <https://doi.org/10.1007/s12083-015-0404-5>
17. Shin S., Kwon T., Jo G.-Y., Park Y., Rhy H. An experimental study of hierarchical intrusion detection for wireless industrial sensor networks. *IEEE Transactions on Industrial Informatics*, 2010, vol. 6, no. 4, pp. 744–757. <https://doi.org/10.1109/TII.2010.2051556>
18. Ayday E., Fekri F. An iterative algorithm for trust management and adversary detection for delay-tolerant networks. *IEEE Transactions on Mobile Computing*, 2012, vol. 11, no. 9, pp. 1514–1531. <https://doi.org/10.1109/TMC.2011.160>
19. Kamvar S., Schlosser M., Garcia-Molina H. The EigenTrust algorithm for reputation management in P2P networks. *Proc. of the 12th International Conference on World Wide Web*, 2003, pp. 640–651. <https://doi.org/10.1145/775152.775242>
20. Umar I.A., Hanapi Z.M., Sali A., Zulkarnain Z.A. TruFiX: A configurable trust-based cross-layer protocol for wireless sensor networks. *IEEE Access*, 2017, vol. 5, pp. 2550–2562. <https://doi.org/10.1109/ACCESS.2017.2672827>
3. Ali R., Pal A.K., Kumari S., Karuppiiah M., Conti M. A secure user authentication and key-agreement scheme using wireless sensor networks for agriculture monitoring // *Future Generation Computer Systems*. 2018. V. 84. P. 200–215. <https://doi.org/10.1016/j.future.2017.06.018>
4. Han L., Zhou M., Jia W., Dalil Z., Xu X. Intrusion detection model of wireless sensor networks based on game theory and an autoregressive model // *Information Sciences*. 2019. V. 476. P. 491–504. <https://doi.org/10.1016/j.ins.2018.06.017>
5. Shajilin Loret J.B., Vijayalakshmi K. Security enrichment with trust multipath routing and key management approach in WMN // *IETE Journal of Research*. 2018. V. 64. N 5. P. 709–721. <https://doi.org/10.1080/03772063.2017.1369365>
6. Usman A.B., Gutierrez J. DATM: A dynamic attribute trust model for efficient collaborative routing // *Annals of Operations Research*. 2019. V. 277. N 2. P. 293–310. <https://doi.org/10.1007/s10479-018-2864-5>
7. Zawaideh F., Salamah M. An efficient weighted trust-based malicious node detection scheme for wireless sensor networks // *International Journal of Communication Systems*. 2019. V. 32. N 3. P. e3878. <https://doi.org/10.1002/dac.3878>
8. Agrawal S., Das M.L., Lopez J. Detection of node capture attack in wireless sensor networks // *IEEE Systems Journal*. 2019. V. 13. N 1. P. 238–247. <https://doi.org/10.1109/JSYST.2018.2863229>
9. Mishra A.K., Turuk A.K. A comparative analysis of node replica detection schemes in wireless sensor networks // *Journal of Network and Computer Applications*. 2016. V. 61. P. 21–32. <https://doi.org/10.1016/j.jnca.2015.12.001>
10. Lin C., Wu G. Enhancing the attacking efficiency of the node capture attack in WSN: a matrix approach // *Journal of Supercomputing*. 2013. V. 66. N 2. P. 989–1007. <https://doi.org/10.1007/s11227-013-0965-0>
11. Wang C., Wang D., Tu Y., Xu G., Wang H. Understanding node capture attacks in user authentication schemes for wireless sensor networks // *IEEE Transactions on Dependable and Secure Computing*. 2022. V. 19. N 1. P. 507–523. <https://doi.org/10.1109/TDSC.2020.2974220>
12. Khan F.A.B., Hannah L., Devi K.S., Rajalakshmi S. A multi-attribute based trusted routing for embedded devices in MANET-IoT // *Microprocessors and Microsystems*. 2022. V. 89. P. 10446. <https://doi.org/10.1016/j.micpro.2022.104446>
13. Vidhya Lakshmi G., Vaishnavi P. An efficient security framework for trusted and secure routing in MANET: A comprehensive solution // *Wireless Personal Communications*. 2022. V. 124. N 1. P. 333–348. <https://doi.org/10.1007/s11277-021-09359-2>
14. Nikokheslat H.D., Ghaffari A. Protocol for controlling congestion in wireless sensor networks // *Wireless Personal Communications*. 2017. V. 95. N 3. P. 3233–3251. <https://doi.org/10.1007/s11277-017-3992-y>
15. Jin X., Liang J., Tong W., Lu L., Li Z. Multi-agent trust-based intrusion detection scheme for wireless sensor networks // *Computers and Electrical Engineering*. 2017. V. 59. P. 262–273. <https://doi.org/10.1016/j.compeleceng.2017.04.013>
16. Wu F., Xu L., Kumari S., Li X. A new and secure authentication scheme for wireless sensor networks with formal proof // *Peer-to-Peer Networking and Applications*. 2017. V. 10. N 1. P. 16–30. <https://doi.org/10.1007/s12083-015-0404-5>
17. Shin S., Kwon T., Jo G.-Y., Park Y., Rhy H. An experimental study of hierarchical intrusion detection for wireless industrial sensor networks // *IEEE Transactions on Industrial Informatics*. 2010. V. 6. N 4. P. 744–757. <https://doi.org/10.1109/TII.2010.2051556>
18. Ayday E., Fekri F. An iterative algorithm for trust management and adversary detection for delay-tolerant networks // *IEEE Transactions on Mobile Computing*. 2012. V. 11. N 9. P. 1514–1531. <https://doi.org/10.1109/TMC.2011.160>
19. Kamvar S., Schlosser M., Garcia-Molina H. The EigenTrust algorithm for reputation management in P2P networks // *Proc. of the 12th International Conference on World Wide Web*. 2003. P. 640–651. <https://doi.org/10.1145/775152.775242>
20. Umar I.A., Hanapi Z.M., Sali A., Zulkarnain Z.A. TruFiX: A configurable trust-based cross-layer protocol for wireless sensor networks // *IEEE Access*. 2017. V. 5. P. 2550–2562. <https://doi.org/10.1109/ACCESS.2017.2672827>

Authors

Jayaraman Kolangiappan — PhD, Research Scholar, Periyar University, Salem, 636011, India, <https://orcid.org/0000-0001-5093-5822>, jkakshiya@gmail.com

Angamuthu Senthil Kumar — PhD, Assistant Professor, Arignar Anna Government Arts College, Namakkal, 637002, India, [sc 57196435711](https://orcid.org/0000-0001-5131-7428), <https://orcid.org/0000-0001-5131-7428>, senthilkumarmca76@gmail.com

Received 15.06.2022

Approved after reviewing 27.09.2022

Accepted 09.11.2022

Авторы

Колангиаппан Джаяраман — PhD, исследователь, Университет Перияр, Салем, 636011, Индия, <https://orcid.org/0000-0001-5093-5822>, jkakshiya@gmail.com

Сентил Кумар Ангамуту — PhD, доцент, Аригнар Анна Говермент Артс Колледж, Намаккал, 637002, Индия, [sc 57196435711](https://orcid.org/0000-0001-5131-7428), <https://orcid.org/0000-0001-5131-7428>, senthilkumarmca76@gmail.com

Статья поступила в редакцию 15.06.2022

Одобрена после рецензирования 27.09.2022

Принята к печати 09.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1136-1142

УДК 004.2, 004.3, 004.383.8, 004.383.8.032.26, 004.4, 004.273

Метод документирования архитектурных решений вычислительных платформ

Ярослав Георгиевич Горбачев^{1,2}✉

¹ ООО ЛМТ, Санкт-Петербург, 199034, Российская Федерация

² Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

yaroslav-go@yandex.ru✉, <https://orcid.org/0000-0001-5419-6422>

Аннотация

Предмет исследования. Представлен метод описания вычислительных механизмов и документирования вычислительных платформ. Новизна метода заключается в использовании унифицированных средств для документирования на разных уровнях гранулярности аппаратного, программного и инструментального обеспечения, а также реконфигурируемых (интеллектуальных, адаптивных) компонент. Метод позволяет представлять в понятном виде вычислительные системы с нестандартными архитектурными решениями.

Метод. Предложенный метод заключается в описании идеальной модели вычислительной платформы с ее последующей итеративной детализацией. Особенность метода — наличие единого для различных систем ядра, которое используется для упрощения описаний и структурирования информации. Ядро включает в себя универсальные элементы и создано на основе анализа большого количества архитектур вычислительных систем. **Основные результаты.** С использованием предложенного метода описаны принципы организации широкого спектра вычислительных платформ. Рассмотрены следующие платформы: обобщенные процессоры с классической архитектурой, которая является развитием принципов фон-Неймана; программно-аппаратные системы на базе микроконтроллеров; операционные системы общего назначения; крупногранулярные и мелкогранулярные реконфигурируемые вычислительные системы; специализированные процессоры и ускорители; искусственные нейронные сети. **Практическая значимость.** Предложенный метод может быть использован для структурирования информации как по традиционным, так и по активно развивающимся направлениям: реконфигурируемым вычислительным системам и специализированным процессорам. На основе метода создана общая база универсальных вычислительных механизмов, пригодных для использования в разных узлах системы, для объектов разной гранулярности программными, аппаратными и иными средствами, на разной элементной базе. Результаты работы могут быть полезны системным архитекторам для документирования сложных вычислительных компонент, состоящих из программных, аппаратных и прочих механизмов. Метод направлен на упрощение повторного использования вычислительных механизмов и призван облегчить генерацию новых архитектурных решений. Также метод может быть полезен при обучении профильных специалистов, поскольку позволяет демонстрировать основные принципы вычислительной техники.

Ключевые слова

архитектура вычислительной системы, архитектурное описание, вычислительная платформа, вычислительный механизм, реконфигурируемая система, динамическая реконфигурация

Ссылка для цитирования: Горбачев Я.Г. Метод документирования архитектурных решений вычислительных платформ // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1136–1142. doi: 10.17586/2226-1494-2022-22-6-1136-1142

A method for documenting architectural solutions of computing platforms

Yaroslav G. Gorbachev^{1,2}✉

¹ LMT Ltd., Saint Petersburg, 199034, Russian Federation

² ITMO University, Saint Petersburg, 197101, Russian Federation

yaroslav-go@yandex.ru✉, <https://orcid.org/0000-0001-5419-6422>

© Горбачев Я.Г., 2022

Abstract

The article describes a method for documenting the principles of functioning and internal organization of computing platforms, including reconfigurable computing systems and non-standard processor architectures. The novelty is in using of unified tools to describe: the design process and the computing process, hardware, software and tools, computing components of different granularity. The proposed approach is to describe computing platform as an ideal model that represents abstract algorithms for fulfilling functional requirements without specifying of how to implement it. Then the iterative model refinement follows including selection of physical implementation options, specifying the technological stacks, and additional mechanisms that provide the specified system qualities. A feature of the method is a kernel used for structuring information, classifying and describing the computational mechanisms the system consists of. The kernel includes elements common for different systems and is based on the analysis of a large number of computing architectures. The method describes the principles of the organization of platforms which are usually not considered together. These are: generalized processors with classical architecture which is an evolution of the von Neumann principles; systems based on microcontrollers; operating systems; large- and small-granular reconfigurable systems; specialized processors and accelerators; artificial neural networks. The proposed method can be used to structure information in both traditional and rapidly developing areas: reconfigurable systems and specialized processors. Based on the method, it is possible to create a common database of computing mechanisms suitable for use in different functional units of the system and at different levels of granularity. The results of the work can be useful for system architects to describe complex computing mechanisms consisting of software, hardware and dynamically generated adaptive “intelligent” components which will simplify their reuse and can be used to generate new architectural solutions. Also, the proposed method can be used in the process of training specialists, for a visual demonstration of the basic principles of computer technology.

Keywords

architecture, architectural description, computing systems, computing mechanisms, reconfigurable systems

For citation: Gorbachev Ya.G. A method for documenting architectural solutions of computing platforms. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1136–1142 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1136-1142

Введение

Современная компьютерная индустрия сложна и разделена на самостоятельные, обособленно развивающиеся направления (например, разработка программного и аппаратного обеспечения). Как следствие, многие специалисты компетентны лишь в какой-то узкой области, плохо представляя, что происходит за ее границами.

Семантическая и технологическая разобщенность специалистов продолжает стремительно расти с развитием реконфигурируемых систем [1] и обилием оригинальных архитектур и концепций, например: NISC-систем [2], вычислений с памятью [3] и в памяти [4], процессоров функционального уровня [5], WARP-процессоров [6] и др. Каждое из нестандартных архитектурных решений требует уникального инструментария и компетенций разработчиков, обладает собственным набором характеристик (производительность, надежность, стоимость, сложность и т. д.).

Необходимость изучения нестандартных архитектур вычислительных систем связана с окончанием действия «Закона Мура» и поиском новых путей повышения эффективности вычислений, в том числе за счет преодоления проблемы «бутылочного горлышка» классических архитектур [4].

Одна из проблем — обмен информацией между сотрудниками об использованных в проектах архитектурных решениях и перенос смешанных программно-аппаратных решений между разными вычислительными платформами [7]. В этом случае общепринятое разделение на программное обеспечение (ПО) и аппаратуру неэффективно и может снизить производительность команды разработчиков.

В некоторых программно-аппаратных решениях отдельные действия, обычно выполняемые во время

разработки, смещаются в область функционирования (или наоборот). Классификации реконфигурируемых систем и систем, реализующих параллельные вычисления, часто оперируют понятием «гранулярность», которое может относиться как к разбиению алгоритмов на составляющие, так и к данным.

В связи с этим необходима разработка способа регистрации наиболее значимой информации о рассматриваемых архитектурах и вычислительных механизмах без искусственного разделения на программное, аппаратное и инструментальное обеспечения, на этапы проектирования и непосредственного исполнения для объектов разной гранулярности.

Если говорить об известных решениях, позволяющих описать принципы функционирования вычислительных систем, в первую очередь следует упомянуть понятие «модели вычислений» [8]. Модели вычислений определяют формальные правила и ограничения, описывающие поведение вычислительной системы или ее виртуального представления.

Для описания аппаратуры вычислительных систем на архитектурном уровне существует целый набор языков описания архитектуры [9]. К сожалению, они в основном узко специализированы и имеют значительные ограничения.

Довольно универсальные архитектурные описания доступны через создание высокоуровневых моделей в рамках методологии модельно-ориентированного проектирования [10]. Данная методология поддерживается рядом широко используемых в индустрии инструментов, таких как Simulink и LabVIEW. Отдельным направлением является генерация аппаратуры с помощью средств HLS [11].

Распространено описание ПО и вычислительных систем на основе языка UML [12] и его подмножества

SysML [13]. Эти языки широко поддерживаются разнообразными инструментальными средствами, например средой моделирования Capella и методом ARCADIA [14].

В слабо формализованной области программной инженерии, где разработчики столкнулись со схожими проблемами (большое количество различных сущностей, которые сложно классифицировать, сравнивать и описывать), используется стандарт Essence¹, созданный для описания и применения на практике методов программной инженерии.

Ни один из описанных подходов в полной мере не удовлетворяет заданным требованиям.

Описание модели

Отправной точкой для описания специализированных вычислительных архитектур и вычислительных механизмов предложено взять некую идеальную модель, аналог абсолютно черного тела и других подобных физических моделей. Модель должна быть пригодна для описания как можно большего количества разнообразных вычислительных механизмов и платформ, а также отображать наиболее значимые аспекты функционирования и жизненного цикла при максимальной простоте для восприятия. Сформулируем определение вычислительной платформы как набора программных и аппаратных средств (или стандарта их организации) для реализации прикладных вычислений, а вычислительного механизма — как упрощенного принципа функционирования и организации отдельного аспекта или элемента вычислительной платформы.

Для создания модели наиболее перспективен подход, предложенный в стандарте Essence. Подобно выделению общего ядра из универсальных элементов для методов и практик программной инженерии в Essence, в области вычислительной техники также можно выделить общее ядро — основу для описания широкого спектра вычислительных систем и механизмов. Ядро Essence включает в себя универсальные элементы и не зависит от конкретных реализаций и их практических применений. Отдельные элементы выделены из большого количества конкретных примеров, а все реализации описаны как надстройки поверх неизменного ядра.

Один из простейших способов описания объекта — его представление в виде «черного ящика», когда описывается взаимодействие объекта с внешним миром и реакции на внешние раздражители, а внутреннее устройство не раскрывается. Такой «черный ящик» принят за основу предложенной модели.

Система, рассматриваемая как «черный ящик», каким-либо образом производит преобразование подаваемой на вход энергии (соответствующие уровни которой могут интерпретироваться как данные) из некоторых входных значений в соответствующие им выходные. Кроме этого, в общем случае вычислительные (и не

только) системы могут иметь возможность сохранять состояние.

Для системы как «черного ящика» можно выделить следующие обобщенные действия или этапы функционирования, которые составляют основу ядра:

- 1) чтение входов;
- 2) вычисление;
- 3) сохранение внутреннего состояния;
- 4) запись выходов.

Отметим, что любая система обязательно должна выполнять хотя бы часть перечисленных действий. Описанное ядро — основа идеальной модели системы, на которую надстраивается ее более подробное описание.

Так как модель идеальная, считаем, что все действия производятся мгновенно, объемы блоков памяти бесконечны, и не требуется дополнительных действий, например обеспечения системы энергией, необходимой для штатного функционирования.

Примем, что переход от действия к действию идет сверху вниз с возможными возвратами назад и повторением отдельных итераций столько раз, сколько необходимо. Отметим, что в отдельных случаях может быть другой порядок действий.

При рассмотрении вычислителя в контексте жизненного цикла, необходимо учесть такие действия, как создание и уничтожение объекта. Контекст вычислительной техники делает оправданным выделение следующих этапов в создании объекта: создание реально существующей в физическом мире и фиксированной основы (аппаратного обеспечения) и ее программирование. В данном случае программирование необходимо понимать как программирование в широком смысле, включая создание и последующую запись в целевое устройство программ, настроек, конфигураций программируемой логики и т. д.

Отметим отличие между созданием физической основы и ее программированием. Создание физической основы производится только один раз, при создании системы, и закладывает неизменные на всех этапах жизненного цикла принципы (если оставить за скобками деградацию компонентов, помехи и т. д.). Программирование может (но не обязательно должно) производиться больше одного раза, явно и заметно изменяя свойства системы.

Обобщенно можно выделить следующие этапы жизненного цикла некоторой системы, подсистемы или элемента ядра. При необходимости они могут варьироваться и дополняться (приведенный список — лишь один из возможных вариантов представления):

- фиксация основных принципов функционирования системы;
- создание инструментария для разработки аппаратуры;
- разработка аппаратуры;
- создание инструментария для разработки ПО;
- разработка ПО;
- создание инструментария для загрузки ПО;
- загрузка ПО;
- инициализация системы;
- штатное функционирование;
- уничтожение.

¹ Essence — Kernel and Language for Software Engineering Methods (Essence), v1.2., 300 с. [Электронный ресурс]. Режим доступа: <https://www.omg.org/spec/Essence/1.2/PDF>, свободный. Яз. англ. (дата обращения: 13.11.2022).

Описанная модель применима для простых систем, или для небольших подсистем, входящих в сложные системы. Для поверхностного понимания архитектуры и особенностей функционирования сложной системы в целом ее необходимо представлять на более детальном уровне. Таким уровнем может стать представление системы как взаимодействующих «черных ящиков» — атомарных вычислителей, каждый из которых осуществляет какое-то простейшее действие в составе системы.

Тем самым можно получить естественную плавающую границу, позволяющую повышать или понижать уровень абстракции. Сначала в самом грубом приближении мы рассматриваем систему как «черный ящик»; когда этого недостаточно — как систему взаимодействующих «черных ящиков»; когда и этого недостаточно — каждый из «черных ящиков», составляющих систему, рассматриваем как систему взаимодействующих «черных ящиков», и так до тех пор, пока не достигается нужный уровень детализации.

Систему «черных ящиков» уже не описать простыми действиями 1–4, рассредоточенными по этапам жизненного цикла. Результаты вычислений системы как целого есть производная от результатов вычислений каждого из составляющих систему «черных ящиков» и зависит от того, как они взаимодействуют.

При переходе от представления системы как «черного ящика» к представлению системы как взаимодействующих «черных ящиков» появляется необходимость описать устройство обеспечивающей среды — то, каким образом вычислительные блоки, блоки памяти, входы и выходы объединяются в единое целое и благодаря чему они получают возможность функционировать вместе, решая некоторые общие задачи.

Обеспечивающая среда идеальной модели должна выполнять две функции: перемещать данные между вычислительными блоками, блоками памяти, входами и выходами, и синхронизировать вычисления отдельных вычислительных блоков.

На этом этапе логично выделить управления данными и вычислениями. При этом область управления вычислениями не включает в себя сами вычисления — она только выявляет причинно-следственные связи и обеспечивает передачу управления конкретным вычислителем, «черным ящиком», в которых и производятся вычисления.

Для получения описания модели всей системы необходимо описать взаимодействие трех областей: «черных ящиков» или «атомарных» вычислений; подсистемы управления вычислителями; подсистемы управления данными. Вместе данные области формируют ядро идеальной модели системы при рассмотрении ее как системы «черных ящиков». Каждое действие, которое формирует ядро вычислительной системы, должно обеспечиваться цепочкой действий и этапов жизненного цикла.

Получение информации о желаемом поведении системы и ее «обучении» (например, анализ функциональных требований, написание удовлетворяющей им программы и ее прошивка), могут выполняться на этапах создания системы или в процессе функционирования для самообучающихся систем. В последнем случае обучение можно добавить в расширенный вариант

ядра для каждой из подсистем, иначе обучение в ядро не включается и входит в процесс проектирования.

Подсистема управления вычислениями включает следующие этапы функционирования (вариант без возможности самообучения): определение следующего вычисления по предыдущему; запуск вычисления.

Подсистемы управления данными включает следующие этапы функционирования: чтение входов системы, блоков памяти, выходов вычислителей; коммутация данных внутри системы; запись данных в вычислители, блоки памяти и выходы системы.

Самообучение может быть добавлено в любую из двух подсистем и в сам вычислитель. Самообучение включает в себя получение и загрузку (реализацию) правил функционирования.

Представленная модель может быть использована для описания механизмов разной гранулярности. В качестве градаций гранулярности для вычислений можно выделить: инструкцию, базовый блок, функцию и процесс. Гранулярность для данных включает: бит данных, любое количество бит данных (байт, слово), пакет данных, файл и др. При необходимости можно наряду с цифровыми рассматривать аналоговые сигналы.

Примеры использования модели

Корректность модели была проверена на следующих примерах: основные архитектурные решения, используемые в типичном современном процессоре; программа для микроконтроллера; программируемая логическая интегральная схема (FPGA); операционная система; процессор функционального уровня (Function-Level Processor); математическая модель перцептрона. Была рассмотрена классификация вычислительных механизмов, обеспечивающих надежную работу подсистемы памяти с разнотипными блоками в контроллерах для ответственных применений.

Приведем пример описания архитектуры типичного современного процессора (рисунок), в котором использованы принципы фон-Неймана и стандартные архитектурные решения. Пример выбран как наиболее универсальный и понятный широкому кругу специалистов, метод может быть использован для описания более специфичных и нестандартных вычислительных архитектур.

На схеме показаны только наиболее важные моменты, дающие базовое понимание устройства процессора — основы, вынесенные на верхний уровень описания. При необходимости более детального описания любой из элементов, представленный на схеме (как этап функционирования, так и этап жизненного цикла) может быть расписан более подробно (на отдельной схеме или в «теле» основной схемы). Аналогичным образом описаны все вспомогательные механизмы — контроллер прерываний, механизмы работы со стеком, кэш, конвейер процессора, средства распараллеливания на уровне инструкций и др.

На «верхний уровень» описания вынесены следующие принципы и механизмы: — универсализация и разделение ресурса во времени — использование одного универсального

функционального блока, арифметико-логического устройства (АЛУ) процессора для выполнения различных операций;

- использование шины данных под управлением отдельно добавленных в программный код инструкций — простое базовое действие, «считать данные со входов системы/из блоков памяти/с выходов вычислителей», разложенное в цепочку действий, выполняемых сначала программистом при написании программы, а затем процессором в процессе ее исполнения;
- формирование последовательности выполнения операций с помощью механизмов управления программным счетчиком и написанного вручную программистом алгоритма, хранящегося в памяти.

На рисунке представлен единый алгоритм функционирования и создания системы с разделением на три этапа жизненного цикла: функционирование; создание аппаратного обеспечения; создание ПО (при необходимости количество этапов можно увеличить, добавив, например: изготовление системы, загрузку ПО, инициализацию и т. д.). Горизонтально ориентированные пятиугольники в области, обозначенной как «RUNTIME PROCESS», описывают вычислительный процесс — действия, производимые системой на этапе функционирования. Вертикально ориентированные пятиугольники в областях, обозначенных как «HARDWARE DESIGN» и «SOFTWARE DESIGN», описывают действия, производимые на этапах создания аппаратуры и ПО соответственно, без которых само функционирование системы невозможно.

В соответствии с предложенным устройством ядра, действия этапа функционирования разделены на три подсистемы: вычисления (Computations); подсистема управления вычислениями (Computation management subsystem, которой в процессоре соответствует блок, обычно называемый «Control Unit», CU); подсистема управления данными (Data management subsystem). Отметим, что каждое действие этапа функционирования относится к одному из базовых действий ядра, обозначенных небольшими прямоугольниками, в которые они вписываются. Например, для подсистемы управления вычислениями это: определение следующего вычисления по предыдущему (Select next computation), которое заключается в чтении регистра счетчика команд (PC) и запуск вычисления (Launch next computation), который подразумевает под собой загрузку соответствующей команды в регистр команд (IR).

В описанном примере обучение вычислительного блока системы — АЛУ (ALU) процессора — происходит на этапах создания аппаратуры и разработки архитектуры команд (ISA). На рисунке этот этап обозначен как TRAINING (ISA design). Последовательность выполнения отдельных операций с помощью АЛУ и коммутация данных внутри системы формируются на этапе создания ПО и обозначены как «Writing Program». На схеме два элемента с таким названием, поскольку в первом случае подразумевается исключительно формирование последовательности вычислений, а во втором — добавление в программу отдельных инструкций, обеспечивающих взаимодействие с памятью. Тем не менее обе эти работы выполняются программистом

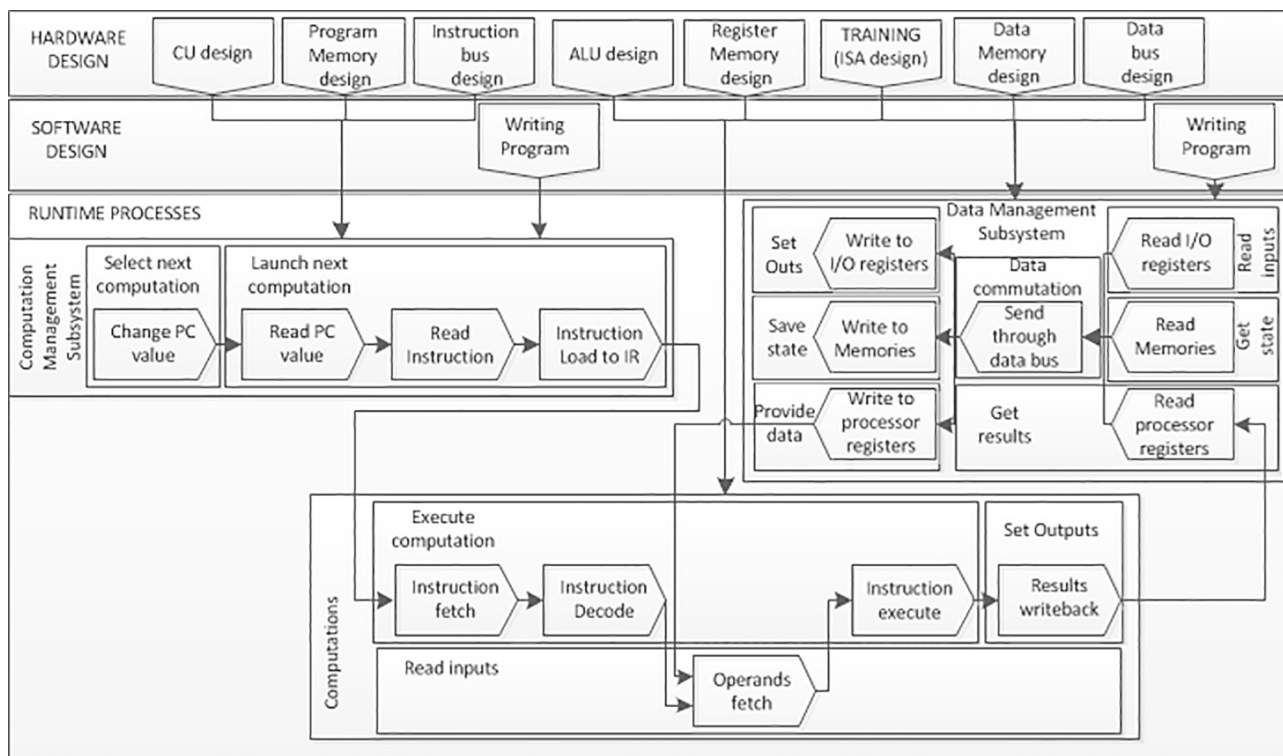


Рисунок. Пример документирования основных архитектурных решений типичного современного микропроцессора

Figure. An example of documenting the basic architectural solutions of a typical modern microprocessor

при написании ПО и обычно не разделяются на самостоятельные сущности.

Заключение

В работе предложен метод документирования архитектурных решений вычислительных платформ безотносительно к способу их конечной реализации, когда в единых терминах описываются: процесс проектирования и вычислительный процесс, аппаратура, программное обеспечение и инструментарий. Плавающий уровень абстракции позволяет получить описания для объектов разной гранулярности вычислений и данных.

Метод предоставляет особые правила структурирования информации об организации вычислительных платформ и о протекающих в них процессах. Основная цель метода — борьба со сложностью и упрощение обмена информацией с одновременным увеличением ее объема (рассматривается одновременно программное

обеспечение, аппаратура и инструментальные аспекты создания системы).

Использование метода дает возможность взглянуть по-новому на многие вопросы и позволяет генерировать новые нестандартные решения. Одно из главных достоинств метода — наглядность представления протекающих в системе процессов и возможность выделения основной сути функционирования системы — образующих ее вычислительных механизмов, для их отдельного описания и дальнейшего повторного использования.

Представленный метод в первую очередь предназначен для системных архитекторов и технических писателей, но может быть использован в процессе обучения специалистов соответствующих направлений. В частности, метод прошел апробацию в Университете ИТМО в рамках магистерской программы «Компьютерные системы и технологии».

Литература

1. Chattopadhyay A. Ingredients of adaptability: A survey of reconfigurable processors // *VLSI Design*. 2013. P. 1–18. <https://doi.org/10.1155/2013/683615>
2. Reshadi M. No-Instruction-Set-Computer (NISC) Technology Modeling and Compilation: PhD dissertation / University of California, Irvine. 2007. 153 p.
3. Somnath P., Swarup B. Computing with Memory for Energy-Efficient Robust Systems. Dordrecht: Springer, 2011. 249 p.
4. Siegl P., Buchty R., Berekovic M. Data-centric computing frontiers: A survey on processing-in-memory // *Proc. of the Second International Symposium on Memory Systems (MEMSYS '16)*. 2016. P. 295–308. <https://doi.org/10.1145/2989081.2989087>
5. Tabkhi H., Bushey R., Schirner G. Function-level processor (FLP): A novel processor class for efficient processing of streaming applications // *Journal of Signal Processing and Systems*. 2016. V. 85. N 1. P. 287–306. <https://doi.org/10.1007/s11265-015-1058-5>
6. Lysecky R., Stitt G., Vahid F. Warp Processors // *ACM Transactions on Design Automation of Electronic Systems*. 2006. V. 11. N 3. P. 659–681. <https://doi.org/10.1145/1142980.1142986>
7. Pinkevich V., Platonov A., Gorbachev Y. Design of embedded and cyber-physical systems using a cross-level microarchitectural pattern of the computational process organization // *CEUR Workshop Proceedings*. 2020. V. 2893.
8. Savage J. Models of Computation: Exploring the Power of Computing. Boston, MA, USA: Addison-Wesley, 1998. 600 p.
9. Processor Description Languages / ed. by M. Prabhat, D.Nikil. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc., 2008. 432 p. <https://doi.org/10.1016/B978-0-12-374287-2.X5001-0>
10. Aarenstrup R. Managing Model-Based Design. Natick, MA: MathWorks Inc., 2015. 86 p.
11. Nane R., Sima V., Pilato C., Choi J., Fort B., Canis A., Chen Y., Hsiao H., Brown S., Ferrandi F., Anderson J., Bertels K. A Survey and evaluation of FPGA high-level synthesis tools // *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*. 2016. V. 35. N 10. P. 1591–1604. <https://doi.org/10.1109/TCAD.2015.2513673>
12. Booch G., Jacobson I., Rumbaugh J. The Unified Modeling Language User Guide. Boston, MA, USA: Addison-Wesley, 1998. 391 p.
13. Delligatti L. SysML Distilled: A Brief Guide to the Systems Modeling Language. Boston, MA, USA: Addison-Wesley, 2013. 267 p.
14. Jean-Luc V. Model-based System and Architecture Engineering with the Arcadia Method. Elsevier, 2017. 388 p.

References

1. Chattopadhyay A. Ingredients of adaptability: A survey of reconfigurable processors. *VLSI Design*, 2013, pp. 1–18. <https://doi.org/10.1155/2013/683615>
2. Reshadi M. *No-Instruction-Set-Computer (NISC) Technology Modeling and Compilation*. PhD dissertation. University of California, Irvine. 2007, 153 p.
3. Somnath P., Swarup B. *Computing with Memory for Energy-Efficient Robust Systems*. Dordrecht, Springer, 2011, 249 p.
4. Siegl P., Buchty R., Berekovic M. Data-centric computing frontiers: A survey on processing-in-memory. *Proc. of the Second International Symposium on Memory Systems (MEMSYS '16)*, 2016, pp. 295–308. <https://doi.org/10.1145/2989081.2989087>
5. Tabkhi H., Bushey R., Schirner G. Function-level processor (FLP): A novel processor class for efficient processing of streaming applications. *Journal of Signal Processing and Systems*, 2016, vol. 85, no. 1, pp. 287–306. <https://doi.org/10.1007/s11265-015-1058-5>
6. Lysecky R., Stitt G., Vahid F. Warp Processors. *ACM Transactions on Design Automation of Electronic Systems*, 2006, vol. 11, no. 3, pp. 659–681. <https://doi.org/10.1145/1142980.1142986>
7. Pinkevich V., Platonov A., Gorbachev Y. Design of embedded and cyber-physical systems using a cross-level microarchitectural pattern of the computational process organization. *CEUR Workshop Proceedings*, 2020, vol. 2893.
8. Savage J. *Models of Computation: Exploring the Power of Computing*. Boston, MA, USA, Addison-Wesley, 1998, 600 p.
9. *Processor Description Languages*. Ed. by M. Prabhat, D.Nikil. San Francisco, CA, USA, Morgan Kaufmann Publishers Inc., 2008, 432 p. <https://doi.org/10.1016/B978-0-12-374287-2.X5001-0>
10. Aarenstrup R. *Managing Model-Based Design*. Natick, MA, MathWorks Inc., 2015, 86 p.
11. Nane R., Sima V., Pilato C., Choi J., Fort B., Canis A., Chen Y., Hsiao H., Brown S., Ferrandi F., Anderson J., Bertels K. A Survey and evaluation of FPGA high-level synthesis tools. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2016, vol. 35, no. 10, pp. 1591–1604. <https://doi.org/10.1109/TCAD.2015.2513673>
12. Booch G., Jacobson I., Rumbaugh J. *The Unified Modeling Language User Guide*. Boston, MA, USA, Addison-Wesley, 1998, 391 p.
13. Delligatti L. *SysML Distilled: A Brief Guide to the Systems Modeling Language*. Boston, MA, USA, Addison-Wesley, 2013, 267 p.
14. Jean-Luc V. *Model-based System and Architecture Engineering with the Arcadia Method*. Elsevier, 2017, 388 p.

Автор

Горбачев Ярослав Георгиевич — инженер, ООО ЛМТ, Санкт-Петербург, 199034, Российская Федерация; преподаватель практики, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация,  57216337562, <https://orcid.org/0000-0001-5419-6422>, yaroslav-go@yandex.ru

Author

Yaroslav G. Gorbachev — Engineer, LMT Ltd., Saint Petersburg, 199034, Russian Federation; Practice Lecturer, ITMO University, Saint Petersburg, 197101, Russian Federation,  57216337562, <https://orcid.org/0000-0001-5419-6422>, yaroslav-go@yandex.ru

Статья поступила в редакцию 15.05.2022
Одобрена после рецензирования 19.10.2022
Принята к печати 22.11.2022

Received 15.05.2022
Approved after reviewing 19.10.2022
Accepted 22.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1143-1149

Improving out of vocabulary words recognition accuracy for an end-to-end Russian speech recognition system

Andrei Yu. Andrusenko¹, Aleksei N. Romanenko²

^{1,2} STC-Innovations Ltd., Saint Petersburg, 194044, Russian Federation

^{1,2} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ andrusenkoau@itmo.ru, <https://orcid.org/0000-0002-8697-832X>

² AlexeySk8@gmail.com, <https://orcid.org/0000-0001-6267-018X>

Abstract

Automatic Speech Recognition (ASR) systems are experiencing an active introduction into our daily lives, simplifying the way we interact with electronic devices. The advent of end-to-end approaches has only accelerated this process. However, the constant evolution and a high degree of inflection of the Russian language lead to the problem of recognizing new words outside the vocabulary (Out Of Vocabulary, OOV) because they did not take part in the training process of the ASR system. In such a case, the ASR model tends to predict the most similar word from the training data which leads to a recognition error. This is especially true for ASR models that use decoding based on a Weighted Finite State Transducer (WFST), since they are obviously limited by the list of vocabulary words that may appear as a result of recognition. In this paper, this problem is investigated on the basis of an open data set of the Russian language (common voice) and an integrated ASR system using the WFST decoder. A method for retraining an integral ASR system based on the discriminative loss function MMI (maximum mutual information) and a method for decoding the integral model using a TG graph are proposed. Discriminative learning allows smoothing the probability distribution of acoustic class prediction, thus adding more variability in the recognition results. Decoding using the TG graph, in turn, is not limited to recognizing only vocabulary words and allows the use of a language model trained on a large amount of external text data. An eight-hour subset from the common voice base is used as a test set. The total number of OOV words in this test sample is 18.1 %. The results show that the use of the proposed methods allows to reduce the word recognition error (Word Error Rate, WER) by 3 % in absolute value relative to the standard method of decoding integral models (beam search), while maintaining the ability to recognize OOV words at a comparable level. The use of the proposed methods should improve the overall quality of recognition of ASR systems and make such systems more resistant to the recognition of new words that were not involved in the learning process.

Keywords

automatic speech recognition, end-to-end ASR, discriminative training, OOV words, weighted finite state transducer

For citation: Andrusenko A. Yu., Romanenko A. N. Improving out of vocabulary words recognition accuracy for an end-to-end Russian speech recognition system. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1143–1149. doi: 10.17586/2226-1494-2022-22-6-1143-1149

УДК 004.93

Повышение точности распознавания внесловарных слов для интегральной системы автоматического распознавания русской речи

Андрей Юрьевич Андрусенко¹, Алексей Николаевич Романенко²

^{1,2} ООО «ЦРТ-инновации», Санкт-Петербург, 194044, Российская Федерация

^{1,2} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ andrusenkoau@itmo.ru, <https://orcid.org/0000-0002-8697-832X>

² AlexeySk8@gmail.com, <https://orcid.org/0000-0001-6267-018X>

Аннотация

Предмет исследования. Системы автоматического распознавания речи (Automatic Speech Recognition, ASR) активно внедряются в нашу повседневную жизнь, тем самым упрощая способ взаимодействия с электронными

© Andrusenko A. Yu., Romanenko A. N., 2022

устройствами. Развитие интегральных (end-to-end) подходов только ускоряет данный процесс. Тем не менее постоянная эволюция и большая степень флективности русского языка приводят к проблеме распознавания новых вне словарных (Out Of Vocabulary, OOV) слов, которые не принимали участие в процессе обучения ASR-системы при ее создании. В связи с этим ASR-модель может прогнозировать наиболее похожее слово из обучающих данных, что влечет к ошибке распознавания. Особенно это касается ASR-моделей, использующих декодирование на основе взвешенного конечного автомата (Weighted Finite State Transducer, WFST), так как они заведомо ограничены списком словарных слов, которые могут появиться в результате распознавания. Выполнено исследование проблемы на основе открытой базы русского языка (common voice) и интегральной ASR-системы, использующей WFST-декодер. **Метод.** Предложен метод дообучения интегральной ASR-системы на основе дискриминативной функции потерь MMI (Maximum Mutual Information) и метода декодирования интегральной модели с помощью TG графа. Дискриминативное обучение позволило сгладить распределение вероятностей предсказания акустических классов, добавив таким образом большую вариативность в результате распознавания. Так как декодирование с помощью TG графа не имеет ограничений на распознавание только словарных слов, оно позволило использовать языковую модель, обученную на большом количестве внешних текстовых данных. **Основные результаты.** В качестве тестового множества использована восьмичасовая подвыборка из базы common voice. Общее число OOV-слов в тестовой выборке составило 18,1 %. Полученные результаты показали, что использование предложенных методов сократило пословную ошибку распознавания на 3 % в абсолютном значении относительно стандартного метода декодирования интегральных моделей. При этом сохранилась возможность распознавания OOV-слов на сравнимом уровне. **Практическая значимость.** Использование предложенных методов может улучшить общее качество распознавания ASR-систем и сделать их более устойчивыми к распознаванию новых слов, которые не участвовали в процессе обучения модели.

Ключевые слова

автоматическое распознавание речи, интегральные системы, дискриминативное обучение, OOV-слова, взвешенный конечный автомат

Ссылка для цитирования: Андрусенко А.Ю., Романенко А.Н. Повышение точности распознавания внесловарных слов для интегральной системы автоматического распознавания русской речи // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1143–1149 (на англ. яз.). doi: 10.17586/2226-1494-2022-22-6-1143-1149

Introduction

The active development of artificial deep neural networks has led to a significant breakthrough in Automatic Speech Recognition (ASR) tasks. For a long time, the conventional hybrid approach to building ASR systems [1] was dominant in this area. However, it is a complex and lengthy process of obtaining separate modules that are eventually combined into a single system. This encouraged the development of end-to-end learning methods for ASR models [2] the task of which is to combine all modules into one deep neural network. Unlike the conventional hybrid approach, the end-to-end method learns to generate text from an input audio signal directly without using an intermediate signal representation. This approach dramatically simplifies building and training an ASR system, showing high recognition accuracy comparable with conventional systems. In many ASR tasks, it is already the best solution [3, 4] which only increases the scientific community's attention to the development of this technology.

However, the constant evolution (emergence of new words) and a high degree of inflection (a large number of different spellings of the same word) of the Russian language leads to the degradation of the recognition accuracy of ASR systems since such “new” words may not be present in the training data. These words are called Out Of Vocabulary (OOV) words. The words in the training data are called IV (in vocabulary) words. Conventional methods, by definition, are not capable of recognizing new words since their decoder is pre-limited by the list of words (lexicon) presented in training data. To tackle this problem, such OOV words can be added to the lexicon graph [5], but it is necessary to know the list of OOV words

in advance. To recognize words previously unknown to the model, there is a method of adding the unk subgraph [6]; it can recognize new words in cases where the ASR model could not find the corresponding IV word. This is achieved by adding a phoneme subgraph with arbitrary transitions within itself which allows generating new sequences of phonemes that are then converted into words. Such unk subgraph is integrated into the lexicon graph which has only known paths based on IV words. However, the accuracy of this method is far from perfect. A comparison of OOV word recognition techniques for classical ASR systems is considered in [7].

End-to-end ASR models in the recognition process produce separate target acoustic units; it can be letters (graphemes) or pieces of words (subwords) consisting of several letters. Such a method is theoretically capable of recognizing OOV words, but in practice, this can lead to degradation of the overall recognition accuracy. The model may be wrong in IV word recognition because it does not have a lexicon constraint during the decoding process, as it does in the conventional hybrid model. With a limited number of training data (less than 100–150 hours of transcribed data), this problem only gets worse. To improve this situation, BPE (Byte pair encoding) augmentation [8] of acoustic classes can be used generating new word divisions into acoustic units during each training epoch. A comparison of conventional hybrid and end-to-end ASR systems for the Russian language can be found in [9].

There is a method for combining an encoder trained in end-to-end mode with a WFST (Weighted Finite State Transducer) decoder used for conventional systems [10]. This approach allows using external text data to train the language model and make fewer mistakes in recognizing IV words due to the presence of a lexicon. This combines the

strengths of end-to-end systems (robust encoders trained in end-to-end mode) and a strong language model integrated into the WFST graph from the classical system. However, the lexicon still limits this approach making it impossible to recognize OOV words.

This paper proposes a WFST graph method to decode an end-to-end model unrestricted by a predetermined lexicon L. Instead of TLG, the TG graph is used in which the language model is based on graphemes. We also use the discriminative MMI (Maximum Mutual Information) loss function for finetuning the pre-trained end-to-end model similarly from [11] which further increases the accuracy of OOV word recognition by smoothing the probabilities of acoustic unit prediction. The final model reduces the overall WER (Word Error Rate) by 3 % in absolute value relative to the baseline end-to-end decoding algorithm while retaining the ability to recognize OOV words on the same level.

Decoding method

The conventional hybrid ASR system is divided into two main components: acoustic and language models. The classifier of target acoustic units (phonemes, graphemes, subwords), based on a deep neural network, usually acts as an acoustic model. Its task is to predict the probability distribution of acoustic units on each short segment (25 ms long) of the input signal. The language model is integrated into the decoding module and generates the most probable word sequences based on the predictions received from the acoustic model. The lexicon sets the rules for dividing words into acoustic units mapping each word from the training data to the corresponding sequence of acoustic units. For example, if graphemes are used as acoustic units, then the lexicon will display the rules for splitting each word into letters. The final decoding graph for the conventional hybrid ASR system is a WFST that sequentially transforms the outputs of the acoustic model into the final recognition result in the text. Such a graph is a composition of four separate graphs:

- H — topology graph for acoustic model outputs; converts acoustic model outputs (tying states) to context-sensitive acoustic classes (such as triphones).
- C — transition graph from context-dependent acoustic classes to context-independent ones (for example, from triphones to phonemes).
- L — lexicon.
- G — word language model.

As a result, we get the HCLG graph which is used in the process of decoding the conventional hybrid system.

For end-to-end ASR systems, it is possible to use such a WFST graph for decoding [10] since its encoder also plays the role of an acoustic model. To do this, it is necessary to replace the HC subgraph of the conventional model with T, the graph of the CTC (Connectionist Temporal Classification) topology. As a result, a TLG graph is obtained which can decode the results of recognition of the end-to-end model. This method uses both approaches strengths: an encoder based on the latest architectures, trained in end-to-end mode, and a WFST graph with a

powerful language model. However, due to the L graph presence, this method cannot recognize OOV words.

To solve this problem, we propose to remove the L graph from the composition of the decoding graph. As a result, we get a TG graph, where G is already a grapheme language model instead of a word one. Using positional graphemes as acoustic units makes it possible to directly restore the sequence of words from the results of such “non-word” recognition. This approach retains all the advantages of the hybrid method of decoding the end-to-end model using the WFST graph and makes it possible to recognize new OOV words by getting rid of the lexicon restriction.

Discriminative finetuning

To improve the recognition accuracy of end-to-end systems, one can use discriminative training methods based on MMI objective function [11]. In contrast to the MMI finetuning for the conventional hybrid system, in the end-to-end approach, the loss function is calculated immediately over the entire input signal, leading to extensive memory consumption. The compact CTC topology can be used to combat this, it described in detail in [12]. Additional finetuning of the pre-trained end-to-end model for a certain number of epochs with the MMI loss function can have a beneficial effect on the accuracy of the final system. Our experiments show that this approach improves the overall recognition accuracy and leads to a noticeable improvement in OOV word recognition. This effect may be because the model trained only on the CTC loss function initially has excessively high probabilities for predictions, which was already noted earlier in [13]. Additional finetuning with MMI allows smoothing out this “peaky” behavior and increases the probability of recognition of alternative acoustic units improving the recognition accuracy of OOV words. An example of comparing the probability distribution by acoustic units for a model trained only on CTC loss and then finetuned with MMI loss is shown in Fig. 1.

Data set

An open database of the Russian language common voice [14] is used to train and test ASR models for the proposed method. This database consists of voice recordings from users’ personal devices (smartphones, PCs, etc.). The sampling rate of audio files is adjusted to 16000 Hz. The entire data available for download from the original source was divided into a train (104 hours and 38 minutes) and a test set (8 hours and 16 minutes). The number of OOV words for the selected test set was 18.1 %. It is important to note that this is a percentage of the total words spoken on the test, not a percentage of unique OOV words.

As an acoustic feature, 80-dimensional log-Mel filterbank coefficients are used, to be calculated with a window of 25 ms and a step of 10 ms. In order to reduce the effect of model overfitting, SpecAugment [15] is used. The number of frequency masks is 2 with a mask range of 30 bins. The number of time masks is 2 with a length of 40 ms.

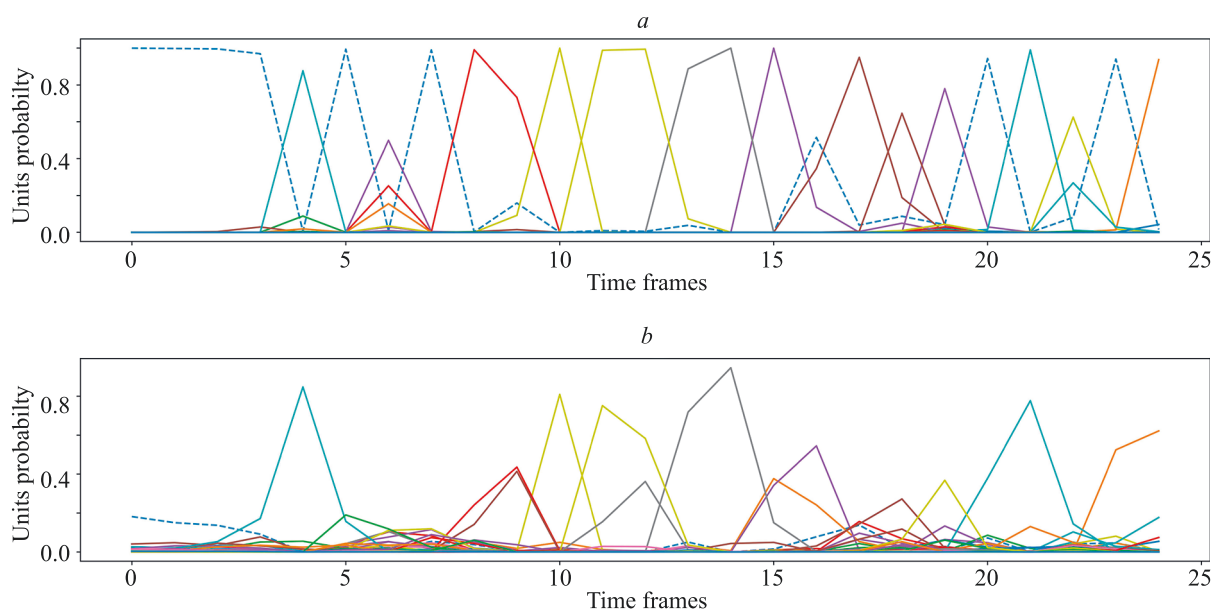


Fig. 1. Comparison of the acoustic units distribution probabilities of the end-to-end model for two training options: only with the CTC loss function (a) and the use of additional training with 10 MMI epochs (b). The dotted line indicates the probability of a blank symbol. Continuous lines correspond to the remaining acoustic units

ASR system

The acoustic model (encoder) of our end-to-end ASR system is Conformer [16] — a deep neural network architecture based on the attention mechanism and local convolution blocks. This architecture works effectively with a long-term and local context within utterances (for example, with whole words and individual graphemes). In this work, we use a 12-layer Conformer with the size of the attention mechanism embedding of 320. The number of attention heads is 8. The size of fully connected layers is 1024, and the size of the conformer convolution block is 5.

The Joint CTC-Attention mode (the weight of the CTC loss function is 0.3) is used for training the end-to-end model. An attention decoder part based on the Transformer architecture [17] consists of 6 layers with an attention mechanism size of 320. The number of attention heads is 8, and the size of fully connected layers is 1024. This block is also used for decoding in the beam search mode. The weight of the encoder block is 0.3. Adam algorithm is used as an optimizer with 16000 warm-up steps to a peak learning rate of 0.002, followed by a quadratic decrease. The total number of training epochs is 100.

For decoding in conventional hybrid mode, a WFST graph is used based on the Kaldi toolkit [18]. For language modeling, a 3-gram word language model is used in the case of the TLG.fst graph and a 7-gram grapheme language model for the TG.fst graph.

The word error rate (WER) metric is used to evaluate the overall accuracy of speech recognition. OOV word recognition accuracy is measured using Character Error Rate (OOV-CER) and Word Error Rate (OOV-WER) for specific OOV words. As an implementation of such a metric, the texterrors¹ are used. It is worth noting that

¹ Available at: <https://github.com/RuABraun/texterrors> (accessed: 15.11.2022).

WER is the highest priority indicator in comparing the accuracy of model recognition. OOV-CER and OOV-WER are secondary metrics that show only OOV words recognition accuracy.

All experiments were carried out in the ESPnet speech recognition toolkit [19].

Experiment results

As baseline values for comparison with subsequent modifications, the speech recognition accuracy of the end-to-end model was investigated using a standard beam search decoding algorithm. The influence of the choice of acoustic units for ASR system modeling was also investigated. The results of these experiments are presented in Table 1. Subwords were obtained using sentencepiece tokenizer with 250 BPE units. Additionally, the effect of BPE augmentation with `bpe_alpha 0.1 (dp0.1)` on the recognition accuracy of OOV words was tested. It can be concluded that the graphemes units work the best for this task. The BPE model shows a slightly better WER, but noticeably loses in OOV word recognition. BPE dropout helps improve OOV recognition but messes up WER a lot.

The following experiments use the WFST graph to decode the predictions of the end-to-end model encoder. Table 2 shows that the transition from TLG.fst to TG.fst can significantly reduce the OOV-CER/WER and the

Table 1. Decoding the end-to-end model using the beam search algorithm for different acoustic units, %

Acoustic units	WER	OOV-CER	OOV-WER
Graphemes	30.5	28.8	58.1
BPE 250	30.4	32.8	62.2
BPE 250 + dp0.1	37.5	32.6	60.0

Table 2. Results for the WFST decoding method of the end-to-end model using TLG.fst and TG.fst, %

Training criterion	Graph	WER	OOV-CER	OOV-WER
CTC-Att	TLG.fst	34.3	51.9	100.0
CTC-Att	TG.fst	29.7	33.6	62.4
CTC-Att+MMI	TLG.fst	33.5	50.9	100.0
CTC-Att+MMI	TG.fst	27.5	31.0	60.4

overall WER. At the same time, additional finetuning for ten epochs using the MMI loss function provides a further improvement in both WER and OOV metrics.

Table 3 compares the baseline beam search with our best result from Table 2. The method of adding the unk subgraph was also tested. It can be seen that baseline beam search is the best for recognition of OOV words but shows that WER is 3 % worse than TG.fst + MMI which still works well with OOV words and is inferior in OOV metrics quite a bit. Model with the unk subgraph is also capable of recognizing OOV words but lags far behind in WER and OOV metrics.

Additional analysis of recognition results indicates that most of the recognized OOV words in about 96.3 % differ from IV words by no more than three letters. It means that the proposed method works well with the inflection of the language (small changes in the ending, word prefix, etc.). Recognition of a new OOV word that is not similar to an existing IV word is rare. As a measure of the proximity of two words, the Levenshtein distance is used. It is defined as the minimum number of single-character operations (insertions, deletions, and replacements) required to convert one character sequence to another. A complete distribution of the number of correctly recognized OOV words depending on the Levenshtein distance to the nearest IV word from the training set can be seen in Fig. 2.

Table 4 shows an example of recognized OOVs and IV words closest to them.

Table 4. An example of recognized OOV words and IV words closest to them. Grapheme differences are highlighted in bold

OOV	IV	Levenshtein distance
обширным	обширный	1
немногими	немногим	1
платформой	платформы	2
подпишут	подпись	3
вознаграждено	вознаградить	4

Table 3. Comparison of the best results for different recognition methods of the same end-to-end ASR model, %

Method	WER	OOV-CER	OOV-WER
Baseline beam search	30.5	28.8	58.1
TG.fst + MMI	27.5	31.0	60.4
TLG.fst + unk subgraph	32.1	37.5	74.1

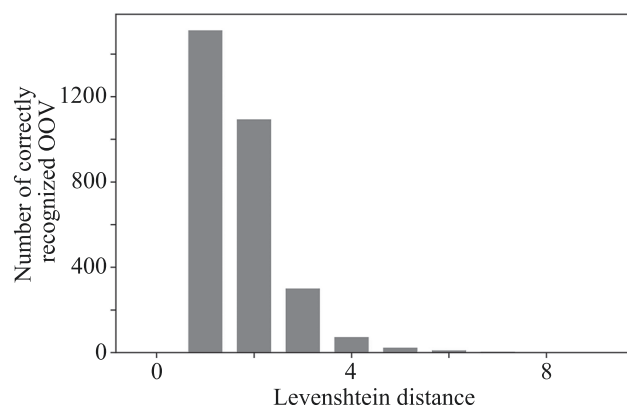


Fig. 2. Distribution of the number of correctly recognized OOV words depending on the Levenshtein distance to the nearest IV word

Conclusion

This paper considers the problem of OOV word recognition for the Russian language using the end-to-end ASR system. The possibilities of OOV recognition were explored using the beam search decoding algorithm compared to the WFST decoding graph. The proposed methods of using the TG.fst graph and discriminative MMI finetuning allowed reducing the overall WER by 3 % in absolute value compared to the beam search decoding while maintaining a high level of recognition ability for OOV words obtained as a result of the inflection of the Russian language. Our method is also significantly superior to the existing approach based on using the unk graph.

As further research, it is planned to improve the grapheme language model (G graph) by increasing the number of N-grams. It is assumed that the bigger context of the language model will help to deal better with OOV word recognition.

References

Литература

1. Hinton G., Deng L., Yu D., Dahl G.E., Mohamed A., Jaitly N., Senior A., Vanhoucke V., Nguyen P., Sainath T.N., Kingsbury B. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Processing Magazine*, 2012, vol. 29, no. 6, pp. 82–97. <https://doi.org/10.1109/MSP.2012.2205597>
2. Graves A., Fernandez S., Gomez F., Schmidhuber J. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks. *Proc. of the 23rd International Conference on Machine Learning (ICML)*, 2006, pp. 369–376. <https://doi.org/10.1145/1143844.1143891>
3. Synnaeve G., Xu Q., Kahn J., Likhomanenko T., Grave E., Pratat V., Sriram A., Liptchinsky V., Collobert R. End-to-end ASR: From supervised to semi-supervised learning with modern architectures. *arXiv*, 2019, ArXiv:1911.08460. <https://doi.org/10.48550/arXiv.1911.08460>
4. Li J., Lavrukhin V., Ginsburg B., Leary R., Kuchaiev O., Cohen J.M., Nguyen H., Gadde R.T. Jasper: An end-to-end convolutional neural acoustic model. *Proc. of the 20th Annual Conference of the International Speech Communication Association: Crossroads of Speech and Language (INTERSPEECH)*, 2019, pp. 71–75. <https://doi.org/10.21437/Interspeech.2019-1819>
5. Khokhlov Y., Tomashenko N., Medennikov I., Romanenko A. Fast and accurate OOV decoder on high-level features. *Proc. of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2017, pp. 2884–2888. <https://doi.org/10.21437/Interspeech.2017-1367>
6. Alumaë A., Tilk O., Ullah A. Advanced rich transcription system for Estonian speech. *Frontiers in Artificial Intelligence and Applications*, 2018, vol. 307, pp. 1–8. <https://doi.org/10.3233/978-1-61499-912-6-1>
7. Braun R., Madikeri S., Motlicek P. A comparison of methods for OOV-word recognition on a new public dataset. *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2021, pp. 5979–5983. <https://doi.org/10.1109/ICASSP39728.2021.9415124>
8. Laptev A., Andrusenko A., Podluzhny I., Mitrofanov A., Medennikov I., Matveev Y. Dynamic acoustic unit augmentation with BPE-dropout for low-resource end-to-end speech recognition. *Sensors (Basel)*, 2021, vol. 21, no. 9, pp. 3063. <https://doi.org/10.3390/s21093063>
9. Andrusenko A., Laptev A., Medennikov I. Exploration of end-to-end ASR for OpenSTT — Russian open speech-to-text dataset. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2020, vol. 12335, pp. 35–45. https://doi.org/10.1007/978-3-030-60276-5_4
10. An K., Xiang H., Ou Z. CAT: A CTC-CRF based ASR toolkit bridging the hybrid and the end-to-end approaches towards data efficiency and low latency. *Proc. of the 21st Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2020, pp. 566–570. <https://doi.org/10.21437/Interspeech.2020-2732>
11. Hadian H., Sameti H., Povey D., Khudanpur S. End-to-end speech recognition using lattice-free MMI. *Proc. of the 19th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2018, pp. 12–16. <https://doi.org/10.21437/Interspeech.2018-1423>
12. Laptev A., Majumdar S., Ginsburg B. CTC variations through new WFST topologies. *Proc. of the 23rd Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2022, pp. 1041–1045. <https://doi.org/10.21437/Interspeech.2022-10854>
13. Zeyer A., Schlüter R., Ney H. Why does CTC result in peaky behavior? *arXiv*, 2021, arXiv:2105.14849. <https://doi.org/10.48550/arXiv.2105.14849>
14. Ardila R., Branson M., Davis K., Henretty M., Kohler M., Meyer J., Morais R., Saunders L., Tyers F.M., Weber G. Common voice: A massively-multilingual speech corpus. *Proc. of the 12th International Conference on Language Resources and Evaluation (LREC)*, 2020, pp. 4218–4222.
15. Park D., Chan W., Zhang Y., Chiu C., Zoph B., Cubuk E.D., Le Q.V. SpecAugment: A simple data augmentation method for automatic speech recognition. *Proc. of the 20th Annual Conference of the International Speech Communication Association: Crossroads of Speech and Language (INTERSPEECH)*, 2019, pp. 2613–2617. <https://doi.org/10.21437/Interspeech.2019-2680>
1. Hinton G., Deng L., Yu D., Dahl G.E., Mohamed A., Jaitly N., Senior A., Vanhoucke V., Nguyen P., Sainath T.N., Kingsbury B. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups // *IEEE Signal Processing Magazine*. 2012. V. 29. N 6. P. 82–97. <https://doi.org/10.1109/MSP.2012.2205597>
2. Graves A., Fernandez S., Gomez F., Schmidhuber J. Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks // *Proc. of the 23rd International Conference on Machine Learning (ICML)*. 2006. P. 369–376. <https://doi.org/10.1145/1143844.1143891>
3. Synnaeve G., Xu Q., Kahn J., Likhomanenko T., Grave E., Pratat V., Sriram A., Liptchinsky V., Collobert R. End-to-end ASR: From supervised to semi-supervised learning with modern architectures // *arXiv*. 2019. ArXiv:1911.08460. <https://doi.org/10.48550/arXiv.1911.08460>
4. Li J., Lavrukhin V., Ginsburg B., Leary R., Kuchaiev O., Cohen J.M., Nguyen H., Gadde R.T. Jasper: An end-to-end convolutional neural acoustic model // *Proc. of the 20th Annual Conference of the International Speech Communication Association: Crossroads of Speech and Language (INTERSPEECH)*. 2019. P. 71–75. <https://doi.org/10.21437/Interspeech.2019-1819>
5. Khokhlov Y., Tomashenko N., Medennikov I., Romanenko A. Fast and accurate OOV decoder on high-level features // *Proc. of the 18th Annual Conference of the International Speech Communication Association (INTERSPEECH)*. 2017. P. 2884–2888. <https://doi.org/10.21437/Interspeech.2017-1367>
6. Alumaë A., Tilk O., Ullah A. Advanced rich transcription system for Estonian speech // *Frontiers in Artificial Intelligence and Applications*. 2018. V. 307. P. 1–8. <https://doi.org/10.3233/978-1-61499-912-6-1>
7. Braun R., Madikeri S., Motlicek P. A comparison of methods for OOV-word recognition on a new public dataset // *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. 2021. P. 5979–5983. <https://doi.org/10.1109/ICASSP39728.2021.9415124>
8. Laptev A., Andrusenko A., Podluzhny I., Mitrofanov A., Medennikov I., Matveev Y. Dynamic acoustic unit augmentation with BPE-dropout for low-resource end-to-end speech recognition // *Sensors (Basel)*. 2021. V. 21. N 9. P. 3063. <https://doi.org/10.3390/s21093063>
9. Andrusenko A., Laptev A., Medennikov I. Exploration of end-to-end ASR for OpenSTT - Russian open speech-to-text dataset // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2020. V. 12335. P. 35–45. https://doi.org/10.1007/978-3-030-60276-5_4
10. An K., Xiang H., Ou Z. CAT: A CTC-CRF based ASR toolkit bridging the hybrid and the end-to-end approaches towards data efficiency and low latency // *Proc. of the 21st Annual Conference of the International Speech Communication Association (INTERSPEECH)*. 2020. P. 566–570. <https://doi.org/10.21437/Interspeech.2020-2732>
11. Hadian H., Sameti H., Povey D., Khudanpur S. End-to-end speech recognition using lattice-free MMI // *Proc. of the 19th Annual Conference of the International Speech Communication Association (INTERSPEECH)*. 2018. P. 12–16. <https://doi.org/10.21437/Interspeech.2018-1423>
12. Laptev A., Majumdar S., Ginsburg B. CTC variations through new WFST topologies // *Proc. of the 23rd Annual Conference of the International Speech Communication Association (INTERSPEECH)*. 2022. P. 1041–1045. <https://doi.org/10.21437/Interspeech.2022-10854>
13. Zeyer A., Schlüter R., Ney H. Why does CTC result in peaky behavior? // *arXiv*. 2021. arXiv:2105.14849. <https://doi.org/10.48550/arXiv.2105.14849>
14. Ardila R., Branson M., Davis K., Henretty M., Kohler M., Meyer J., Morais R., Saunders L., Tyers F.M., Weber G. Common voice: A massively-multilingual speech corpus // *Proc. of the 12th International Conference on Language Resources and Evaluation (LREC)*. 2020. P. 4218–4222.
15. Park D., Chan W., Zhang Y., Chiu C., Zoph B., Cubuk E.D., Le Q.V. SpecAugment: A simple data augmentation method for automatic speech recognition // *Proc. of the 20th Annual Conference of the International Speech Communication Association: Crossroads of Speech and Language (INTERSPEECH)*. 2019. P. 2613–2617. <https://doi.org/10.21437/Interspeech.2019-2680>

16. Gulati A., Qin J., Chiu C., Parmar N., Zhang Y., Yu J., Han W., Wang S., Zhang Z., Wu Y., Pang R. Conformer: Convolution-augmented transformer for speech recognition. *Proc. of the 21st Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2020, pp. 5036–5040. <https://doi.org/10.21437/Interspeech.2020-3015>
17. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser Ł., Polosukhin I. Attention is all you need. *Proc. of the 31st Annual Conference on Neural Information Processing Systems (NIPS)*, 2017, pp. 5998–6008.
18. Povey D., Ghoshal A., Boulianne G., Burget L., Glembek O., Goel N., Hannemann M., Motlicek P., Qian Y., Schwarz P., Silovsky J., Stemmer G., Vesely K. The Kaldi speech recognition toolkit. *Proc. of the IEEE 2011 Workshop on Automatic Speech Recognition and Understanding*, 2011.
19. Watanabe S., Hori T., Karita S., Hayashi T., Nishitoba J., Unno Y., Soplin N.E.Y., Heymann J., Wiesner M., Chen N., Renduchintala A., Ochiaiet T. ESPnet: End-to-end speech processing toolkit. *Proc. of the 19th Annual Conference of the International Speech Communication (INTERSPEECH)*, 2018, pp. 2207–2211. <https://doi.org/10.21437/Interspeech.2018-1456>
16. Gulati A., Qin J., Chiu C., Parmar N., Zhang Y., Yu J., Han W., Wang S., Zhang Z., Wu Y., Pang R. Conformer: Convolution-augmented transformer for speech recognition // *Proc. of the 21st Annual Conference of the International Speech Communication Association (INTERSPEECH)*. 2020. P. 5036–5040. <https://doi.org/10.21437/Interspeech.2020-3015>
17. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A.N., Kaiser Ł., Polosukhin I. Attention is all you need // *Proc. of the 31st Annual Conference on Neural Information Processing Systems (NIPS)*. 2017. P. 5998–6008.
18. Povey D., Ghoshal A., Boulianne G., Burget L., Glembek O., Goel N., Hannemann M., Motlicek P., Qian Y., Schwarz P., Silovsky J., Stemmer G., Vesely K. The Kaldi speech recognition toolkit // *Proc. of the IEEE 2011 Workshop on Automatic Speech Recognition and Understanding*. 2011.
19. Watanabe S., Hori T., Karita S., Hayashi T., Nishitoba J., Unno Y., Soplin N.E.Y., Heymann J., Wiesner M., Chen N., Renduchintala A., Ochiaiet T. ESPnet: End-to-end speech processing toolkit // *Proc. of the 19th Annual Conference of the International Speech Communication (INTERSPEECH)*. 2018. P. 2207–2211. <https://doi.org/10.21437/Interspeech.2018-1456>

Authors

Andrei Yu. Andrusenko — PhD Student, Scientific Researcher, STC-Innovations Ltd., Saint Petersburg, 194044, Russian Federation; Software Developer, ITMO University, 197101, Russian Federation, <https://orcid.org/0000-0002-8697-832X>, andrusenkoau@itmo.ru

Aleksei N. Romanenko — PhD, Leading Researcher, STC-Innovations Ltd., Saint Petersburg, 194044, Russian Federation; Senior Researcher, ITMO University, 197101, Russian Federation, <https://orcid.org/0000-0001-6267-018X>, AlexeySk8@gmail.com

Авторы

Андрусенко Андрей Юрьевич — аспирант, научный сотрудник, ООО «ЦРТ-инновации», Санкт-Петербург, 194044, Российская Федерация; программист, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-8697-832X>, andrusenkoau@itmo.ru

Романенко Алексей Николаевич — кандидат технических наук, ведущий научный сотрудник, ООО «ЦРТ-инновации», Санкт-Петербург, 194044, Российская Федерация; старший научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-6267-018X>, AlexeySk8@gmail.com

Received 22.08.2022

Approved after reviewing 25.10.2022

Accepted 22.11.2022

Статья поступила в редакцию 22.08.2022

Одобрена после рецензирования 25.10.2022

Принята к печати 22.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1150-1158

УДК 004.021

Метод мониторинга состояния элементов киберфизических систем на основе анализа временных рядов

Виктор Викторович Семенов✉

Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация

v.semenov@spcras.ru✉, <https://orcid.org/0000-0002-7216-769X>

Аннотация

Предмет исследования. Широкое распространение киберфизических систем, а также повсеместная интеграция вычислительных ресурсов в физические сущности привели к повышению рисков преднамеренных и случайных инцидентов безопасности. В связи с этим приобретает особую актуальность разработка новых и совершенствование существующих методов и средств мониторинга таких систем. Создаваемые и модернизируемые методы должны обладать повышенной полнотой и точностью идентификации, в особенности для объектов критической инфраструктуры. **Метод.** Предложен оригинальный метод обработки данных мониторинга состояния киберфизических систем на основе анализа временных рядов с применением весовых коэффициентов значимости в качестве постобработки результатов классификации. Метод отличается от существующих комбинированным подходом, сочетающим применение в системах мониторинга событий информационной и функциональной безопасности. Характеризуется использованием ансамбля деревьев решений, а также параллельно работающих классификаторов и весовых коэффициентов Фишберна при анализе совокупности наиболее информативных признаков, полученных из временных рядов. **Основные результаты.** Применимость метода обоснована при помощи вычислительного эксперимента на известном наборе данных, характеризующем функционирование информационной и физической составляющих при осуществлении различных типов атак на компоненты экспериментального стенда киберфизической системы водоочистки. Точность идентификации по сравнению с лучшими подходами, представленными в научных работах при использовании разработанного метода, увеличилась на 1,45 %, полнота — на 4,45 % и составила 99,85 % для обоих показателей. **Практическая значимость.** Полученные результаты адаптированы для практического использования в системах идентификации состояния киберфизических систем. Теоретическая значимость состоит в возможности использования результатов исследования при проектировании систем мониторинга информационной и функциональной безопасности киберфизических систем.

Ключевые слова

системы мониторинга, анализ временных рядов, киберфизические системы, выявление аномалий, информационная безопасность, функциональная безопасность, решающие деревья

Ссылка для цитирования: Семенов В.В. Метод мониторинга состояния элементов киберфизических систем на основе анализа временных рядов // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1150–1158. doi: 10.17586/2226-1494-2022-22-6-1150-1158

Method for monitoring the state of elements of cyber-physical systems based on time series analysis

Viktor V. Semenov✉

St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation

v.semenov@spcras.ru✉, <https://orcid.org/0000-0002-7216-769X>

Abstract

The wide spread of cyber-physical systems, as well as the widespread integration of computing resources into physical entities, have led to an increase in the risks of deliberate and accidental security incidents. In this regard, the development

© Семенов В.В., 2022

of new methods and tools and improvement of the existing ones for monitoring such systems is of particular relevance. The methods being created and modernized should have increased recall and precision of identification, especially for critical infrastructure objects. An original method for processing data for monitoring the state of cyber-physical systems based on time series analysis using significance weights as a post-processing of classification results was proposed. The method differs from the existing ones by the combined approach that combines the use events of information security and functional safety in monitoring systems. It is characterized by the use of an ensemble of decision trees as well as parallel classifiers and Fishburn weight coefficients in the analysis of the set of the most informative features obtained from time series. The applicability of the method was substantiated by conducting of a computational experiment on a known data set which characterizes the functioning of the information and physical components in the implementation of various types of attacks on the components of the experimental stand of the cyber-physical water treatment system. When using the developed method, the identification precision increased by 1.45 % compared to the best approaches presented in other scientific works, and the recall increased by 4.45 % and amounted to 99.85 % for both indicators. The results obtained are adapted for practical use in systems for identifying the state of cyber-physical systems. The theoretical significance lies in the possibility of using the results of the study in the design of systems for monitoring the information security and functional safety of cyber-physical systems.

Keywords

monitoring systems, time series analysis, cyber-physical systems, identification of anomalies, information security, functional safety, decision trees

For citation: Semenov V.V. Method for monitoring the state of elements of cyber-physical systems based on time series analysis. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1150–1158 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1150-1158

Введение

Слияние информационных технологий и промышленных процессов, наблюдаемое в последние годы в развитии технологической инфраструктуры, привело к трансформации принципов построения производственных объектов и широкому распространению киберфизических систем (КФС).

На современном этапе развития КФС отмечается повышение степени интеллектуальности систем управления, их автономности и адаптивности, вместе с этим стремительно возрастает объем обрабатываемой информации, передаваемой от различных сенсоров и датчиков [1]. Данные системы являются сложными и

распределенными, что также ведет к возникновению ряда проблем, связанных с их работоспособностью и информационной безопасностью (ИБ) [2]. Таким образом, необходимо обеспечение постоянного мониторинга состояния КФС, в том числе оценки защищенности, при этом крайне важным является учет временных параметров потенциальных инцидентов безопасности.

Среди основных уязвимостей КФС можно выделить: возможность прослушивания каналов, посылку «внешних» пакетов, осуществление физического доступа злоумышленника к объекту КФС, недостаточную стандартизацию интеллектуальных алгоритмов маршрутизации, учитывающих состояние сети и др. На рис. 1 приведены типовые деструктивные воздействия



Рис. 1. Угрозы безопасности на различных уровнях киберфизических систем

Fig. 1. Security threats at different levels of the cyber-physical systems

на элементы КФС на различных уровнях: физическом, сетевом и уровне приложений.

Управление инцидентами безопасности, включая их выявление, фиксацию, предсказание — постоянный сложный процесс наблюдения и анализа результатов событий безопасности и иных данных. Подобного рода мониторинг представляет собой комплексную задачу по работе с угрозами и нарушениями ИБ, а также технологическими сбоями и отказами, осложненными разнородностью промышленных сетевых устройств и протоколов, количеством данных и скоростью их поступления. Дополнительную проблему предоставляет необходимость адаптации средств мониторинга в динамически меняющихся условиях. Исходя из этого, перспективной задачей является разработка методов мониторинга состояния элементов КФС с целью обеспечения их комплексной информационно-функциональной безопасности, а также устойчивого функционирования в условиях наличия информационных угроз и атак.

Существующие методы и технологии в большей мере ориентированы на классические компьютерные или информационные системы [3–5], что ограничивает возможность их применения в КФС. Используемые на сегодняшний день решения [6, 7] не обладают достаточным функционалом, обеспечивающим эффективный мониторинг в режиме реального времени, что вызывает ряд проблем обеспечения информационно-функциональной безопасности, связанных с анализом состояния отдельных устройств КФС. В связи с этим возникает объективная необходимость развития и адаптации методов математического обеспечения специализированных информационных систем, интегрируемых в КФС, в целях противодействия внешним и внутренним деструктивным воздействиям.

Настоящая работа — логическое продолжение работ [8, 9].

Постановка задачи

Рассмотрим задачу классификации временных рядов, характеризующих состояние информационно-функциональной безопасности КФС. Пусть имеется временной ряд $\mathbf{X} = \{\{x_1(t_1), x_2(t_1), \dots, x_S(t_1)\}, \{x_1(t_2), x_2(t_2), \dots, x_S(t_2)\}, \dots, \{x_1(t_m), x_2(t_m), \dots, x_S(t_m)\}\}$, каждому кортежу которого соответствует набор характеристик информационных или физических процессов КФС в дискретный момент времени; $\{c_1, \dots, c_l\}$ — множество состояний КФС, l — число идентифицируемых состояний КФС; $C = \{C_0, C_1\}$ — множество рассматриваемых состояний. Каждый элемент КФС может находиться в опасном (C_1) или безопасном (разрешенном) (C_0) состоянии. $C_0 = \{c_1, c_2, \dots, c_k\}$, $C_1 = \{c_{k+1}, c_{k+2}, \dots, c_l\}$, k — число безопасных состояний КФС.

Требуется определить состояния КФС (метка класса c), к которому относится подаваемый на вход элемент временного ряда. Для обучения модели необходимо получить характеристику протекающих информационных и физических процессов каждого рассматриваемого состояния КФС. Формирование обучающей выборки может производиться, например, с

использованием программного обеспечения для автоматизированного анализа сетевого трафика.

Метка класса состояния КФС c в дискретный момент времени t описывается с использованием предлагаемого в работе метода при помощи ряда отобранных согласно работе [8] наиболее информативных признаков:

$$c = \mu(a_1(x_{t,1}, x_{t,2}, \dots, x_{t,S}), a_2(x_{t,1}, x_{t,2}, \dots, x_{t,S}), \dots, a_n(x_{t,1}, x_{t,2}, \dots, x_{t,S})), \\ c \in C, x_{t,i} \in D_f, s \leq S,$$

где C — множество рассматриваемых состояний КФС; μ — агрегирующая функция; $a_1, a_2, \dots, a_n$ — классифицирующие алгоритмы; $x_{t,i}$ — значения признаков в дискретный момент времени; D_f — множество допустимых значений признаков; s — количество отобранных наиболее информативных признаков; S — общее количество доступных признаков.

Цель работы — разработка метода, обеспечивающего увеличение полноты и точности идентификации состояния информационно-функциональной безопасности КФС за счет использования в системе мониторинга значений временных рядов за предшествующие моменты времени с применением весовых коэффициентов значимости.

Метод решения поставленной задачи

В работе применен и исследован алгоритм на основе деревьев решений, который относится к группе логических классификаторов [10]. Суть алгоритма заключается в построении бинарного дерева, во внутренних узлах которого располагаются предикаты, а в листах — метки классов c_i ($i = 1, \dots, l$). Корень деревьев решений содержит узел принятия решения на основе анализа наиболее информативного признака, по мере удаления и при дальнейшем построении деревьев решений используются менее информативные признаки (по убыванию информативности). Листовые вершины содержат значения классов, определенных на этапе формирования обучающей выборки. Выбор предикатов осуществлен с помощью критериев информативности [11].

В бинарном дереве решений:

- каждой внутренней вершине v приписана функция (или предикат) $\beta_v: \mathbf{X} \rightarrow \{0, 1\}$;
- каждой листовой вершине v приписан прогноз — одно из l возможных состояний КФС $c_i \in C$.

Предикаты β_v сравнивают значение одного из признаков с порогом τ_v :

$$\beta_v(x; j, \tau_v) = [x_j < \tau_v].$$

Алгоритм $a(x)$, начиная с корневой вершины v_0 , вычисляет значение функции β_{v_0} . Если оно равно нулю, то алгоритм переходит в левую дочернюю вершину, иначе в правую, после чего вычисляет значение предиката в новой вершине и делает переход или влево, или вправо. Процесс продолжается, пока не будет достигнута листовая вершина; алгоритм возвращает тот класс, который приписан этой вершине.

Таким образом, подавая на вход исходные значения в момент времени t , построенный алгоритм $a(x)$ на основе обучающего набора формирует ответ — одну из меток класса $c_i \in C$, ассоциированную с состоянием КФС.

КФС — сложная система, каждый элемент которой может подвергаться отдельным видам деструктивных воздействий и атак. Данные, поступающие от различных элементов КФС, могут обладать индивидуальными свойствами. В связи с этим возникает задача идентификации состояния для классификаторов, обладающих своими компетенциями на подвыборках.

Временной ряд X состоит из значений признаков в m моментах. По методу бутстрэпа (bootstrap) из всего множества равновероятно выбирается m/n векторов признаков, каждый из которых соответствует определенному моменту времени. Отметим, что из-за возвращения некоторые элементы в подмножества могут повторяться. Обозначим новую выборку через X_1 . Повторяя процедуру n раз, сгенерируем n подвыборок $X_1, X_2, \dots, X_n$.

Применим n параллельно работающих классификаторов для увеличения скорости получения итоговых результатов. Используем агрегирующую функцию μ для улучшения стабильности и точности алгоритмов. Принципы построения ансамбля параллельно работающих классификаторов:

- генерация с помощью бутстрэпа n выборок размерностью $m/n \times s$ для каждого классификатора a_1-a_n ;
- независимое обучение каждого элементарного классификатора (алгоритм a_1-a_n , определенный на своем подпространстве) на заранее размеченном наборе данных (обучение с учителем);
- независимая классификация каждой подвыборки $X_1, X_2, \dots, X_n$ на каждом из подпространств;
- принятие окончательного решения о принадлежности объекта одному из состояний.

В классических подходах при использовании ансамблей классификаторов окончательное решение о принадлежности элементов временного ряда определенному состоянию КФС принимается одним из следующих методов.

- Консенсус: если все элементарные классификаторы присвоили одну и ту же метку множеству значений признаков в момент времени t , то такой объект будет отнесен к выбранному классу. Консенсус достигим не всегда.
- Простое большинство: объекту присваивается метка того класса, который определило для него большинство элементарных классификаторов.

При использовании «моментальных снимков» в дискретный момент времени не всегда достигается необходимая полнота и точность идентификации состояния КФС. Кроме того, на практике наиболее важными являются значения состояний, которые приближены к текущему моменту времени [9]. Для увеличения полноты и точности предлагается ввести временной отрезок идентификации N , представляющий собой скользящее окно от t_{i-N+1} до текущего момента времени t_i , и весовые коэффициенты p_i , учитывающие степень предпочтения одних результатов идентификации другим (рис. 2).

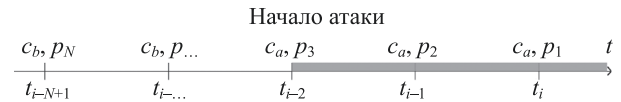


Рис. 2. Временной график идентификации атаки с учетом коэффициентов значимости

Fig. 2. Time schedule of attack identification which takes into account the coefficients of significance

В рассматриваемом случае c_b — безопасное состояние КФС; c_a — состояние, в котором КФС находится под атакующим информационным или физическим воздействием. Весовые коэффициенты должны удовлетворять следующим требованиям:

- учитывать временной отрезок идентификации N , представляющий собой значение величины временного интервала, за который анализируемые данные подаются на вход алгоритма, реализующего предложенный метод;
- любой коэффициент p_{i+1} должен быть меньше p_i ($\forall p_{i-1} < p_i, i \in [1, N]$).

Для системы убывающего предпочтения, состоящей из N альтернатив, предложено использовать весовые коэффициенты, снижающиеся по правилу арифметической прогрессии. Весовые коэффициенты Фишберна — рациональные дроби, в числителе которых расположены убывающие на единицу элементы натурального ряда от N до 1, а в знаменателе — сумма арифметической прогрессии N первых членов натурального ряда с шагом 1.

$$r_1 = N, r_i = r_{i-1} - 1, K = \sum_{i=1}^N r_i, p_i = \frac{r_i}{K}, \quad (1)$$

где r_i — элементы натурального ряда; K — сумма арифметической прогрессии N первых членов натурального ряда с шагом 1; p_i — весовой коэффициент значимости результата идентификации для дискретного момента времени; N — временной отрезок идентификации.

Апробация метода

С целью практической реализации разработанного метода идентификации состояния элементов КФС на основе анализа временных рядов выполним расчетный эксперимент над набором данных [12]. Исследователи из Сингапурского университета технологии и дизайна (Singapore University of Technology and Design) смоделировали различные типы атак на компоненты экспериментального стенда КФС водоочистки. Для исследования выбраны атаки, которые происходили: на разных уровнях КФС; включали в себя один или несколько этапов; имели различную продолжительность и затрагивали разные уровни КФС; большинство атак имели воздействие на технологический процесс.

В табл. 1 представлены типы и количество атак (общее количество — 41).

Из полученного сетевого трафика между системой SCADA (Supervisory Control And Data Acquisition) и программируемыми логическими контроллерами сформированы временные ряды, содержащие информацию со всех доступных датчиков КФС. Для анализа вре-

Таблица 1. Типы атак на киберфизические системы водоочистки

Table 1. Types of attacks on cyber-physical systems of water treatment

Типы атак	Количество атак
Одностадийные и однонаправленные	26
Одностадийные многонаправленные	4
Многостадийные однонаправленные	2
Многостадийные и многонаправленные	4
Без воздействия на технологический процесс	5

менных рядов, характеризующих функционирование КФС, применено программное обеспечение MATLAB R2022a. Формирование деревьев решений, их ансамблей и последующая классификация выполнена при помощи приложения Classification Learner App.

Исходные данные для реализации разработанного метода представляют численный двумерный массив $944\,919 \times 51$. В строках расположены значения временного ряда, регистрируемые раз в секунду, а столбцы упорядочены по источникам получения информации от КФС. Набор данных размечен по классам (состояниям), каждому временному ряду поставлена в соответствие метка класса $\{c_1, \dots, c_{41}\}$.

В начале процесса оценивания состояния элементов КФС формируется обучающая выборка из временных рядов, составленных из значений параметров функционирования КФС и соответствующих им меток состояний (классов). Методом равномерного случайного сэмпинга с возвратом формируется n подвыборок.

Алгоритмы a_1-a_n на основе деревьев решений обучаются каждый на своей подвыборке независимо друг от друга, классификаторы не исправляют ошибки друг друга, а компенсируют их при голосовании. Базовые классификаторы в таком случае являются независимыми за счет обучения на различных подвыборках. На стадии идентификации состояния анализируемые показатели за отрезок времени N , представляющие собой кортеж $\mathbf{X}^* = \{\{x_1(t_{i-N+1}), x_2(t_{i-N+1}), \dots, x_s(t_{i-N+1})\}, \dots, \{x_1(t_{i-1}), x_2(t_{i-1}), \dots, x_s(t_{i-1})\}, \{x_1(t_i), x_2(t_i), \dots, x_s(t_i)\}\}$, подаются на вход вышеуказанных классифицирующих алгоритмов.

Каждый алгоритм a_1-a_n генерирует N ответов на заданном временном отрезке, которые постобрабатываются на первом этапе с использованием весовых коэффициентов Фишберна, придающих больший вес более поздним результатам идентификации. Таким образом, метка класса определяется путем взвешенного обобщения результатов классификации на отрезке времени N . Весовые коэффициенты p_i для различных моментов времени суммируются в том случае, если предсказанные классы совпадают. В табл. 2 приведен пример применения рассчитанных по формуле (1) коэффициентов для классификатора a_1 .

Сумма коэффициентов для c_{37} равна $\frac{5}{15}$, $c_{35} - \frac{4}{15}$, $c_{18} - \frac{3}{15}$. С учетом весовых коэффициентов значи-

Таблица 2. Пример применения весовых коэффициентов для $N = 5$ Table 2. An example of applying weigh coefficients for $N = 5$

Время	Весовой коэффициент	Метка класса состояния
t_i	$\frac{5}{15}$	c_{37}
t_{i-1}	$\frac{4}{15}$	c_{35}
t_{i-2}	$\frac{3}{15}$	c_{35}
t_{i-3}	$\frac{2}{15}$	c_{18}
t_{i-4}	$\frac{1}{15}$	c_{18}

мости результатом будет метка класса c_{35} . В случае равенства коэффициентов перед классами выберем тот класс, который был определен для более позднего момента времени.

Окончательное решение примем за счет обобщения результатов, полученных на предыдущем этапе и по каждому классификатору a_1-a_n . Итоговый результат определим простым большинством результатов. Применим нечетные n , чтобы избежать случаи равного числа голосов для отличающихся классов c . Блок-схема алгоритма представлена на рис. 3.

Полученные результаты

Разработанный метод за счет комбинированного подхода позволил, не увеличивая числа отобранных признаков, существенно повысить показатели качества классификации и быстроту реагирования на инциденты ИБ и физические воздействия на КФС. На рис. 4 в сравнении показаны результаты применения предложенного подхода и стандартного метода, не учитывающего степень предпочтения результатов идентификации в различные дискретные моменты времени. При $N = 1$ решение задачи сводится к определению состояния КФС в дискретный момент времени. Применение разработанного подхода позволило повысить F-меру по сравнению со стандартным методом. Исходя из анализа гистограммы сделаем вывод, что оптимальное значение — величина отрезка идентификации $N = 4$. Отметим, что дальнейшее увеличение отрезка практически не приводит к увеличению F-меры. Решающий фактор для выбора величины скользящего окна N и количества классификаторов n — максимизация F-меры результатов классификации для контролируемой КФС.

Выполним сравнение результатов настоящего исследования с данными, полученными в научных работах [13–19]. Сравнение проведено на идентичном наборе исходных данных, выделены примененные другими исследователями методы (табл. 3).

На рис. 5, а представлена диаграмма точности (precision). Как видно, точность идентификации состо-

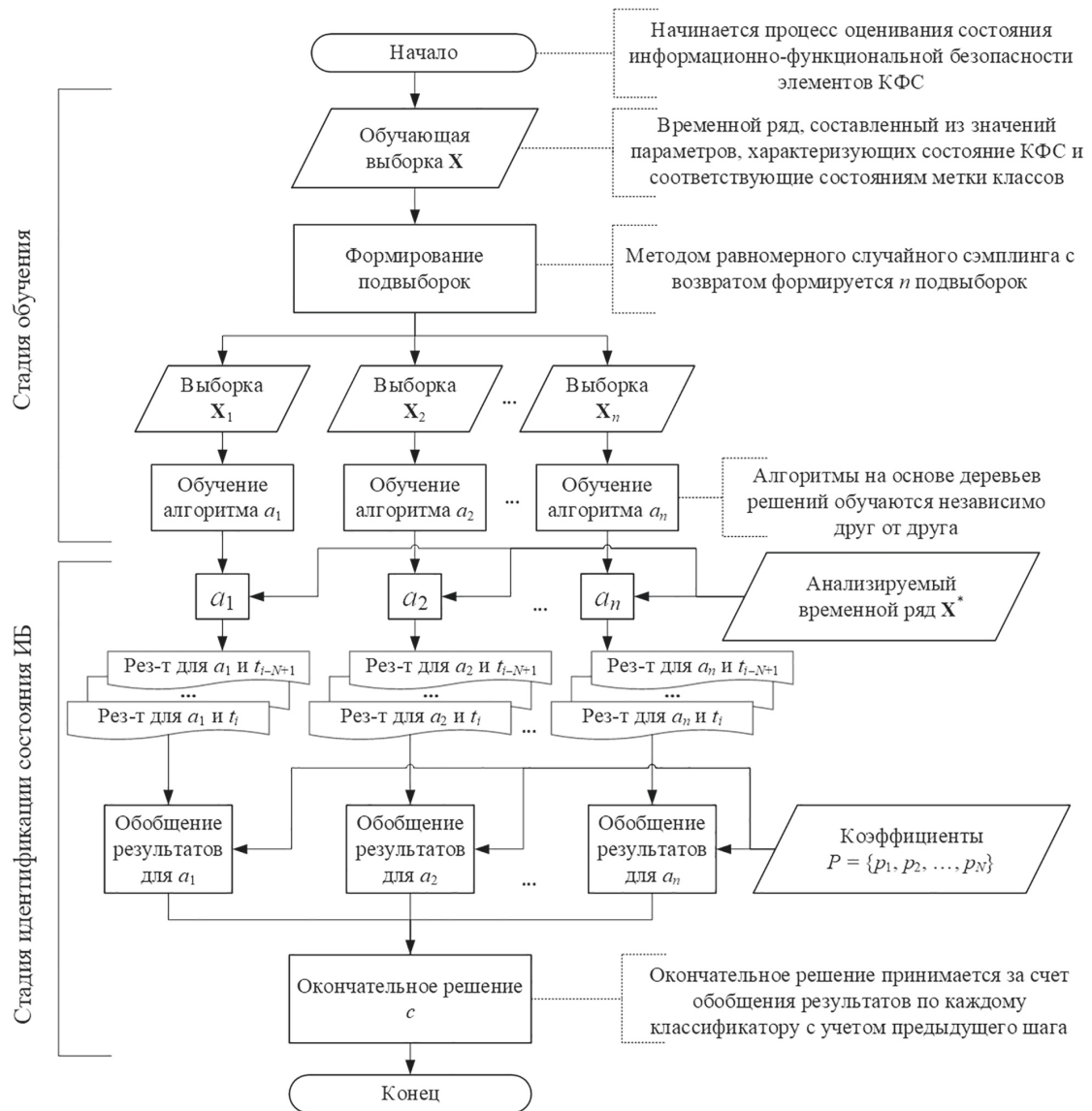


Рис. 3. Блок-схема алгоритма идентификации состояния киберфизических систем

Fig. 3. Block diagram of the algorithm for identifying the state of cyber-physical systems state

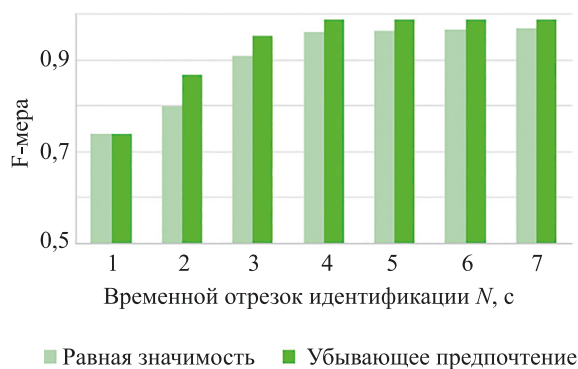


Рис. 4. F-мера при варьировании временного отрезка идентификации

Fig. 4. F-measure under varying the time interval of identification

яния элементов КФС с применением разработанного метода существенно выше, чем в работах других исследователей, применивших иные по природе классификаторы и методы предварительной обработки данных.

Разработанный метод позволил также повысить полноту (recall) идентификации состояния КФС (рис. 5, b). Методики с высокой полнотой классификации предпочтительнее для распознавания ранее неизвестных типов аномалий [20].

Таким образом, проанализированные работы других исследователей характеризуются относительно низкой полнотой идентификации состояний КФС, что является их существенным недостатком, поскольку такие модели могут идентифицировать значительное число атак на КФС как безопасные состояния.

Таблица 3. Исследования, с которыми произведено сравнение результатов

Table 3. Studies against which the results were compared

Авторы исследований	Применяемый метод	Обозначение	Источник
Kravchik M., Shabtai A.	Одномерные сверточные нейронные сети (One-Dimensional Convolutional Neural Networks)	1D CHC (1D CNN)	[13]
Shalyga D., Filonov P., Lavrentyev A.	Многослойный перцептрон (Multilayered Perceptron)	МП (MLP)	[14]
	Сверточные нейронные сети (Convolutional Neural Networks)	CHC (CNN)	
	Рекуррентные нейронные сети (Recurrent Neural Networks)	PHC (RNN)	
Inoue J., Yamagata Y., Chen Y., Poskitt C.M., Sun J.	Глубинные Нейронные Сети (Deep Neural Networks)	ГHC (DNN)	[15]
	Одноклассовый метод опорных векторов (One-Class Support Vector Machines)	MOB (OCSVM)	
Kravchik M., Shabtai A.	Автоэнкодер (Autoencoder)	АК (AE)	[16]
Elnour M., Meskin N., Khan K., Jain R.	Изолирующие леса (Isolation Forests)	ИЛ (IF)	[17]
Li D., Chen D., Jin B., Shi L., Goh J., Ng S.K.	Генеративно-сопоставительные сети (Generative Adversarial Networks)	ГCHC (GAN)	[18]
Gomez A., Maimo L., Celdran A., Clemente F.	Нейронные сети с долгой краткосрочной памятью (Long Short-Term Memory Neural Networks)	LSTM ИHC (LSTM NN)	[19]

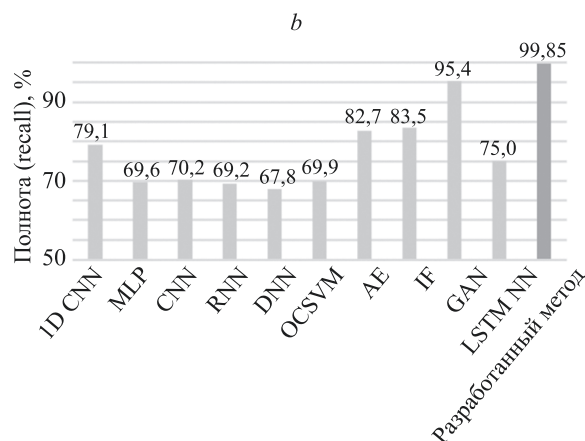
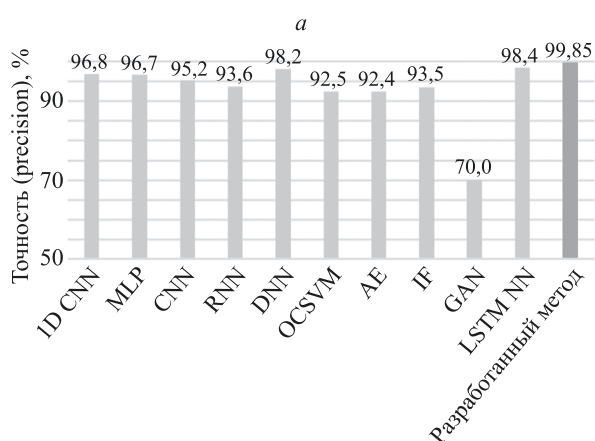


Рис. 5. Сравнение точности (a) и полноты (b) идентификации состояния киберфизических систем

Fig. 5. Comparison of precision (a) and recall (b) of identification of the cyber-physical systems state

Заключение

В работе представлен метод обработки данных мониторинга состояния информационно-функциональной безопасности киберфизических систем на основе анализа временных рядов с применением весовых коэффициентов значимости. Метод имеет важное значение для совершенствования средств обеспечения аудита и мониторинга состояния объекта, находящегося под воздействием угроз нарушения его информационной и функциональной безопасности, а также расследования инцидентов информационной безопасности в автоматизированных информационных системах.

Предложенный метод позволил существенно повысить скорость идентификации состояния элементов

киберфизических систем за счет применения ансамбля параллельно работающих классификаторов. Ошибки из-за случайных отклонений параметров функционирования киберфизических систем сокращаются с помощью обобщения и взвешенного усреднения результатов идентификации на временном отрезке. Точность идентификации, по сравнению с лучшими подходами, представленными в научных работах, при использовании разработанного метода увеличилась на 1,45 %, полнота — на 4,45 % и составила 99,85 % для обоих показателей. Снижение вычислительных затрат и увеличение скорости идентификации состояния элементов киберфизических систем — решающие факторы при осуществлении мониторинга и восстановлении безопасного функционирования.

Разработанный метод является новой альтернативой и дополнением к существующим программным и программно-аппаратным средствам. В качестве дальнейших перспектив исследования можно отметить раз-

работку методов и методик противодействия выявленным нарушениям на основе принципа обратной связи в режиме реального времени.

Литература

1. Shukalov A.V., Zakoldaev D.A., Zharinov I.O., Zharinov O.O. Control, computing and communication in industrial cyberphysical systems with feedback // *Journal of Physics: Conference Series*. 2021. V. 2094. N 4. P. 042036. <https://doi.org/10.1088/1742-6596/2094/4/042036>
2. Котенко И.В., Крибель А.М., Лаута О.С., Саенко И.Б. Анализ процесса самоподобия сетевого трафика как подход к обнаружению кибератак на компьютерные сети // *Электросвязь*. 2020. № 12. С. 54–59. <https://doi.org/10.34832/ELSV.2020.13.12.008>
3. Васильев В.И., Вульфин А.М., Гвоздев В.Е., Картак В.М., Атарская Е.А. Обеспечение информационной безопасности киберфизических объектов на основе прогнозирования и обнаружения аномалий их состояния // *Системы управления, связи и безопасности*. 2021. № 6. С. 90–119. <https://doi.org/10.24412/2410-9916-2021-6-90-119>
4. Зегжда Д.П., Павленко Е.Ю. Гомеостатическая стратегия безопасности киберфизических систем // *Проблемы информационной безопасности. Компьютерные системы*. 2017. № 3. С. 9–23.
5. Зайцева Е.А., Зегжда Д.П., Полтавцева М.А. Использование графового представления и прецедентного анализа для оценки защищенности компьютерных систем // *Проблемы информационной безопасности. Компьютерные системы*. 2019. № 2. С. 136–148.
6. Lavrova D.S. An approach to developing the SIEM system for the Internet of Things // *Automatic Control and Computer Sciences*. 2016. V. 50. N 8. P. 673–681. <https://doi.org/10.3103/S0146411616080125>
7. Васильев Ю.С., Зегжда Д.П., Зегжда Д.П. Обеспечение безопасности автоматизированных систем управления технологическими процессами на объектах гидроэнергетики // *Известия Российской академии наук. Энергетика*. 2016. № 3. С. 49–61.
8. Семенов В.В. Подход к формированию информативных признаков в задачах мониторинга информационной безопасности киберфизических систем // *Научно-технический вестник информационных технологий, механики и оптики*. 2021. Т. 21. № 6. С. 887–894. <https://doi.org/10.17586/2226-1494-2021-21-6-887-894>
9. Семенов В.В. Оценивание состояния информационной безопасности на основе анализа временных рядов // *Научно-технический вестник Поволжья*. 2021. № 10. С. 127–129.
10. Kruegel C., Toth T. Using decision trees to improve signature-based intrusion detection // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2003. V. 2820. P. 173–191. https://doi.org/10.1007/978-3-540-45248-5_10
11. Cagli E., Dumas C., Prouff E. Convolutional neural networks with data augmentation against jitter-based countermeasures // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2017. V. 10529. P. 45–68. https://doi.org/10.1007/978-3-319-66787-4_3
12. Goh J., Adepu S., Junejo K.N., Mathur A. A dataset to support research in the design of secure water treatment systems // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2017. V. 10242. P. 88–99. https://doi.org/10.1007/978-3-319-71368-7_8
13. Kravchik M., Shabtai A. Detecting cyber attacks in industrial control systems using convolutional neural networks // *Proc. of the 47th Workshop on Cyber-Physical Systems Security and Privacy*. 2018. P. 72–83. <https://doi.org/10.1145/3264888.3264896>
14. Shalyga D., Filonov P., Lavrentyev A. Anomaly detection for water treatment system based on neural network with automatic architecture optimization // *arXiv*. 2018. arXiv:1807.07282. <https://doi.org/10.48550/arXiv.1807.07282>
15. Inoue J., Yamagata Y., Chen Y., Poskitt C.M., Sun J. Anomaly detection for a water treatment system using unsupervised machine learning // *Proc. of the 17th IEEE International Conference on Data Mining Workshops (ICDMW)*. 2017. P. 1058–1065. <https://doi.org/10.1109/ICDMW.2017.149>

References

1. Shukalov A.V., Zakoldaev D.A., Zharinov I.O., Zharinov O.O. Control, computing and communication in industrial cyberphysical systems with feedback. *Journal of Physics: Conference Series*, 2021, vol. 2094, no. 4, pp. 042036. <https://doi.org/10.1088/1742-6596/2094/4/042036>
2. Kotenko I.V., Kribel A.M., Laut O.S., Saenko I.B. Analysis of the process of selfsimilarity of network traffic as an approach to detecting cyber attacks on computer networks. *Electrosvyaz Magazine*, 2020, no. 12, pp. 54–59. (in Russian). <https://doi.org/10.34832/ELSV.2020.13.12.008>
3. Vasilyev V.I., Vulfin A.M., Gvozdev V.E., Kartak V.M. Atarskaya E.A. Ensuring information security of cyber-physical objects based on predicting and detecting anomalies in their state. *Systems of Control, Communication and Security*, 2021, no. 6, pp. 90–119. (in Russian). <https://doi.org/10.24412/2410-9916-2021-6-90-119>
4. Zegzhda D.P., Pavlenko E.Y. Homeostatic security of cyber-physical systems. *Information Security Problems. Computer Systems*, 2017, no. 3, pp. 9–23. (in Russian)
5. Zaitceva E.A., Zegzhda D.P., Poltavtseva M.A. Applying of graph representation and case-based reasoning for security evaluation of computer systems. *Information Security Problems. Computer Systems*, 2019, no. 2, pp. 136–148. (in Russian)
6. Lavrova D.S. An approach to developing the SIEM system for the Internet of Things. *Automatic Control and Computer Sciences*, 2016, vol. 50, no. 8, pp. 673–681. <https://doi.org/10.3103/S0146411616080125>
7. Vasiliev Y.S., Zegzhda P.D., Zegzhda D.P. Providing security for automated process control systems at hydropower engineering facilities. *Thermal Engineering*, 2016, vol. 63, no. 13, pp. 948–956. <https://doi.org/10.1134/S0040601516130073>
8. Semenov V.V. An approach to the identification of the state of elements in cyber-physical systems based on principal component analysis. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2021, vol. 21, no. 6, pp. 887–894. (in Russian). <https://doi.org/10.17586/2226-1494-2021-21-6-887-894>
9. Semenov V.V. Assessment of information security state based on analysis of time series. *Scientific and Technical Volga region Bulletin*, 2021, no. 10, pp. 127–129. (in Russian)
10. Kruegel C., Toth T. Using decision trees to improve signature-based intrusion detection. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2003, vol. 2820, pp. 173–191. https://doi.org/10.1007/978-3-540-45248-5_10
11. Cagli E., Dumas C., Prouff E. Convolutional neural networks with data augmentation against jitter-based countermeasures. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10529, pp. 45–68. https://doi.org/10.1007/978-3-319-66787-4_3
12. Goh J., Adepu S., Junejo K.N., Mathur A. A dataset to support research in the design of secure water treatment systems. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10242, pp. 88–99. https://doi.org/10.1007/978-3-319-71368-7_8
13. Kravchik M., Shabtai A. Detecting cyber attacks in industrial control systems using convolutional neural networks. *Proc. of the 47th Workshop on Cyber-Physical Systems Security and Privacy*, 2018, P. 72–83. <https://doi.org/10.1145/3264888.3264896>
14. Shalyga D., Filonov P., Lavrentyev A. Anomaly detection for water treatment system based on neural network with automatic architecture optimization. *arXiv*, 2018, arXiv:1807.07282. <https://doi.org/10.48550/arXiv.1807.07282>
15. Inoue J., Yamagata Y., Chen Y., Poskitt C.M., Sun J. Anomaly detection for a water treatment system using unsupervised machine learning. *Proc. of the 17th IEEE International Conference on Data Mining Workshops (ICDMW)*, 2017, P. 1058–1065. <https://doi.org/10.1109/ICDMW.2017.149>

16. Kravchik M., Shabtai A. Efficient cyber attack detection in industrial control systems using lightweight neural networks and PCA // *IEEE Transactions on Dependable and Secure Computing*. 2022. V. 19. N 4. P. 2179–2197. <https://doi.org/10.1109/TDSC.2021.3050101>
17. Elnour M., Meskin N., Khan K., Jain R. A dual-isolation-forests-based attack detection framework for industrial control systems // *IEEE Access*. 2020. V. 8. P. 36639–36651. <https://doi.org/10.1109/ACCESS.2020.2975066>
18. Li D., Chen D., Jin B., Shi L., Goh J., Ng S.-K. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2019. V. 11730. P. 703–716. https://doi.org/10.1007/978-3-030-30490-4_56
19. Gómez A., Maimó L., Celdrán A., Clemente F. MADICS: A methodology for anomaly detection in industrial control systems // *Symmetry*. 2020. V. 12. N 10. P. 1583. <https://doi.org/10.3390/sym12101583>
20. Гайфулина Д.А., Котенко И.В. Анализ моделей глубокого обучения для задач обнаружения сетевых аномалий Интернета вещей // *Информационно-управляющие системы*. 2021. № 1. С. 28–37. <https://doi.org/10.31799/1684-8853-2021-1-28-37>
16. Kravchik M., Shabtai A. Efficient cyber attack detection in industrial control systems using lightweight neural networks and PCA. *IEEE Transactions on Dependable and Secure Computing*, 2022, vol. 19, no. 4, pp. 2179–2197. <https://doi.org/10.1109/TDSC.2021.3050101>
17. Elnour M., Meskin N., Khan K., Jain R. A dual-isolation-forests-based attack detection framework for industrial control systems. *IEEE Access*, 2020, vol. 8, pp. 36639–36651. <https://doi.org/10.1109/ACCESS.2020.2975066>
18. Li D., Chen D., Jin B., Shi L., Goh J., Ng S.-K. MAD-GAN: Multivariate anomaly detection for time series data with generative adversarial networks. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2019, vol. 11730, pp. 703–716. https://doi.org/10.1007/978-3-030-30490-4_56
19. Gómez A., Maimó L., Celdrán A., Clemente F. MADICS: A methodology for anomaly detection in industrial control systems. *Symmetry*, 2020, vol. 12, no. 10, pp. 1583. <https://doi.org/10.3390/sym12101583>
20. Gaifulina D.A., Kotenko I.V. Analysis of deep learning models for network anomaly detection in Internet of Things. *Informatsionno-Upravliaiushchie Sistemy*, 2021, no. 1, pp. 28–37. (in Russian). <https://doi.org/10.31799/1684-8853-2021-1-28-37>

Автор

Семенов Виктор Викторович — кандидат технических наук, младший научный сотрудник, Санкт-Петербургский Федеральный исследовательский центр Российской академии наук, Санкт-Петербург, 199178, Российская Федерация, [sc 57204123255](https://orcid.org/0000-0002-7216-769X), <https://orcid.org/0000-0002-7216-769X>, v.semenov@spcras.ru

Author

Viktor V. Semenov — PhD, Junior Researcher, St. Petersburg Federal Research Center of the Russian Academy of Sciences, Saint Petersburg, 199178, Russian Federation, [sc 57204123255](https://orcid.org/0000-0002-7216-769X), <https://orcid.org/0000-0002-7216-769X>, v.semenov@spcras.ru

Статья поступила в редакцию 20.06.2022
Одобрена после рецензирования 03.10.2022
Принята к печати 22.11.2022

Received 20.06.2022
Approved after reviewing 03.10.2022
Accepted 22.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1159-1165

УДК 004.9

Применение волновой модели текста к задаче сентимент-анализа

Анастасия Сергеевна Груздева^{1✉}, Родион Николаевич Юрьев²,
Игорь Александрович Бессмертный³

^{1,2,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ prog.anastasia@gmail.com[✉], <https://orcid.org/0000-0003-4963-0823>

² rodion@juryev.ru, <https://orcid.org/0000-0003-1146-2617>

³ bessmertny@itmo.ru, <https://orcid.org/0000-0001-6711-6399>

Аннотация

Предмет исследования. Исследована волновая модель представления коротких текстов на русском языке. Модель является одной из реализаций дистрибутивной семантики. В модели учтены не только частоты встречаемости слов в тексте, но и их взаимное влияние. Итогом реализации модели служит повышение точности анализа тональности коротких текстов. **Метод.** Основу определения взаимосвязей между текстом и термином составляет расчет амплитуды вероятности близости текста к термину с использованием волновой модели. Термин, обладающий наибольшей амплитудой вероятности, считается наиболее точно соответствующим смыслу текста. Волновая модель позволяет учесть тот факт, что известные методы определяют антонимы как семантически близкие лексические единицы. **Основные результаты.** Для экспериментального исследования предложенного метода выбрано решение задачи сентимент-анализа, то есть нахождения соответствия отзывов пользователей о покупках на маркетплейсе классам «позитивный» и «негативный». В результате эксперимента получена точность оценки тональности текста до 76,4 %, что превышает точность классического подхода, а также известных методик сентимент-анализа для русского языка. Выявлено значительное влияние таких параметров модели, как выбор базисной дистрибутивно-семантической модели, выбор контрольной точки для расчета волновых чисел, учет влияния антонимов на точность классификации. Представленная модель показала высокую точность в выявлении взаимосвязей текста с не присутствующими в нем явно понятиями. **Практическая значимость.** Предложенный метод может успешно применяться как математическая основа решения задач сентимент-анализа. Полученные результаты показали потенциальные возможности использования волновой модели в областях, требующих классификации текстов по косвенным признакам, например, для определения элементов психологического портрета автора.

Ключевые слова

сентимент-анализ, классификация, обработка естественного языка, волновая модель, квантово-подобная модель

Благодарности

Работа выполнена в рамках магистерско-аспирантской НИР № 620164 «Методы искусственного интеллекта для киберфизических систем».

Ссылка для цитирования: Груздева А.С., Юрьев Р.Н., Бессмертный И.А. Применение волновой модели текста к задаче сентимент-анализа // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1159–1165. doi: 10.17586/2226-1494-2022-22-6-1159-1165

Application of the text wave model to the sentiment analysis problem

Anastasia S. Gruzdeva^{1✉}, Rodion N. Iurev², Igor A. Bessmertny³

^{1,2,3} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ prog.anastasia@gmail.com[✉], <https://orcid.org/0000-0003-4963-0823>

² rodion@juryev.ru, <https://orcid.org/0000-0003-1146-2617>

³ bessmertny@itmo.ru, <https://orcid.org/0000-0001-6711-6399>

Abstract

Authors researched the wave model of text representation which is one of the implementations of distributive semantics. This model takes into account not only the frequency of words occurrence in the text, but also their mutual location.

© Груздева А.С., Юрьев Р.Н., Бессмертный И.А., 2022

The purpose of the study: to increase the accuracy of the analysis of the tonality of short texts based on the wave model. The method of determining the relationship between the text and the term is based on the calculation of the probability amplitude of the text and term proximity using a wave model. The term with the highest probability amplitude is considered to correspond most closely to the meaning of the text. The wave model allowed taking into account the fact that well-known methods define antonyms as semantically close lexical units. For the experimental study of this technique, a solution to the problem of sentiment analysis was chosen, exactly, finding the correspondence of user reviews about the product to the classes “positive” and “negative”. As a result, the accuracy of the text tonality defining was obtained up to 76.4 %, which exceeds the accuracy of the classical approach as well as the well-known methods of sentiment analysis for the Russian language. In addition, authors detected significant influence on classification accuracy of such model parameters as the choice of a basic distributive semantic model, the choice of a control point for calculating wave numbers, taking into account the influence of antonyms. The presented model has shown high accuracy in identifying the relationships of the text with concepts that are not explicitly present in it and can be successfully used as a mathematical basis for solving problems of sentiment analysis. In addition, the results obtained indicate the potential use of the wave model in other areas that require the classification of texts by indirect signs, for example, to determine the elements of author psychological portrait.

Keywords

sentiment analysis, classification, natural language processing, wave model, quantum-like model

Acknowledgements

The work was carried out within the framework of the project no. 620164 (Artificial intelligence methods for cyberphysical systems).

For citation: Gruzdeva A.S., Iurev R.N., Bessmertny I.A. Application of the text wave model to the sentiment analysis problem. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1159–1165 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1159-1165

Введение

Проблемы современных подходов к сентимент-анализу связаны с высокой размерностью данных и неизбежно возникающей неоднозначностью в текстах. Результативность техник сентимент-анализа ограничена в настоящее время для английского языка точностью около 70–80 % в зависимости от применяемого инструментария [1]. Для русского языка точность распознавания продемонстрирована в отдельных исследованиях от 60 до 73 % [2]. В то же время востребованность сентимент-анализа в мировом и российском сегментах растет с каждым годом, что показывает статистика запросов на популярном ресурсе Google Trends¹. График роста числа запросов за последние годы показан на рисунке.

Цель работы — совершенствование методов сентимент-анализа для выявления элементов онтологий в поведении пользователей

Состояние проблемы и постановка задачи

Сентимент-анализ — одна из популярных задач компьютерной лингвистики. Первые методы сентимент-анализа базировались на размеченных словарях оценочных слов [3, 4]. Применение словарей позволяет не только упростить оценку тональности текста, но и скорректировать его при необходимости [5]. Ограничением является словарь, составление которого требует времени, а одни и те же слова в разных контекстах могут иметь разную эмоциональную окраску. Альтернатива словарному подходу — корпусная лингвистика, и, в частности, дистрибутивная гипотеза Харриса [6], согласно которой слова, часто встречающиеся в похожем контексте, имеют схожий смысл. В данном случае слова и тексты представлены векторами

в гильбертовом пространстве, а близость слов и текстов определены евклидовой мерой или косинусным расстоянием. Последние исследования в области сентимент-анализа описаны в работах [7–10], где общей проблемой является семантическая неразличимость антонимов.

Около 30 лет назад к анализу текстов стали применять квантовую теорию вероятностей [11–16]. Основное отличие квантовой теории от колмогоровской — наличие дополнительного параметра — фазы, за счет которой происходит интерференция амплитуд вероятностей. В связи с этим предложено использовать аппарат квантовой теории для классификации текстов, чтобы в отличие от известных методов, базирующихся на «мешке слов», учитывать взаимное расположение слов в тексте. В работе [17] разработана волновая модель, которая позволила учесть интерференцию при анализе текстов. Задачи сентимент-анализа часто связаны с необходимостью выявления неявных взаимосвязей и классификации текстов по косвенным признакам. В тексте могут отсутствовать явные указания на то, к какому эмоциональному полюсу — позитивному или негативному — он ближе. Тем не менее, набор используемых слов, строение предложений позволяют человеку понять общую эмоциональную тональность текста. Задача настоящей работы — проверить, может ли введение дополнительного параметра — фазы — оказать положительное влияние на точность классификации текстов по косвенным признакам.

Краткое описание волновой модели

Рассмотрим возможность применения квантово-подобной волновой модели представления текстовой информации для выявления тональности текста. Отметим основные моменты волновой модели, подробное описание которой приведено в работе [17].

В рамках волновой модели текст представлен в виде ансамбля элементарных частиц, где каждое сло-

¹ Google Trends. [Электронный ресурс]. URL: <https://trends.google.es/trends/explore?date=all&q=sentiment%20analysis> (дата обращения: 14.09.2022).

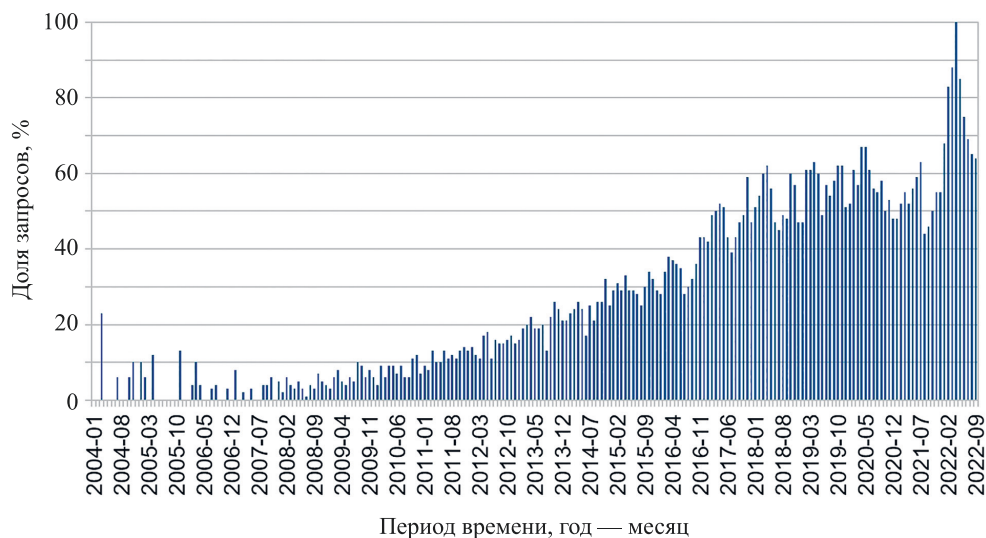


Рисунок. Диаграмма роста запросов «сентимент-анализа» с 2004 по 2022 годы (во всем мире). Чередование цветов использовано для улучшения зрительного восприятия диаграммы. Доля запросов рассчитана относительно максимального числа запросов в месяц за исследуемый период

Figure. Diagram of “sentiment analysis” requests growth from 2004 to 2022 (worldwide). The share of the requests is calculated relative to the maximum number of requests per month for the study period

во — отдельная частица. Поведение частицы описано волновой функцией, соответствующей сферической волне [18], интенсивность которой определяет амплитуду вероятности детектирования частицы в точке пространства. Поведение ансамбля элементарных частиц описано пакетом сферических волн, суммарная интенсивность которого отражает амплитуду вероятности обнаружения всего ансамбля в точках пространства. Такой принцип лежит в основе классификации текстов с использованием волновой модели. Формула расчета амплитуды вероятности принадлежности текста к некоторому классу l имеет вид [17]

$$I_l = \sum_{j=1}^M \left(\frac{A_j}{r_{lj}} \right)^2 + 2 \times \sum_{j=1}^{M-1} \sum_{n=j+1}^M \frac{A_j}{r_{lj}} \frac{A_n}{r_{ln}} \cos(k_j r_{lj} - k_n r_{ln} + \varphi_{j0} - \varphi_{n0}), \quad (1)$$

где I_l — интенсивность волнового пакета в точке пространства, соответствующей классу l ; M — число волн в пакете (число слов в тексте); A_j, A_n — число вхождений соответствующих слов в исходный текст; r_{lj} — семантическое расстояние между текстом l и словом j (величина, обратная семантической близости, рассчитанной на предобученной дистрибутивно-семантической языковой модели [19]); k_j, k_n — некоторые волновые числа.

В рамках используемых базисных моделей семантическая близость принимает значения в интервале от -1 до 1 . Семантические расстояния изменяются в диапазонах от 1 до $+\infty$ и от -1 до $-\infty$. Рассчитанная амплитуда вероятности может быть больше 1 , что не противоречит принципам квантовоподобных моделей.

Пренебрежем начальными фазами $\varphi_{j0}, \varphi_{n0}$ и для расчетов используем упрощенную модификацию уравнения (1):

$$I_l = \sum_{j=1}^M \left(\frac{A_j}{r_{lj}} \right)^2 + 2 \times \sum_{j=1}^{M-1} \sum_{n=j+1}^M \frac{A_j}{r_{lj}} \frac{A_n}{r_{ln}} \cos(k_j r_{lj} - k_n r_{ln}). \quad (2)$$

Исследуем применение нескольких моделей, предоставляемых интернет-ресурсом RusVectors¹: НКРЯ + Википедия (Национальный корпус русского языка + русскоязычная Википедия) за ноябрь 2021 года; Новостной корпус и корпус Тайга за 2019 год. Причина выбора данных моделей — наличие словарей достаточно большого объема (около 249 000 слов каждый). Кроме того, на портале RusVectors реализован программный интерфейс приложения, который упрощает обработку данных для анализа.

Для вычисления волновых чисел k_j, k_n используем следующие соображения. Найдем контрольную точку пространства, в которой суммарная интенсивность волнового пакета будет предположительно максимальной. Для достижения максимума интенсивности необходимо, чтобы фазы всех волн в данной точке были равны — волновые числа должны быть обратно пропорциональны расстояниям до контрольной точки. В работе [17] в качестве контрольной точки был использован центроид ансамбля частиц (текста). Здесь также применим расчет волновых чисел относительно центроида текста. Дополнительно, в качестве контрольной точки, используем заранее определенный контекст повествования, что позволит повысить точность классификации на 3,7 %.

В исходный алгоритм внесем изменение для коррекции проблемы высокой семантической близости антонимов в базисных дистрибутивно-семантических

¹ RusVectors [Электронный ресурс]. URL: <https://rusvectors.org/> (дата обращения: 05.05.2022).

моделях. Заметим, что часто пары противоположных по значению слов показывают высокие положительные величины семантической близости вместо ожидаемых значений, близких к -1 . Так, например, в модели НКРЯ+Википедия слова «белый» и «черный» обладают семантической близостью 0,732, а «плохой» и «хороший» — 0,770, т. е. они интерпретируются как синонимы, что противоречит человеческой логике. Очевидно, что высокая смысловая близость терминов, близких к понятиям «плохой» и «хороший», может особенно негативно сказаться на результатах анализа текста с целью определения отношения пользователя к качеству товара. Эта проблема отмечена во всех примененных в настоящей работе базисных дистрибутивно-семантических моделях. Предварительный анализ моделей одного из лучших проектов Russian Distributional Thesaurus [20] показал, что они также не лишены указанного недостатка. Для частичной компенсации проблемы синонимичности антонимов использован простой прием инверсии. Если пара слов обладает положительной семантической близостью и в то же время распознается как пара антонимов, знак семантической близости меняется на противоположный. Такой подход позволил увеличить точность классификации текстов.

Классификация отзывов пользователей

Для экспериментального исследования возможностей применения волновой модели в решении задач оценки общественного мнения отобраны комментарии пользователей, посвященные различным моделям смартфонов в интернет-магазине «Wildberries»¹. Комментарии распределены по двум классам: негативный и позитивный. Изначально распределение выполнено в соответствии с количеством «звезд», выставленных пользователями. Четыре и пять «звезд» расценивались как позитивный комментарий, одна и две — как негативный. Отзывы, отмеченные тремя «звездами», были отнесены к условному классу «нейтральный», не участвовавшему в классификации. В процессе классификации к «нейтральному» классу были присоединены комментарии, которые не удавалось отнести к позитивному или негативному классам в соответствии с заданными критериями. Изначальное распределение по классам пришлось пересмотреть, так как «звезды» выставлялись пользователями не столько за сам продукт, сколько за работу сервиса в целом, поэтому в целом позитивный комментарий могла сопровождать одна «звезда» из-за проблем, например, с доставкой. Окончательное распределение по классам было скорректировано в соответствии со смыслом отзывов. Всего было отобрано 100 позитивных отзывов (четыре и пять «звезд») и 100 негативных (одна и две «звезды»). Встретившиеся при отборе 16 отзывов, отмеченные тремя «звездами» были отнесены к «нейтральному» классу. В целом нейтральные отзывы наблюдались редко. Человек скорее всего не будет писать отзыв, если ему нечего сказать, а товар не произвел

ни положительного, ни отрицательного впечатления. После уточнения объективных классов в соответствии со смыслом текстов получено следующее распределение: 102 позитивных отзыва, 106 — негативных и 8 — нейтральных. Длина отзывов варьировалась от двух до 124 слов.

В процессе классификации выполнен расчет амплитуды вероятности принадлежности каждого отзыва к негативному и позитивному классам с использованием волновой модели. На основании полученных амплитуд вероятности для каждого класса, осуществлен расчет вероятности принадлежности текста:

$$P_l = \frac{I_l}{\sum_{i=1}^C I_i}, \quad (3)$$

где P_l — вероятность принадлежности текста к классу l ; I_l — амплитуда вероятности принадлежности текста к классу; C — количество классов. При предварительных расчетах обнаружено, что модель стабильно относит к позитивному классу почти в два раза больше отзывов, чем к негативному, при том, что объективно каждому из классов соответствовало примерно равное число комментариев. В связи с этим при использовании волновой модели классам присвоен понижающий коэффициент и учтен в расчете вероятностей. При этом уравнение (3) приняло вид

$$P_l = \frac{(1 - D_l) \times I_l}{\sum_{i=1}^C (1 - D_i) \times I_i},$$

где D_l — понижающий коэффициент класса l . Расчеты показали, что оптимальные значения понижающего коэффициента для позитивного класса располагаются в интервале от 0,4 до 0,6, для негативного — 0 при использовании волновой модели. Для классической модели (евклидова мера семантической близости (ЕМ)) применение понижающих коэффициентов не привело к повышению точности, поэтому для нее расчет вероятности выполнен по формуле (3). По результатам расчетов исследованный текст отнесен к классу l , при условии, что вероятность принадлежности к нему превысила пороговую величину 50,5 %. Если такой класс не обнаружен, то текст считается неклассифицированным и относится к «нейтральному» классу.

Выполним расчеты с использованием базисной модели НКРЯ + Википедия. В расчетах применим алгоритмы: расчет евклидовой меры семантической близости (ЕМ); волновая модель, в которой в качестве контрольной точки для расчета волновых чисел использован центр масс текста (ВМ ЦМТ); волновая модель, в которой волновые числа рассчитаны относительно контекста «смартфон» (ВМ К); волновая модель, в которой волновые числа рассчитаны относительно контекста «смартфон» с учетом инверсии антонимов (ВМ КИ). Полученная точность классификации приведена в табл. 1.

Наилучшую точность классификации показала волновая модель ВМ КИ, с использованием которой выполнен серию расчетов исследования влияния базисной

¹ Маркетплейс «Wildberries» [Электронный ресурс]. URL: <https://www.wildberries.ru/> (дата обращения: 22.05.2022).

Таблица 1. Точность классификации отзывов пользователей на базисной модели НКРЯ + Википедия, %
 Table 1. Precision of user reviews classification based on the distributive-semantic model NRC + Wikipedia, %

Класс	Алгоритмы			
	ЕМ	ВМ ЦМТ $D_n = 0,6$	ВМ К $D_n = 0,6$	ВМ КИ $D_n = 0,6$
Позитивный	78,4	77,5	68,6	70,6
Негативный	47,1	68,9	78,3	79,2
Нейтральный	25,0	12,5	0	0
Общий итог	61,1	70,8	70,8	72,2

Примечание. D_n — выбранные значения понижающего коэффициента позитивного класса

Таблица 2. Точность классификации для волновой модели с расчетом волновых чисел относительно контекста, %
 Table 2. Classification precision for wave model with wave number calculation based on context, %

Класс	Базисные модели			
	НКРЯ + Википедия $D_n = 0,6$	Новостной корпус $D_n = 0,45$	Корпус Тайга $D_n = 0,45$	НКРЯ + Википедия и Тайга $D_n = 0,45$
Позитивный	70,6	53,9	80,4	80,4
Негативный	79,2	87,8	68,9	78,3
Нейтральный	0	0	0	0
Общий итог	72,2	68,5	71,8	76,4

Примечание. D_n — выбранные значения понижающего коэффициента позитивного класса

дистрибутивно-семантической модели. В расчетах применим модели интернет-ресурсов: НКРЯ + Википедия, Новостной корпус, корпус Тайга и комбинацию моделей НКРЯ + Википедия и Тайга. В комбинации моделей для расчетов использовано среднее арифметическое семантических близостей. Точность классификации приведена в табл. 2.

В табл. 1 и 2 приведены наилучшие значения точностей, полученные для каждого алгоритма и модели. В рамках выбранных параметров волновой модели ВМ КИ не был верно распознан ни один из отзывов, относящихся к «нейтральному» классу. В данном исследовании не предпринимались никакие шаги для повышения точности определения данного класса, так как он одновременно и малочисленный, и малозначимый для целей sentiment-анализа. Тем не менее, вопрос об отсутствии верно распознанных данных на этом классе требует дополнительного осмысления и проверки, возможно, это дополнительно указывает на интерференционную природу самого исследуемого объекта.

Выводы

В результате работы волновая модель не только показала работоспособность, но и продемонстрировала превышение точности классификации по сравнению с традиционными моделями. Как показано в табл. 2, по отдельным типам классификации точность достигает 87,8 %, при этом в общем итоге средняя точность достигла 76,4 %, что на 3,4 % превышает аналогичные показатели традиционных моделей для русскоязычных текстов. Данные результаты получены благодаря простым математическим вычислениям, не требующим

повышенных вычислительных мощностей от пользователя, что является дополнительным преимуществом при внедрении новой системы на практике и позволяет осуществить распараллеливание процесса анализа. Алгоритмы, разработанные на базе волновой модели, могут стать хорошим дополнением к спектру существующих методик sentiment-анализа [7–10, 21–23].

Полученные экспериментальные данные позволяют сделать вывод о положительном влиянии учета интерференции на классификацию текстов по косвенным признакам. Как видно из уравнения (2) волновая модель отличается от евклидовой меры семантической близости только наличием члена, отвечающего за интерференцию, поэтому именно учет интерференции позволяет достичь заметного увеличения точности классификации.

Высокая точность, достигнутая в классификации по неявным признакам, позволяет допустить последующее расширение применимости волновой модели до любых других объектов исследования (например, в литературе представлены предложения по анализу музыкальных данных [24]).

Заключение

Безусловно, для полномасштабного применения волновой модели остаются ограничения, в первую очередь это касается правильного подбора вычислительного базиса, что во много зависит от искусства аналитика, работающего с материалом. Тем не менее, достижение вычислительной точности, продемонстрированной в работе, позволяет сделать еще один шаг в направлении решения задачи классификации текстов.

Литература

References

- Garg N., Sharma K. Text pre-processing of multilingual for sentiment analysis based on social network data // *International Journal of Electrical and Computer Engineering*. 2022. V. 12. N 1. P. 776–784. <https://doi.org/10.11591/ijece.v12i1.pp776-784>
- Корней А.О., Крючкова Е.Н. Проблемы эффективности сентимент-анализа русскоязычных текстов в социальных сетях // *Высокопроизводительные вычислительные системы и технологии*. 2018. Т. 2. № 2. С. 87–92.
- Пазельская А.Г., Соловьев А.Н. Метод определения эмоций в текстах на русском языке // *Тезисы докладов Международной конференции по компьютерной лингвистике и интеллектуальным технологиям «Диалог 2011»*. М.: РГТУ, 2011. С. 510–522.
- Клековкина М.В., Котельников Е.В. Метод автоматической классификации текстов по тональности, основанный на словаре эмоциональной лексики // *Электронные библиотеки: перспективные методы и технологии, электронные коллекции: материалы XIV Всероссийской научной конференции (RCDL-2012)*. 2012. С. 118–123.
- Бессмертный И.А., Джалишвили З.О., Максимов В.В., Маркин Д.А. Лингвооценочное управление текстом // *Тезисы докладов X Международной конференции «Применение новых технологий в образовании»*. Троицк: Фонд новых технологий в образовании «Байтик», 1999.
- Harris Z. *Mathematical Structures of Language*. Wiley, 1968. 230 p.
- Богданов А.Л., Дуля И.С. Сентимент-анализ коротких русскоязычных текстов в социальных медиа // *Вестник Томского государственного университета. Экономика*. 2019. № 47. С. 220–241. <https://doi.org/10.17223/19988648/47/17>
- Рябыкин А.С., Сухов Е.А. Нейросетевые методы в задаче сентимент-анализа // *DSPA*. 2022. Т. 2. № 2. С. 41–57.
- Воробьев А.А., Рыбак А.М., Середкин Р.А., Андросов А.Ю., Соловьев Б.И. Методика сбора и обработки социологической информации из сети Интернет // *Известия Тульского государственного университета. Технические науки*. 2022. № 2. С. 208–213. <https://doi.org/10.24412/2071-6168-2022-2-208-214>
- Фомина И.К., Татаурова А.С. Повышение эффективности модели классификации определения тональности текста // *Актуальные проблемы экономики и управления*. 2022. № 1. С. 55–58.
- Zucco G., Azzopardi L., van Rijsbergen K. The quantum probability ranking principle for information retrieval // *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2009. V. 5766. P. 232–240. https://doi.org/10.1007/978-3-642-04417-5_21
- Melucci M. *Introduction to Information Retrieval and Quantum Mechanics*. Springer-Verlag Berlin Heidelberg, 2015. 247 p. <https://doi.org/10.1007/978-3-662-48313-8>
- Хренников А.Ю. Введение в квантовую теорию информации. М.: Физматлит, 2008. 283 с.
- Blacoe W., Kashefi E., Lapata M. A quantum-theoretic approach to distributional semantics // *Proc. of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT)*. 2013. P. 847–857.
- Jaiswal A.K., Holdack G., Frommholz I., Liu H. Quantum-like Generalization of Complex Word Embedding: a lightweight approach for textual classification // *CEUR Workshop Proceedings*. 2018. V. 2191. P. 159–168.
- Surov I.A., Semenenko E., Platonov A.V., Bessmertny I.A., Galofaro F., Toffano Z., Khrennikov A.Yu., Alodjants A.P. Quantum semantics of text perception // *Scientific Reports*. 2021. V. 11. N 1. P. 4193. <https://doi.org/10.1038/s41598-021-83490-9>
- Груздева А.С., Бессмертный И.А. Классификация коротких текстов с использованием волновой модели // *Научно-технический вестник информационных технологий, механики и оптики*. 2022. Т. 22. № 2. С. 287–293. <https://doi.org/10.17586/2226-1494-2022-22-2-287-293>
- Френкель Я.И. Волновая механика. Ч. 1. Элементарная теория. Квантовая физика. М., 2019. 386 с.
- Kutuzov A., Kuzmenko E. Webvectors: A toolkit for building web interfaces for vector semantic models // *Communications in Computer and Information Science*. 2017. V. 661. P. 155–161. https://doi.org/10.1007/978-3-319-52920-2_15
- Panchenko A., Ustalov D., Arefyev N., Paperno D., Konstantinova N., Loukachevitch N., Biemann C. Human and machine judgements about russian semantic relatedness // *Communications in Computer and Information Science*. 2016. V. 661. P. 221–235. https://doi.org/10.1007/978-3-319-52920-2_21
- Pang B., Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. *Proc. of the 42nd*
- Garg N., Sharma K. Text pre-processing of multilingual for sentiment analysis based on social network data. *International Journal of Electrical and Computer Engineering*, 2022, vol. 12, no. 1, pp. 776–784. <https://doi.org/10.11591/ijece.v12i1.pp776-784>
- Korney A.O., Kryuchkova E.N. Performance aspects of the russian texts sentiment analysis in social networks. *High-performance computing systems and technologies*, 2018, vol. 2, no. 2, pp. 87–92. (in Russian)
- Pazelskaia A.G., Solovev A.N. Method for determining emotions in Russian texts. *Proc. of the International conference on computational linguistics and intellectual technologies “Dialogue 2011”*. Moscow, RSUH, 2011, pp. 510–522. (in Russian)
- Klekovkina M., Kotelnikov E. The automatic sentiment text classification method based on emotional vocabulary. *Digital Libraries: Advanced Methods and Technologies. Proceedings of the RCDL 2012*, 2012, pp. 118–123. (in Russian)
- Bessmertnyi I.A., Dzhaliashvili Z.O., Maksimov V.V., Markin D.A. Linguistic evaluation of the text. *Abstracts of the X International Conference “Application of New Technologies in Education”*, Troitsk, BYTIC, 1999. (in Russian)
- Harris Z. *Mathematical Structures of Language*. Wiley, 1968, 230 p.
- Bogdanov A.L., Dulya I.S. Sentiment analysis of short Russian texts in social media. *Tomsk State University Journal of Economics*, 2019, no. 47, pp. 220–241. (in Russian). <https://doi.org/10.17223/19988648/47/17>
- Riabykin A.S., Sukhov E.A. Neural network methods in the Sentiment Analysis problem. *DSPA*, 2022, vol. 2, no. 2, pp. 41–57. (in Russian)
- Vorobyov A.A., Rybak A.M., Seredkin R.A., Androsov A.Yu., Solovoyov B.I. Methods of collecting and processing sociological information from the Internet. *Izvestiya Tula State University. Technical Sciences*, 2022, no. 2, pp. 208–213. (in Russian). <https://doi.org/10.24412/2071-6168-2022-2-208-214>
- Fomina I.K., Tataurova A.S. Increasing the efficiency of the classification model for determining the text key. *Aktual'nye problemy ekonomiki i upravleniya*, 2022, no. 1, pp. 55–58. (in Russian)
- Zucco G., Azzopardi L., van Rijsbergen K. The quantum probability ranking principle for information retrieval. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2009, vol. 5766, pp. 232–240. https://doi.org/10.1007/978-3-642-04417-5_21
- Melucci M. *Introduction to Information Retrieval and Quantum Mechanics*. Springer-Verlag Berlin Heidelberg, 2015, 247 p. <https://doi.org/10.1007/978-3-662-48313-8>
- Khrennikov A. *Introduction to Quantum Information Theory*. Moscow, Fizmatlit Publ., 2008, 283 p. (in Russian)
- Blacoe W., Kashefi E., Lapata M. A quantum-theoretic approach to distributional semantics. *Proc. of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL HLT)*, 2013, pp. 847–857.
- Jaiswal A.K., Holdack G., Frommholz I., Liu H. Quantum-like Generalization of Complex Word Embedding: a lightweight approach for textual classification. *CEUR Workshop Proceedings*, 2018, vol. 2191, pp. 159–168.
- Surov I.A., Semenenko E., Platonov A.V., Bessmertny I.A., Galofaro F., Toffano Z., Khrennikov A.Yu., Alodjants A.P. *Quantum semantics of text perception*. *Scientific Reports*, 2021, vol. 11, no. 1, pp. 4193. <https://doi.org/10.1038/s41598-021-83490-9>
- Gruzdeva A.S., Bessmertny I.A. Classification of short texts using a wave model. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 2, pp. 287–293. (in Russian). <https://doi.org/10.17586/2226-1494-2022-22-2-287-293>
- Frenkel I.A. *Wave Mechanics. Part I. Elementary Theory*. *Quantum physics*. Moscow, 2019, 386 p. (in Russian)
- Kutuzov A., Kuzmenko E. Webvectors: A toolkit for building web interfaces for vector semantic models. *Communications in Computer and Information Science*, 2017, vol. 661, pp. 155–161. https://doi.org/10.1007/978-3-319-52920-2_15
- Panchenko A., Ustalov D., Arefyev N., Paperno D., Konstantinova N., Loukachevitch N., Biemann C. Human and machine judgements about russian semantic relatedness. *Communications in Computer and Information Science*, 2016, vol. 661, pp. 221–235. https://doi.org/10.1007/978-3-319-52920-2_21
- Pang B., Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. *Proc. of the 42nd*

- and Information Science. 2016. V. 661. P. 221–235. https://doi.org/10.1007/978-3-319-52920-2_21
21. Pang B., Lee L. A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts // Proc. of the 42nd Annual Meeting Association for Computational Linguistics (ACL). 2004. P. 271–278. <https://doi.org/10.3115/1218955.1218990>
 22. Меньшиков И.Л. Анализ тональности текста на русском языке при помощи графовых моделей // Доклады всероссийской научной конференции АИСТ'2013 «Анализ Изображений, Сетей и Текстов». Екатеринбург, 2013. С. 151–155.
 23. González F.A., Caicedo J.C. Quantum latent semantic analysis // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2011. V. 6931. P. 52–63. https://doi.org/10.1007/978-3-642-23318-0_7
 24. Dalla Chiara M.L., Giuntini R., Luciani A.R., Negri E. A quantum-like semantic analysis of ambiguity in music // Soft Computing. 2017. V. 21. N 6. P. 1473–1481. <https://doi.org/10.1007/s00500-015-1895-y>
 - Annual Meeting Association for Computational Linguistics (ACL), 2004, pp. 271–278. <https://doi.org/10.3115/1218955.1218990>
 22. Menshikov I. Sentiment analysis of the text in Russian using graph models. Reports of the All-Russian Scientific Conference “Analysis of Images, Networks and Texts, AINT (AIST)”, 2013, pp. 151–155. (in Russian)
 23. González F.A., Caicedo J.C. Quantum latent semantic analysis. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2011, vol. 6931, pp. 52–63. https://doi.org/10.1007/978-3-642-23318-0_7
 24. Dalla Chiara M.L., Giuntini R., Luciani A.R., Negri E. A quantum-like semantic analysis of ambiguity in music. Soft Computing, 2017, vol. 21, no. 6, pp. 1473–1481. <https://doi.org/10.1007/s00500-015-1895-y>

Авторы

Груздева Анастасия Сергеевна — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57674037100](https://orcid.org/0000-0003-4963-0823), <https://orcid.org/0000-0003-4963-0823>, prog.anastasia@gmail.com

Юрьев Родион Николаевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57485730300](https://orcid.org/0000-0003-1146-2617), <https://orcid.org/0000-0003-1146-2617>, rodion@juryev.ru

Бессмертный Игорь Александрович — доктор технических наук, профессор, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 36661767800](https://orcid.org/0000-0001-6711-6399), <https://orcid.org/0000-0001-6711-6399>, bessmertny@itmo.ru

Authors

Anastasia S. Gruzdeva — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57674037100](https://orcid.org/0000-0003-4963-0823), <https://orcid.org/0000-0003-4963-0823>, prog.anastasia@gmail.com

Rodion N. Iurev — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57485730300](https://orcid.org/0000-0003-1146-2617), <https://orcid.org/0000-0003-1146-2617>, rodion@juryev.ru

Igor A. Bessmertny — D. Sc., Full Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 36661767800](https://orcid.org/0000-0001-6711-6399), <https://orcid.org/0000-0001-6711-6399>, bessmertny@itmo.ru

Статья поступила в редакцию 27.05.2022

Одобрена после рецензирования 27.09.2022

Принята к печати 12.11.2022

Received 27.05.2022

Approved after reviewing 27.09.2022

Accepted 12.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1166-1177

УДК 004.9 + 616.1/9

Автоматизированная оценка параметров электрокардиограмм в условиях пандемии COVID-19

Александра Сергеевна Ватьян¹, Наталия Федоровна Гусарова²✉,
Наталия Викторовна Добренко³, Данил Александрович Змиевский⁴,
Максим Васильевич Кабышев⁵, Татьяна Андреевна Полевая⁶,
Анна Андреевна Татаринова⁷, Иван Вячеславович Томилов⁸

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

⁷ Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-Петербург, 197341, Российская Федерация

¹ alexvatyan@gmail.com, <https://orcid.org/0000-0002-5483-716X>

² natfed@list.ru✉, <https://orcid.org/0000-0002-1361-6037>

³ graziokisa@yandex.ru, <https://orcid.org/0000-0001-6206-8033>

⁴ danil.zmievskiy@gmail.com, <https://orcid.org/0000-0002-0077-794X>

⁵ maxk6971@gmail.com, <https://orcid.org/0000-0002-1006-0408>

⁶ tanusha2406@gmail.com, <https://orcid.org/0000-0001-6131-0019>

⁷ antsnet.18@mail.ru, <https://orcid.org/0000-0002-9046-2457>

⁸ ivan-tomilov3@yandex.ru, <https://orcid.org/0000-0003-1886-2867>

Аннотация

Предмет исследования. Рассмотрены алгоритмы оперативной автоматизированной оценки параметров электрокардиограммы в условиях отсутствия специализированного оборудования и профильных специалистов.

Метод. Электрокардиограмма пациента записана на бумажную ленту, которая сфотографирована на мобильный телефон врача первичного звена и обработана специализированным приложением. С помощью приложения выполнена оцифровка фотографируемого изображения электрокардиограммы, оценка ее основных параметров, а также расчет критериев для дифференциальной диагностики отдельных заболеваний по приближенным формулам. Кроме того, оцифрованное изображение электрокардиограммы передается на сервер и обрабатывается с помощью системы машинного обучения. **Основные результаты.** Разработаны алгоритмы оцифровки и анализа электрокардиограммы, которые могут быть использованы для оценки ее элементов, важных для диагностики. Средняя ошибка определения положения наиболее сложных (сглаженных) пиков — зубцов Р и Т — составила не более 0,1 мм. Предложен алгоритм критериального анализа электрокардиограммы для поддержки дифференциальной диагностики острого инфаркта миокарда с подъемом сегмента ST и синдрома ранней реполяризации желудочков, который обеспечил значения точности 0,85 и F-меры 0,74. Представлен альтернативный алгоритм на основе глубокой нейронной сети, который обеспечивает лучшие значения точности и F-меры — 0,96 и 0,88, но требует больших затрат вычислительных ресурсов и выполнения расчетов на сервере.

Практическая значимость. Алгоритмы реализованы в виде набора библиотечных функций. Они могут быть использованы как самостоятельно, так и в составе полномасштабной системы поддержки принятия клинических решений для автоматизированной оценки параметров электрокардиограммы на основе клиент-серверной архитектуры. Все результаты расчетов совместно с фотографией исходной электрокардиограммы могут быть оперативно переданы квалифицированному кардиологу с целью дистанционного получения консультативного заключения.

Ключевые слова

COVID-19, система поддержки клинических решений, автоматизированная оценка параметров электрокардиограммы, поддержка дифференциальной диагностики

Благодарности

Работа поддержана грантом Президента Российской Федерации № МК-5723.2021.1.6.

Ссылка для цитирования: Ватян А.С., Гусарова Н.Ф., Добренко Н.В., Змиевский Д.А., Кабышев М.В., Полевая Т.А., Татаринова А.А., Томилов И.В. Автоматизированная оценка параметров электрокардиограмм в условиях пандемии COVID-19 // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1166–1177. doi: 10.17586/2226-1494-2022-22-6-1166-1177

Automated evaluation of ECG parameters during the COVID-19 pandemic

Alexandra S. Vatian¹, Natalia F. Gusarova²✉, Natalia V. Dobrenko³,
Danil A. Zmievsky⁴, Maxim V. Kabyshev⁵, Tatiana A. Polevaya⁶,
Anna A. Tatarinova⁷, Ivan V. Tomilov⁸

^{1,2,3,4,5,6,8} ITMO University, Saint Petersburg, 197101, Russian Federation

⁷ Almazov National Medical Research Centre, Saint Petersburg, 197341, Russian Federation

¹ alexvatyan@gmail.com, <https://orcid.org/0000-0002-5483-716X>

² natfed@list.ru✉, <https://orcid.org/0000-0002-1361-6037>

³ graziokisa@yandex.ru, <https://orcid.org/0000-0001-6206-8033>

⁴ danil.zmievskiy@gmail.com, <https://orcid.org/0000-0002-0077-794X>

⁵ maxk6971@gmail.com, <https://orcid.org/0000-0002-1006-0408>

⁶ tanusha2406@gmail.com, <https://orcid.org/0000-0001-6131-0019>

⁷ antsvet.18@mail.ru, <https://orcid.org/0000-0002-9046-2457>

⁸ ivan-tomilov3@yandex.ru, <https://orcid.org/0000-0003-1886-2867>

Abstract

Algorithms for prompt automated evaluation of electrocardiogram parameters in the absence of specialized equipment and specialized specialists are considered. The patient's electrocardiogram is recorded on a paper tape, then it is photographed on the primary care doctor's mobile phone and processed by a specialized application. The application digitizes the photographed image of the electrocardiogram, evaluates its main parameters as well as calculates criteria for the differential diagnosis of certain diseases using approximate formulas. In addition, the digitized electrocardiogram image is transmitted to the server and processed using a machine learning system. Algorithms for digitizing and analyzing an electrocardiogram have been developed that make it possible to evaluate its elements that are important for diagnosis, and the average error in determining the position of the most complex (smoothed) peaks — P and T waves — was no more than 0.1 mm. An algorithm for the criteria analysis of an electrocardiogram is proposed to support the differential diagnosis of acute myocardial infarction with ST segment elevation and early ventricular repolarization syndrome which provides accuracy values of 0.85 and F-scores of 0.74. An alternative algorithm based on a deep neural network is proposed which provides the best values — 0.96 and 0.88, respectively, but requires large computing resources and is executed on the server. The algorithms are implemented as a set of library functions. They can be used both independently and as part of a full-scale clinical decision support system for automated evaluation of electrocardiogram parameters based on a client-server architecture. In addition, all calculation results, together with a photograph of the original electrocardiogram, can be promptly transferred to a qualified cardiologist in order to receive an advisory opinion remotely.

Keywords

COVID-19, clinical decision support system, automated evaluation of electrocardiogram parameters, support for differential diagnosis

Acknowledgements

The work was supported by the Grant of the President of the Russian Federation No. MK-5723.2021.1.6.

For citation: Vatian A.S., Gusarova N.F., Dobrenko N.V., Zmievsky D.A., Kabyshev M.V., Polevaya T.A., Tatarinova A.A., Tomilov I.V. Automated evaluation of ECG parameters during the COVID-19 pandemic. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1166–1177 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1166-1177

Введение

Мировой опыт борьбы с пандемией COVID-19 показал, что наиболее тяжелое течение болезни и наиболее неблагоприятные прогнозы приходятся на больных, имеющих сопутствующие заболевания. Среди сопутствующих заболеваний ведущее место занимают сердечно-сосудистые заболевания (ССЗ). Согласно оценкам [1], наличие ССЗ в 3,4 раза увеличивает тяжесть течения COVID-19 и более чем в 23 раза — смертность.

В отличие от медленно развивающихся заболеваний, таких как онкология, ССЗ часто проявляют себя как катастрофы, требующие принятия немедленного

и обоснованного решения на уровне звена скорой и неотложной медицинской помощи. В условиях пандемии COVID-19 в этом звене наблюдалась нехватка специалистов, которую системы здравоохранения во всем мире компенсируют, в первую очередь, за счет привлечения врачей непрофильных специальностей или интернов [2, 3], которым сложно принять квалифицированное решение относительно ССЗ. Кроме того, перепрофилирование медицинских учреждений под нужды пандемии привело к сокращению числа специализированных ССЗ-коек и осуществлению мер против такой ранее широко практиковавшейся тактики, как гипердиагностика острых ССЗ с последующим уточнением диагноза в стационаре [4].

Основное средство диагностики ССЗ на уровне скорой и неотложной медицинской помощи — электрокардиограмма (ЭКГ) (рис. 1). Однако точность интерпретации ЭКГ существенно отличается в зависимости от квалификации интерпретатора. Например, согласно исследованию в работе [5], точность расшифровки ЭКГ составляет: у кардиологов 63,2–86,7 %, у практикующих врачей 57,6–79,5 %, а у медиков низшего звена и непрофильной квалификации (доля которых при пандемии возрастает) 48,1–63,6 %.

Постановка задачи

Таким образом, в условиях пандемии COVID-19 особую актуальность приобрели системы поддержки принятия клинических решений (СПКР) по ССЗ, причем важно предотвратить как ошибки первого рода (избыточно тяжелый диагноз и госпитализация без острой необходимости), так и ошибки второго рода (недооценка серьезности состояния больного). Перспективным подходом здесь признаны системы автоматизированной диагностики ССЗ [6, 7], в первую очередь, на основе автоматизированного определения параметров ЭКГ [8]. Отметим, что несмотря на впечатляющие успехи, в данной области остается еще много нерешенных проблем [9], среди которых в аспекте пандемии COVID-19 важно выделить следующие.

Недостаточная точность оценки параметров ЭКГ. Диагностика ССЗ по ЭКГ основана на результатах измерений временных параметров всех зубцов и интервалов, которые существенно различаются по степени выраженности и вариабельности (рис. 1, а). Например, автоматизированная оценка отклонений синусового ритма (т. е. RR-интервала) может быть выполнена с точностью до 99 % [10]. При этом интервалы PQ, PR и QT измерить гораздо сложнее из-за сглаженной формы зубцов Р и Т; в результате при нахождении связанных с ними признаков на ЭКГ точность падает до 79,2 % и ниже [8]. В связи с этим, как показывает анализ научных работ последних лет [11, 12], фокус исследований все больше сдвигается с повышения точности оценки ЭКГ в целом на поддержку диагностики отдельных заболеваний, связанных с анализом конкретных фрагментов ЭКГ.

Необходимость специализированного оборудования. Практически все представленные в современных научных работах и на коммерческом уровне системы автоматизированного анализа ЭКГ предполагают использование цифрового электрокардиографа, оснащенного соответствующей компьютерной системой [13]. В звене первичной медико-санитарной помощи в отдельных регионах России, а также в развивающихся странах наблюдается дефицит такого оборудования, проявившийся в большей степени в условиях пандемии COVID-19. Многие бригады скорой помощи имеют в своем распоряжении только электрокардиограф с записью ЭКГ на бумажной ленте (рис. 1, б) и мобильный телефон в качестве компьютерной системы. Систем автоматизированного анализа ЭКГ для такого комплекта оборудования в доступной авторам научной литературе не найдено.

Необходимость оперативного получения квалифицированной консультации. Любые СПКР принципиально являются консультативными, т. е. не снимают необходимости оценки ЭКГ квалифицированным специалистом-кардиологом. Известно, что организовать такую оценку посредством личного контакта узкого специалиста и пациента (на дому или в стационаре) в условиях пандемии COVID-19 проблематично. Для этого необходимо использовать дистанционные формы взаимодействия, в том числе посредством мобильного телефона как наиболее доступного в этой ситуации средства передачи изображения.

Перспективный путь преодоления указанных проблем — построение СПКР для автоматизированной оценки параметров ЭКГ на основе клиент-серверной архитектуры. При этом клиентская часть реализована в виде мобильного приложения, установленного на телефон врача первичного звена. Врач снимает ЭКГ пациента на бумажную ленту кардиографа и фотографирует ее на мобильный телефон. Приложение выполняет оцифровку изображения и автоматизированную оценку основных параметров ЭКГ, которые отображаются на экране телефона, а также производит расчет критериев для дифференциальной диагностики отдельных заболеваний по приближенным формулам. Для уточнения диагноза оцифрованное изображение ЭКГ может передаваться на сервер, где обрабатывается

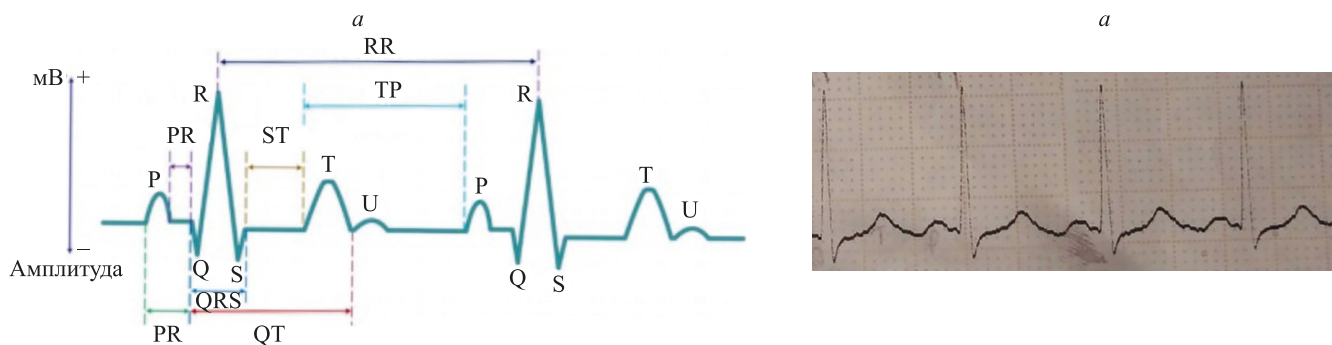


Рис. 1. Характерные интервалы (а) и фрагмент электрокардиограммы в записи на бумаге (б).

P, R, Q, S, T, U — зубцы; PR, ST — сегменты; QRS, RR, PR, QT, TP — интервалы

Fig. 1. Typical ECG intervals (a); fragment of an ECG recorded on paper (b)

с помощью системы машинного обучения. Все результаты расчетов совместно с фотографией исходной ЭКГ могут быть переданы квалифицированному кардиологу с целью дистанционного получения консультативного заключения.

Алгоритм обработки ЭКГ с использованием мобильного телефона был представлен авторами на международной конференции [14]. Цель настоящей работы — развитие идеи в направлении создания полномасштабной СПКР. Рассмотрены ключевые аспекты создания СПКР: оцифровка исходной ЭКГ на бумажной ленте, автоматизированная оценка основных параметров ЭКГ и повышение робастности оценки. Определены возможности поддержки дифференциальной диагностики при остром коронарном синдроме на основе критериальной оценки и машинного обучения. Рассмотрен пример клинически значимой задачи выявления признаков для поддержки дифференциальной диагностики острого инфаркта миокарда с подъемом сегмента ST и синдрома ранней реполяризации желудочков.

Автоматизированная оценка основных параметров ЭКГ

В аналоговых электрокардиографах ЭКГ регистрируется на специальной тепловой бумаге с нанесенной миллиметровой сеткой красного оттенка, которая используется для визуальной оценки параметров ЭКГ (амплитудного и временного масштаба сигналов). Таким образом, для оцифровки фотографии бумажной ЭКГ необходимо разделить изображения сигнала и регистрационной сетки. С этой целью выполнена сравнительная оценка трех подходов:

- 1) спектральной фильтрации посредством маскирования в цветовой модели HSV;
- 2) пороговой фильтрации канала grayscale в полутоновом изображении [13];
- 3) адаптивной бинаризации (автоматическое определение порога для отдельных зон изображения) [15].

Отметим, что подход 3 показал наилучшие результаты. Как показали результаты экспериментов, во избежание возможных шумовых эффектов к отделенному изображению сигнала целесообразно применять фильтр Гаусса и медианный фильтр.

Разработанный алгоритм оцифровки фотографии ЭКГ состоит из следующих шагов: адаптивная бинаризация изображения ЭКГ; фильтрация полученного сигнала ЭКГ; расчет масштаба сетки по ее отделенному изображению. Результатом оцифровки является одномерный массив амплитудных значений сигнала ЭКГ, автоматически приведенный к одному из стандартизованных временных масштабов. Пример визуализации оцифрованного сигнала ЭКГ представлен на рис. 2.

Как показывает практика, медианные значения (изолинии) ЭКГ могут претерпевать низкочастотные изменения как за счет нарушения режима питания электрокардиографа, так и за счет физиологических проявлений пациента. Для устранения этого эффекта в процессе предобработки сигнала использован фильтр Баттерворта (рис. 3).

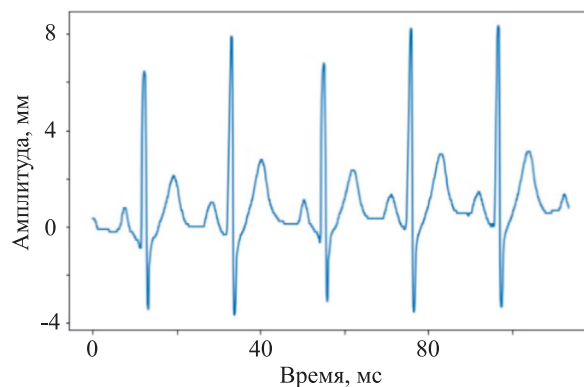


Рис. 2. Оцифрованный сигнал электрокардиограммы

Fig. 2. Digitized ECG signal

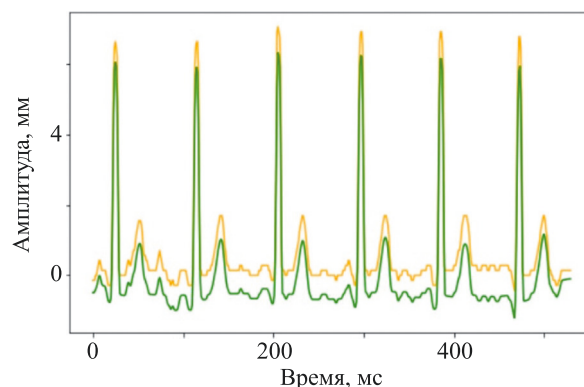


Рис. 3. Устранение смещения изолинии: зеленый — сигнал со смещением изолинии, желтый — сигнал с устраненным смещением изолинии

Fig. 3. Elimination of iseline shift: green — signal with iseline shift, yellow — signal with corrected iseline shift

В работах [16, 17] показано, что для оценки временного положения отдельных зубцов ЭКГ предпочтительны методы на основе вейвлет-преобразований, реализованные, в частности, в широко распространенном пакете NeuroKit¹. Выполненные эксперименты в настоящей работе подтвердили эффективность такого подхода для поиска пиков зубцов ЭКГ (рис. 4, а). Между тем при определении границ зубцов, особенно таких «мягких», как Р и Т, подход показал недостаточную устойчивость (наблюдается смещение и переналожение граничных точек — рис. 4, б). Статистическая обработка полученных значений по методу [18] также оказалась недостаточно эффективной. В связи с этим границы зубцов вычислены на основе предложенного в клинических рекомендациях [19, 20] метода [21]. Метод состоит в поиске максимального или минимального градиента на одной из условных половин оцениваемого зубца ЭКГ и построении касательной, проходящей через точку градиента (рис. 5). Проекция точки пересечения касательной и изолинии непосредственно на сам сигнал принимается за границу зубца.

¹ NeuroKit's documentation [Электронный ресурс]. Режим доступа: <https://rpanderson-neurokit2.readthedocs.io/en/latest/index.html> (дата обращения: 20.11.2022).

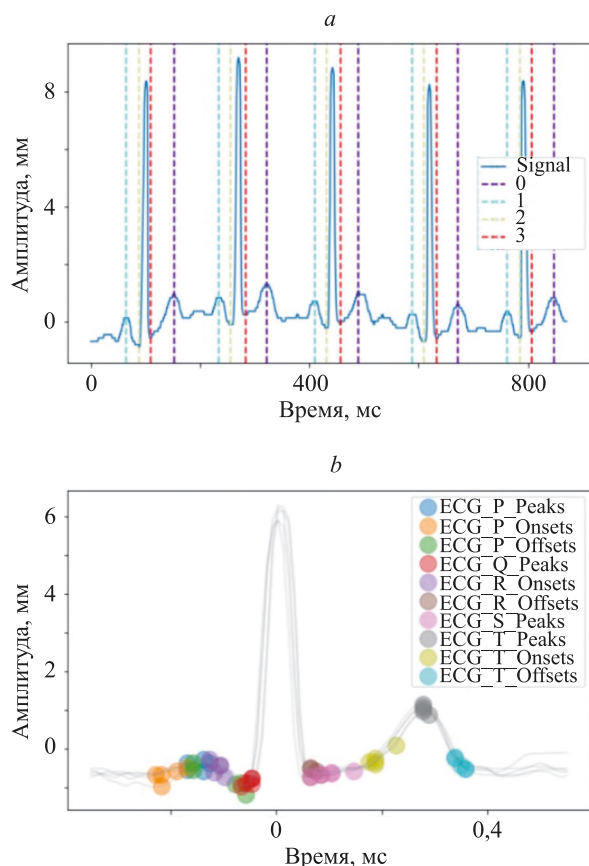


Рис. 4. Определения: пиков (пакет NeuroKit, функция delineate, метод peak) (a) и границ (пакет NeuroKit, функция delineate, метод dwt) (b) зубцов электрокардиограммы

Fig. 4. a — determination of ECG wave peaks (NeuroKit package, delineate function, peak method); b — determination of ECG wave boundaries (NeuroKit package, delineate function, dwt method)

Для оценки разработанных алгоритмов использован публично доступный датасет [22], состоящий из записей ЭКГ, на которых врачами-кардиологами вручную размечены точки пиков R, P, T зубцов. В результате экспериментов средняя ошибка определения положения пиков ЭКГ составила не более 0,1 мм для зубцов P и T

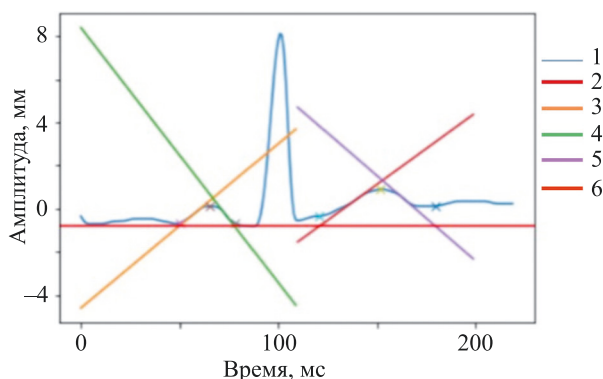


Рис. 5. Схема определения границ зубцов электрокардиограммы по методу касательных

Fig. 5. Scheme for determining the boundaries of ECG teeth using the tangent method

и 0,6 мм для R. Выборочное визуальное обследование использованного датасета выявило наличие выбросов (неверно размеченных положений пиков) именно для зубца R.

Разработанные алгоритмы оцифровки и анализа ЭКГ реализованы на языке Python в виде набора библиотечных функций. При этом на входе функции находится фотографическое изображение ЭКГ в формате .jpg, а на выходе — набор значений элементов ЭКГ: амплитуды зубцов R, P, T, длительность интервалов RR и QT, а также сегментов ST и PR.

Выявление признаков для поддержки дифференциальной диагностики инфаркта миокарда с подъемом сегмента ST и синдрома ранней реполяризации желудочков на основе автоматизированной оценки параметров ЭКГ

В качестве одной из клинически значимых задач рассмотрим дифференциальную диагностику таких состояний, как острый инфаркт миокарда (ОИМ) с подъемом (элевацией) сегмента ST (ST-ОИМ) и синдром ранней реполяризации желудочков (СРРЖ/ЕР), который также проявляется на ЭКГ в виде подъема сегмента ST (рис. 6). Если ST-ОИМ — жизнеугрожающее состояние, которое требует немедленной госпитализации, то СРРЖ/ЕР рассматривается как фенотипический вариант нормы, хотя и ассоциируется с повышением риска внезапной сердечной смерти. В условиях пандемии COVID-19 дифференциальная диагностика состояний ST-ОИМ и СРРЖ/ЕР непосредственно на месте, на основе анализа только ЭКГ, т. е. без госпитализации и дополнительных обследований, приобретает особую важность и одновременно вызывает особые затруднения у медиков звена скорой и неотложной медицинской помощи.

Отметим, что в научной литературе представлен ряд работ, в которых задача дифференциальной диа-

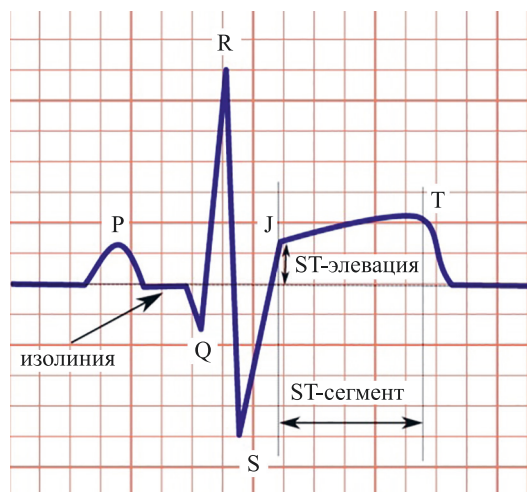


Рис. 6. Элевация сегмента ST на электрокардиограмме во II отведении

Fig. 6. Elevation of the ST segment on the electrocardiogram in lead II

гностики состояний ST-ОИМ и СРРЖ/ЕР решается как двухклассовая классификация на наборе признаков, извлекаемых из ЭКГ. Как показал сравнительный анализ литературы, наилучшая эффективность установлена в работе [22], где на основе многомерной линейной регрессии построен критерий разделения следующего вида:

$$(1,196 \times STE60 \text{ V3}) + (0,059 \times QTc) - (0,326 \times RA \text{ V4}) > Q, \quad (1)$$

где $STE60 \text{ V3}$ — значение ST-элевации в отведении V3, мм; QTc — значение скорректированного интервала QT, мс; $RA \text{ V4}$ — значение амплитуды пика R в отведении V4, мм (приведение к безразмерным единицам учтено в нормировочных коэффициентах). При пороговом значении $Q > 23,4$ принимается решение о наличии ST-ОИМ, при $Q < 23,4$ — о СРРЖ/ЕР.

В [23] для вывода регрессионной формулы (1) использован датасет, содержащий только случаи ОИМ передней локализации. В настоящей работе, исходя из результатов работы [23], проведено уточнение регрессионных коэффициентов на основе открытого датасета [24], содержащего ЭКГ с инфарктами различных типов. Для этого использован фрагмент набора данных [24], содержащий 159 записей ЭКГ, принадлежащих к состояниям ST-ОИМ (49 записей) и СРРЖ/ЕР (110 записей). Применение формулы (1) на этом датасете дало значительно более низкие показатели эффективности разделения, чем в [23] (табл. 1).

Для решения проблемы эффективности выполнена модификация формулы (1): оптимизация порогового значения Q без изменения остальных коэффициентов и ограничение максимальной амплитуды пика R. В результате формула (1) преобразована к виду

$$(1,0 \times STE60 \text{ V3}) + (0,06 \times QTc) - (0,5 \times \min(RA \text{ V4}, 10,5)) > 27,1. \quad (2)$$

Для снижения вероятности переобучения произведено дополнительное дообучение модели на датасете [24] при помощи логистической регрессии с кросс-валидацией на трех разбиениях обучающей выборки. Тогда полученная модификация исходной формулы (1) имеет вид

$$(2,9 \times STE60 \text{ V3}) + (0,3 \times QTc) - (1,7 \times \min(RA \text{ V4}, 19)) \geq 126,9. \quad (3)$$

Из анализа табл. 1 видно, что модификация критерия (1) к форме (2) позволила получить практически исходно заявленные значения точности при существенном увеличении робастности. При необходимости дальнейшего увеличения робастности критериального метода (формула (3)) приходится идти на некоторое снижение точности и F1-меры. Важно подчеркнуть, что разработанные алгоритмы оценки критерия (2) и (3) нетребовательны к вычислительным ресурсам и могут быть легко размещены в виде автономного приложения на большинстве моделей мобильных телефонов, распространенных в России.

Отметим, что вопрос влияния конкретной локализации очага нарушения кровоснабжения, как и других предикторов, представляет самостоятельный интерес и выходит за рамки данной работы, цель которой — показать возможности инструментальной поддержки дифференциальной диагностики острых состояний сердечно-сосудистой системы на примере конкретной пары состояний.

Альтернативный подход к поддержке дифференциальной диагностики состояний ST-ОИМ и СРРЖ/ЕР — подход на основе глубоких нейронных сетей [25, 26]. Для эксперимента использована сеть с архитектурой ConvNet [27], которая включает в себя четыре группы сверточных слоев и слой линейной трансформации.

Результаты настройки сети СРРЖ/ЕР и ST-ОИМ, соответственно, иллюстрируются на графиках f1-меры (рис. 7, а, b и рис. 8, а, b), потерь (рис. 7, c, d и рис. 8, c, d), точности (рис. 7, e, f и рис. 8, e, f), полноты (рис. 7, g, h и рис. 8, g, h), специфичности (рис. 7, i, j и рис. 8, i, j). Графики демонстрируют достижение необходимых показателей нейронной сети при сравнительно небольшом количестве итераций обучения, что можно считать достоинством используемой архитектуры. Кроме того, следует подчеркнуть, что эти результаты достигнуты на достаточно малом количестве обучающих примеров, что является очень ценным для поддержки дифференциальной диагностики сравнительно редко встречающихся заболеваний, для которых количество клинических случаев в популяции принципиально невелико (даже в таком объемном датасете, как [22], таких случаев всего 220).

Количественные оценки эффективности разработанного алгоритма поддержки дифференциальной диагностики состояний ST-ОИМ и СРРЖ/ЕР на основе глубокого обучения представлены в табл. 2.

Несмотря на то, что конкретная реализация алгоритма (в виде обученной нейронной сети) может быть

Таблица 1. Сравнительная оценка эффективности различных модификаций критерия

Table 1. Comparative evaluation of the effectiveness of various criterion modifications

Датасет	Вид критерия	Метрики			
		Точность (accuracy)	Точность (precision)	Полнота	F1-мера
[23]	(1)	0,88	0,84	0,86	0,85
[24]	(1)	0,64	0,44	0,78	0,56
	(2)	0,85	0,76	0,71	0,74
	(3)	0,79	0,63	0,69	0,66

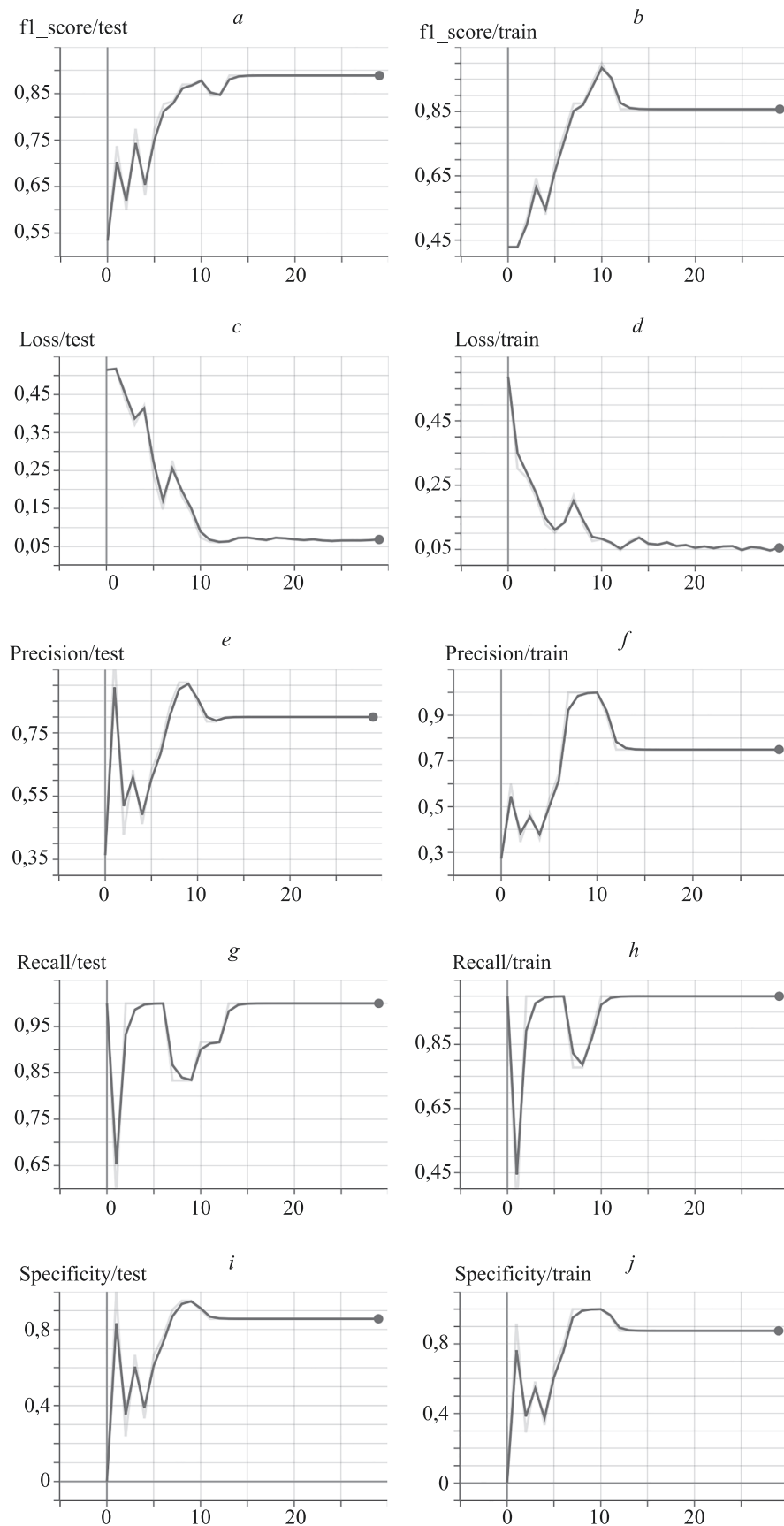


Рис. 7. Графики результатов настройки модели при определении состояния СРРЖ: F1-меры (f1-score) (a, b), потери (loss) (c, d), точности (precision) (e, f), полноты (recall) (g, h), специфичности (specificity) (i, j). По горизонтали — число итераций
 Fig. 7. Results of model tuning when determining the EVRS: graphs of the F1-score (a, b), loss (c, d), precision (e, f), recall (g, h), specificity (i, j). Horizontal — number of iterations

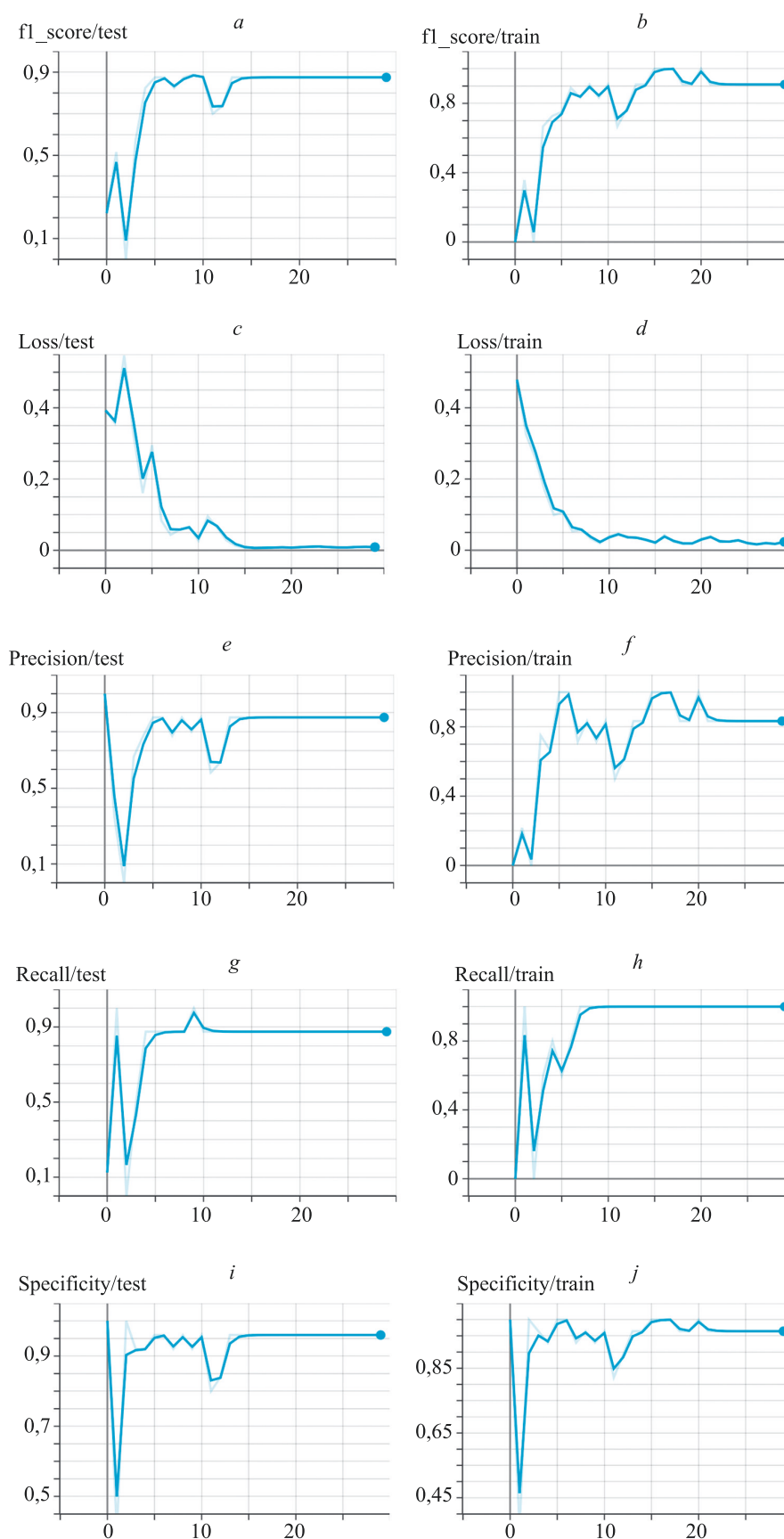


Рис. 8. Графики результатов настройки модели при определении состояния острого инфаркта миокарда: графики F1-меры (f1-score) (*a, b*), потери (loss) (*c, d*), точности (precision) (*e, f*), полноты (recall) (*g, h*), специфичности (specificity) (*i, j*). По горизонтали — число итераций

Fig. 8. Results of model tuning when determining the AMI: graphs of the f1-score (*a, b*), loss (*c, d*), precision (*e, f*), recall (*g, h*), specificity (*i, j*). Horizontal — number of iterations

Таблица 2. Оценка эффективности алгоритмов на основе глубокого обучения

Table 2. Evaluation of the effectiveness of algorithms based on deep learning

Датасет	Состояние	Точность (accuracy)	Точность (precision)	Полнота	F1-мера
[24]	ST-ОИМ	0,96	0,85	0,86	0,88
	СРРЖ/ER	0,91	0,85	0,92	0,88

реализована в виде автономного приложения для мобильного телефона, с точки зрения возможностей модификации и дообучения на вновь поступающие клинические данные целесообразно размещение приложения на сервере и обращение к нему через сеть Интернет. Отметим, что алгоритм реализован на языке Python в виде набора библиотечных функций, входом которого является сигнал ЭКГ в цифровом виде, а выходом — булево значение признака заболевания с маркированной вероятностью.

Сравнительная оценка эффективности различных модификаций критериального подхода (табл. 1) и подхода на основе глубокого обучения (табл. 2) позволила сделать следующие выводы:

- эффективность автоматизированной дифференциальной диагностики состояний ST-ОИМ и СРРЖ/ER может сильно зависеть от данных, на которых проводится обучение метода;
- несмотря на то, что критерий вида (2) показал наилучшие показатели эффективности, с точки зрения робастности к данным целесообразно использовать модель (3);
- подход на основе глубокого обучения показывает лучшую эффективность, но требует дополнительной оценки робастности на альтернативных датасетах, что является предметом дальнейших исследований.

Заключение

В условиях пандемии COVID-19 скорая кардиологическая помощь во всем мире испытывает трудности, связанные с нехваткой профильных специалистов и специализированного оборудования. В связи с этим особую значимость приобретает построение систем поддержки принятия клинических решений для автоматизированной оценки параметров электрокардиограммы на основе клиент-серверной архитектуры. Рассмотрены ключевые аспекты создания системы поддержки принятия клинических решений: оцифровка исходной бумажной электрокардиограммы и автомати-

зированная оценка основных ее параметров с помощью приложения, устанавливаемого на мобильный телефон врача первичного звена, а также поддержка дифференциальной диагностики заболеваний на основах критериальной оценки и машинного обучения. В качестве примера приведена клинически значимая задача дифференциальной диагностики инфаркта миокарда с подъемом (элевацией) сегмента ST (ST-ОИМ) и синдрома ранней реполяризации желудочков (СРРЖ/ER).

Разработанные алгоритмы оцифровки и анализа электрокардиограммы позволили в автоматическом режиме и без подключения к сети Интернет оценить наиболее значимые элементы электрокардиограммы (амплитуды зубцов R, P, T, длительность интервалов RR и QT, а также сегментов ST и PR). При этом средняя ошибка определения положения наиболее сложных (сглаженных) пиков электрокардиограммы (зубцов R и T) — не более 0,1 мм.

Разработан алгоритм для поддержки дифференциальной диагностики состояний ST-ОИМ и СРРЖ/ER на основе критериального анализа электрокардиограммы. Предложенный алгоритм работает без подключения к сети Интернет и обеспечивает значения точности 0,85 и F-меры 0,74. Разработан альтернативный алгоритм на основе глубокой нейронной сети, который обеспечивает лучшие значения — 0,96 и 0,88 соответственно, но требует подключения к сети Интернет.

Все алгоритмы реализованы на языке Python в виде набора библиотечных функций и входят в состав создаваемой в Университете ИТМО библиотеки ECG Recognition Library, размещенной в открытом доступе¹. Они могут быть использованы как самостоятельно, так и в составе полномасштабной системы поддержки принятия клинических решений для автоматизированной оценки параметров электрокардиограммы на основе клиент-серверной архитектуры, которая является предметом дальнейшей работы авторов.

¹ [Электронный ресурс]. Режим доступа: <https://github.com/Med-AI-Lab/ECG-Recognition-Library> (дата обращения: 20.11.2022).

Литература

1. Yu J.-N., Wu B.-B., Yang J., Lei X.-L., Shen W.-Q. Cardio-cerebrovascular disease is associated with severity and mortality of COVID-19: A systematic review and meta-analysis // *Biological Research for Nursing*. 2021. V. 23. N 2. P. 258–269. <https://doi.org/10.1177/1099800420951984>
2. Abir M., Nelson Ch., Chan E.W., Al-Ibrahim H., Cutter Ch., Patel K., Bogar A. *Critical care surge response strategies for the 2020 COVID-19 outbreak in the United States*. Santa Monica, CA: RAND Corporation, 2020 [Электронный ресурс]. URL: https://www.rand.org/pubs/research_reports/RRA164-1.html (accessed: 09.09.2021).

References

1. Yu J.-N., Wu B.-B., Yang J., Lei X.-L., Shen W.-Q. Cardio-cerebrovascular disease is associated with severity and mortality of COVID-19: A systematic review and meta-analysis. *Biological Research for Nursing*, 2021, vol. 23, no. 2, pp. 258–269. <https://doi.org/10.1177/1099800420951984>
2. Abir M., Nelson Ch., Chan E.W., Al-Ibrahim H., Cutter Ch., Patel K., Bogar A. *Critical care surge response strategies for the 2020 COVID-19 outbreak in the United States*. Santa Monica, CA: RAND Corporation, 2020. Available at: https://www.rand.org/pubs/research_reports/RRA164-1.html (accessed: 09.09.2021).

- org/pubs/research_reports/RRA164-1.html (дата обращения: 09.09.2021).
3. Health systems resilience during COVID-19: Lessons for building back better / ed. by A. Sagan, E. Webb, I. de la Mata, J. Figueras, M. McKee, N. Azzopardi-Muscat. WHO Regional Office for Europe, 2021.
4. Wang N.C., Jain S.K., Estes N.A.M., Barrington W.W., Bazaz R., Bhonsale A., Kancharla K., Shalaby A.A., Voigt A.H., Saba S. Priority plan for invasive cardiac electrophysiology procedures during the coronavirus disease 2019 (COVID-19) pandemic // *Journal of Cardiovascular Electrophysiology*. 2020. V. 31. N 6. P. 1255–1258. <https://doi.org/10.1111/jce.14478>
5. Cook D.A., Oh S., Pusic M.V. Accuracy of physicians' electrocardiogram interpretations: A systematic review and meta-analysis // *JAMA Internal Medicine*. 2020. V. 180. N 11. P. 1461–1471. <https://doi.org/10.1001/jamainternmed.2020.3989>
6. Javed A., Khan S.U., Ali L., Ali S., Imrana Y., Rahman A. Machine learning-based automated diagnostic systems developed for heart failure prediction using different types of data modalities: A systematic review and future directions // *Computational and Mathematical Methods in Medicine*. 2022. V. 2022. P. 9288452. <https://doi.org/10.1155/2022/9288452>
7. Martin-Isla C., Campello V.M., Izquierdo C., Raisi-Estabragh Z., Baeßler B., Petersen S.E., Lekadir K. Image-based cardiac diagnosis with machine learning: A review // *Frontiers in Cardiovascular Medicine*. 2020. V. 7. P. 1. <https://doi.org/10.3389/fcvm.2020.00001>
8. Attia Z.I., Noseworthy P.A., Lopez-Jimenez F., Asirvatham S.J., Deshmukh A.J., Gersh B.J., Carter R.E., Yao X., Rabinstein A.A., Erickson B.J., Kapa S., Friedman P.A. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: A retrospective analysis of outcome prediction // *Lancet*. 2019. V. 394. P. 861–867. [https://doi.org/10.1016/S0140-6736\(19\)31721-0](https://doi.org/10.1016/S0140-6736(19)31721-0)
9. Katoh T., Yashima M., Takahashi N., Watanabe E., Ikeda T., Kasamaki Y., Sumitomo N., Ueda N., Morita H., Hiraoka M. Expert consensus document on automated diagnosis of the electrocardiogram: The task force on automated diagnosis of the electrocardiogram in Japan. Part 2: Current status of inappropriate automated diagnosis is widely used electrocardiographs in Japan // *Journal of Arrhythmia*. 2021. V. 37. N 6. P. 1427–1433. <https://doi.org/10.1002/joa3.12646>
10. Sangaiah A., Arumugam M., Bian G. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis // *Artificial Intelligence in Medicine*. 2020. V. 103. P. 101788. <https://doi.org/10.1016/j.artmed.2019.101788>
11. Tse G., Lee S., Li A., Chang D., Li G., Zhou J., Liu T., Zhang Q. Automated electrocardiogram analysis identifies novel predictors of ventricular arrhythmias in brugada syndrome // *Frontiers in Cardiovascular Medicine*. 2021. V. 7. P. 618254. <https://doi.org/10.3389/fcvm.2020.618254>
12. Rueda C., Fernández I., Larriba Y., Rodríguez-Collado A., Canedo C. Compelling new electrocardiographic markers for automatic diagnosis // *Computer Methods and Programs in Biomedicine*. 2022. V. 221. P. 106807. <https://doi.org/10.1016/j.cmpb.2022.106807>
13. Fortune J.D., Coppa N.E., Haq K.T., Patel H., Tereshchenko L.G. Digitizing ECG image: new fully automated method. 2021 [Электронный ресурс]. URL: <https://www.medrxiv.org/content/10.1101/2021.07.13.21260461v1.full.pdf> (дата обращения: 25.04.2022)
14. Vatian A., Peredreev D., Rodiontsev K., Murzina A., Klevtsova E., Tatarinova A., Treshkur T., Shalyto A., Gusarova N. Helping paramedics in assessing a patient's condition based on ECG by means of mobile phone // *Proc. of the International Conferences ICT, Society, and Human Beings 2021; Web Based Communities and Social Media 2021; and e-Health 2021*. 2021. P. 144–151. https://doi.org/10.33965/eh2021_2021061018
15. Agrawal S. *Image Processing in Python – The Computer Vision Techniques*. 2021 [Электронный ресурс]. <https://www.analyticsvidhya.com/blog/2021/08/image-processing-in-python-the-computer-vision-techniques/> (дата обращения: 25.04.2022)
16. Park J.-S., Lee S.-W., Park U. R Peak detection method using wavelet transform and modified shannon energy envelope // *Journal of Healthcare Engineering*. 2017. V. 2017. P. 4901017. <https://doi.org/10.1155/2017/4901017>
17. Nouira I., Abdallah A.B., Bedoui M.H., Dogui M. A robust R peak detection algorithm using wavelet transform for heart rate variability studies // *International Journal on Electrical Engineering and Informatics*. 2013, vol. 5, no. 3, pp. 270–284. <https://doi.org/10.15676/ijeei.2013.5.3.3>
3. *Health systems resilience during COVID-19: Lessons for building back better*. Ed. by A. Sagan, E. Webb, I. de la Mata, J. Figueras, M. McKee, N. Azzopardi-Muscat. WHO Regional Office for Europe, 2021.
4. Wang N.C., Jain S.K., Estes N.A.M., Barrington W.W., Bazaz R., Bhonsale A., Kancharla K., Shalaby A.A., Voigt A.H., Saba S., Priority plan for invasive cardiac electrophysiology procedures during the coronavirus disease 2019 (COVID-19) pandemic. *Journal of Cardiovascular Electrophysiology*, 2020, vol. 31, no. 6, pp. 1255–1258. <https://doi.org/10.1111/jce.14478>
5. Cook D.A., Oh S., Pusic M.V. Accuracy of physicians' electrocardiogram interpretations: A systematic review and meta-analysis. *JAMA Internal Medicine*, 2020, vol. 180, no. 11, pp. 1461–1471. <https://doi.org/10.1001/jamainternmed.2020.3989>
6. Javed A., Khan S.U., Ali L., Ali S., Imrana Y., Rahman A. Machine learning-based automated diagnostic systems developed for heart failure prediction using different types of data modalities: A systematic review and future directions. *Computational and Mathematical Methods in Medicine*, 2022, vol. 2022, pp. 9288452. <https://doi.org/10.1155/2022/9288452>
7. Martin-Isla C., Campello V.M., Izquierdo C., Raisi-Estabragh Z., Baeßler B., Petersen S.E., Lekadir K. Image-based cardiac diagnosis with machine learning: A review. *Frontiers in Cardiovascular Medicine*, 2020, vol. 7, pp. 1. <https://doi.org/10.3389/fcvm.2020.00001>
8. Attia Z.I., Noseworthy P.A., Lopez-Jimenez F., Asirvatham S.J., Deshmukh A.J., Gersh B.J., Carter R.E., Yao X., Rabinstein A.A., Erickson B.J., Kapa S., Friedman P.A. An artificial intelligence-enabled ECG algorithm for the identification of patients with atrial fibrillation during sinus rhythm: A retrospective analysis of outcome prediction. *Lancet*, 2019, vol. 394, pp. 861–867. [https://doi.org/10.1016/S0140-6736\(19\)31721-0](https://doi.org/10.1016/S0140-6736(19)31721-0)
9. Katoh T., Yashima M., Takahashi N., Watanabe E., Ikeda T., Kasamaki Y., Sumitomo N., Ueda N., Morita H., Hiraoka M. Expert consensus document on automated diagnosis of the electrocardiogram: The task force on automated diagnosis of the electrocardiogram in Japan. Part 2: Current status of inappropriate automated diagnosis is widely used electrocardiographs in Japan. *Journal of Arrhythmia*, 2021, vol. 37, no. 6, pp. 1427–1433. <https://doi.org/10.1002/joa3.12646>
10. Sangaiah A., Arumugam M., Bian G. An intelligent learning approach for improving ECG signal classification and arrhythmia analysis. *Artificial Intelligence in Medicine*, 2020, vol. 103, pp. 101788. <https://doi.org/10.1016/j.artmed.2019.101788>
11. Tse G., Lee S., Li A., Chang D., Li G., Zhou J., Liu T., Zhang Q. Automated electrocardiogram analysis identifies novel predictors of ventricular arrhythmias in brugada syndrome. *Frontiers in Cardiovascular Medicine*, 2021, vol. 7, pp. 618254. <https://doi.org/10.3389/fcvm.2020.618254>
12. Rueda C., Fernández I., Larriba Y., Rodríguez-Collado A., Canedo C. Compelling new electrocardiographic markers for automatic diagnosis. *Computer Methods and Programs in Biomedicine*, 2022, vol. 221, pp. 106807. <https://doi.org/10.1016/j.cmpb.2022.106807>
13. Fortune J.D., Coppa N.E., Haq K.T., Patel H., Tereshchenko L.G. *Digitizing ECG image: new fully automated method*. 2021. Available at: <https://www.medrxiv.org/content/10.1101/2021.07.13.21260461v1.full.pdf> (accessed: 25.04.2022)
14. Vatian A., Peredreev D., Rodiontsev K., Murzina A., Klevtsova E., Tatarinova A., Treshkur T., Shalyto A., Gusarova N. Helping paramedics in assessing a patient's condition based on ECG by means of mobile phone. *Proc. of the International Conferences ICT, Society, and Human Beings 2021; Web Based Communities and Social Media 2021; and e-Health 2021*, 2021, pp. 144–151. https://doi.org/10.33965/eh2021_2021061018
15. Agrawal S. *Image Processing in Python – The Computer Vision Techniques*. 2021. Available at: <https://www.analyticsvidhya.com/blog/2021/08/image-processing-in-python-the-computer-vision-techniques/> (accessed: 25.04.2022)
16. Park J.-S., Lee S.-W., Park U. R Peak detection method using wavelet transform and modified shannon energy envelope. *Journal of Healthcare Engineering*, 2017, vol. 2017, pp. 4901017. <https://doi.org/10.1155/2017/4901017>
17. Nouira I., Abdallah A.B., Bedoui M.H., Dogui M. A robust R peak detection algorithm using wavelet transform for heart rate variability studies. *International Journal on Electrical Engineering and Informatics*, 2013, vol. 5, no. 3, pp. 270–284. <https://doi.org/10.15676/ijeei.2013.5.3.3>

- Informatics. 2013. V. 5. N 3. P. 270–284. <https://doi.org/10.15676/ijeei.2013.5.3.3>
18. Bae T.W., Kwon K.K. ECG PQRST complex detector and heart rate variability analysis using temporal characteristics of fiducial points // *Biomedical Signal Processing and Control*. 2021. V. 66. P. 102291. <https://doi.org/10.1016/j.bspc.2020.102291>
 19. Camm A.J., Malik M., Yap Y.G. *Acquired Long QT Syndrome*. Blackwell Futura, 2004. 208 p.
 20. Национальные российские рекомендации по применению методики холтеровского мониторирования в клинической практике (проект) [Электронный ресурс]. URL: <https://scardio.ru/content/images/recommendation/HM.pdf>. (дата обращения: 20.09.2022).
 21. Goldenberg I., Moss A.J., Zareba W. QT interval: how to measure it and what is “normal” // *Journal of Cardiovascular Electrophysiology*. 2006. V. 17. N 3. P. 333–336. <https://doi.org/10.1111/j.1540-8167.2006.00408.x>
 22. Kalyakulina A., Yusipov I., Moskalenko V., Nikolskiy A., Kosonogov K., Zolotykh N., Ivanchenko M. *Lobachevsky University Electrocardiography Database* [Электронный ресурс]. URL: <https://physionet.org/content/ludb/1.0.1/> (дата обращения: 20.09.2022).
 23. Smith S.W., Khalil A., Henry T.D., Rosas M., Chang R.J., Heller K., Scharrer E., Ghorashi M., Pearce L.A. Electrocardiographic differentiation of early repolarization from subtle anterior ST-segment elevation myocardial infarction // *Annals of Emergency Medicine*. 2012. V. 60. N 1. P. 45–56.e2. <https://doi.org/10.1016/j.annemergmed.2012.02.015>
 24. Liu F.F., Liu C., Zhao L., Zhang X., Wu X., Xu X., Liu Y., Ma C., Wei S., He Z., Li J., Yin K., Eddie N. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection // *Journal of Medical Imaging and Health Informatics*. 2018. V. 8. N 7. P. 1368–1373. <https://doi.org/10.1166/jmih.2018.2442>
 25. Choi H.Y., Kim W., Kang G.H., Jang Y.S., Lee Y., Kim J.G., Lee N., Shin D.G., Bae W., Song Y. Diagnostic accuracy of the deep learning model for the detection of ST elevation myocardial infarction on electrocardiogram // *Journal of Personalized Medicine*. 2022. V. 12. N 3. P. 336. <https://doi.org/10.3390/jpm12030336>
 26. Chang K.-C., Hsieh P.-H., Wu M.-Y., Wang Y.-C., Wei J.-T., Shih E.S.-C., Hwang M.-J., Lin W.-Y., Lin W.-T., Lee K.-J., Wang T.-H. Usefulness of multi-labelling artificial intelligence in detecting rhythm disorders and acute ST-elevation myocardial infarction on 12-lead electrocardiogram // *European Heart Journal - Digital Health*. 2021. V. 2. N 2. P. 299–310. <https://doi.org/10.1093/ehjdh/ztab029>
 27. Liu Z., Mao H., Wu C.-Y., Feichtenhofer C., Darrell T., Xie S. A ConvNet for the 2020s // *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR*. 2022. P. 11966–11976. <https://doi.org/10.1109/CVPR52688.2022.01167>
 28. Bae T.W., Kwon K.K. ECG PQRST complex detector and heart rate variability analysis using temporal characteristics of fiducial points. *Biomedical Signal Processing and Control*, 2021, vol. 66, pp. 102291. <https://doi.org/10.1016/j.bspc.2020.102291>
 29. Camm A.J., Malik M., Yap Y.G. *Acquired Long QT Syndrome*. Blackwell Futura, 2004, 208 p.
 30. *National Russian recommendations for the use of Holter monitoring in clinical practice (draft)*. Available at: <https://scardio.ru/content/images/recommendation/HM.pdf>. (accessed: 20.09.2022). (in Russian)
 31. Goldenberg I., Moss A.J., Zareba W. QT interval: how to measure it and what is “normal”. *Journal of Cardiovascular Electrophysiology*, 2006, vol. 17, no. 3, pp. 333–336. <https://doi.org/10.1111/j.1540-8167.2006.00408.x>
 32. Kalyakulina A., Yusipov I., Moskalenko V., Nikolskiy A., Kosonogov K., Zolotykh N., Ivanchenko M. *Lobachevsky University Electrocardiography Database*. Available at: <https://physionet.org/content/ludb/1.0.1/> (accessed: 20.09.2022).
 33. Smith S.W., Khalil A., Henry T.D., Rosas M., Chang R.J., Heller K., Scharrer E., Ghorashi M., Pearce L.A. Electrocardiographic differentiation of early repolarization from subtle anterior ST-segment elevation myocardial infarction. *Annals of Emergency Medicine*, 2012, vol. 60, no. 1, pp. 45–56.e2. <https://doi.org/10.1016/j.annemergmed.2012.02.015>
 34. Liu F.F., Liu C., Zhao L., Zhang X., Wu X., Xu X., Liu Y., Ma C., Wei S., He Z., Li J., Yin K., Eddie N. An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. *Journal of Medical Imaging and Health Informatics*, 2018, vol. 8, no. 7, pp. 1368–1373. <https://doi.org/10.1166/jmih.2018.2442>
 35. Choi H.Y., Kim W., Kang G.H., Jang Y.S., Lee Y., Kim J.G., Lee N., Shin D.G., Bae W., Song Y. Diagnostic accuracy of the deep learning model for the detection of ST elevation myocardial infarction on electrocardiogram. *Journal of Personalized Medicine*, 2022, vol. 12, no. 3, pp. 336. <https://doi.org/10.3390/jpm12030336>
 36. Chang K.-C., Hsieh P.-H., Wu M.-Y., Wang Y.-C., Wei J.-T., Shih E.S.-C., Hwang M.-J., Lin W.-Y., Lin W.-T., Lee K.-J., Wang T.-H. Usefulness of multi-labelling artificial intelligence in detecting rhythm disorders and acute ST-elevation myocardial infarction on 12-lead electrocardiogram. *European Heart Journal — Digital Health*, 2021, vol. 2, no. 2, pp. 299–310. <https://doi.org/10.1093/ehjdh/ztab029>
 37. Liu Z., Mao H., Wu C.-Y., Feichtenhofer C., Darrell T., Xie S. A ConvNet for the 2020s. *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR*, 2022, pp. 11966–11976. <https://doi.org/10.1109/CVPR52688.2022.01167>

Авторы

Ватьян Александра Сергеевна — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-5483-716X>, alexvatyan@gmail.com

Гусарова Наталья Федоровна — кандидат технических наук, старший научный сотрудник, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-1361-6037>, natfed@list.ru

Добренко Наталья Викторовна — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-6206-8033>, graziokisa@yandex.ru

Змиевский Данил Александрович — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-0077-794X>, daniel.zmievskiy@gmail.com

Кабышев Максим Васильевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-1006-0408>, maxk6971@gmail.com

Полевая Татьяна Андреевна — программист, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-6131-0019>, tanusha2406@gmail.com

Татарникова Анна Андреевна — кандидат медицинских наук, старший научный сотрудник, старший научный сотрудник, Национальный медицинский исследовательский центр им. В.А. Алмазова, Санкт-

Authors

Alexandra S. Vatian — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-5483-716X>, alexvatyan@gmail.com

Natalia F. Gusarova — PhD, Senior Researcher, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-1361-6037>, natfed@list.ru

Natalia V. Dobrenko — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-6206-8033>, graziokisa@yandex.ru

Danil A. Zmievisky — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-0077-794X>, daniel.zmieviskiy@gmail.com

Maxim A. Kabyshev — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-1006-0408>, maxk6971@gmail.com

Tatiana A. Polevaya — Software Developer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-6131-0019>, tanusha2406@gmail.com

Anna A. Tatarinova — PhD (Medicine), Senior Researcher, Senior Researcher, Almazov National Medical Research Centre, Saint Petersburg,

Петербург, 197341, Российская Федерация,  6603195545, <https://orcid.org/0000-0002-9046-2457>, antsvet.18@mail.ru

Томилов Иван Вячеславович — старший лаборант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0003-1886-2867>, ivan-tomilov3@yandex.ru

197341, Russian Federation,  6603195545, <https://orcid.org/0000-0002-9046-2457>, antsvet.18@mail.ru

Ivan V. Tomilov — Senior Laboratory Assistant, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0003-1886-2867>, ivan-tomilov3@yandex.ru

Статья поступила в редакцию 15.08.2022

Одобрена после рецензирования 28.10.2022

Принята к печати 29.11.2022

Received 15.08.2022

Approved after reviewing 28.10.2022

Accepted 29.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1178-1186

УДК 004.8+ 65.011.56

Мультиагентная адаптивная маршрутизация агентами-клонами на основе многоголового внутреннего внимания с использованием обучения с подкреплением

Тимофей Александрович Грибанов¹, Андрей Александрович Фильченков²✉, Артур Александрович Азаров³, Анатолий Абрамович Шалыто⁴

^{1,2,3,4} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

³ Северо-Западный институт управления – филиал РАНХиГС, Санкт-Петербург, 199178, Российская Федерация

¹ t.hrybanau@gmail.com, <https://orcid.org/0000-0002-1151-3405>

² afilchenkov@itmo.ru✉, <https://orcid.org/0000-0002-1133-8432>

³ artur-azarov@yandex.ru, <https://orcid.org/0000-0003-3240-597X>

⁴ shalyto@mail.ifmo.ru, <https://orcid.org/0000-0002-2723-2077>

Аннотация

Предмет исследования. Регулярным условием, характерным для пакетной маршрутизации, а также задач транспортировки грузов и управления потоками, является изменчивость графа, на котором осуществляется маршрутизация. Это условие учитывают алгоритмы адаптивной маршрутизации, использующие обучение с подкреплением. Однако при значительных изменениях графа существующим алгоритмам маршрутизации требуется полное переобучение. **Метод.** Предложен новый метод, основанный на мультиагентном моделировании с агентами-клонами, с использованием новой архитектуры нейронной сети с многоголовым внутренним вниманием, которая предобучена в рамках парадигмы обучения с нескольких взглядов. Агент в такой парадигме использует вершину как вход, а его клоны помещены в вершины графа и осуществляют выбор соседа, которому следует передать объект. **Основные результаты.** Выполнен сравнительный анализ с существующим алгоритмом мультиагентной маршрутизации *DQN-LE-routing* по следующим этапам: предобучение и симуляция. Для каждого этапа рассмотрены запуски с помощью изменения топологии в процессе тестирования или симуляции. Эксперименты показали, что предложенный метод повышения адаптивности обеспечивает глобальную адаптивность, увеличивая время доставки при глобальных изменениях не более чем на 14,5 % от оптимального. **Практическая значимость.** Предложенный метод может быть использован для решения задач маршрутизации со сложными функциями оценки пути и динамически меняющимися топологиями графов, например, в транспортной логистике и для управления конвейерными лентами на производстве.

Ключевые слова

маршрутизация, мультиагентное обучение, обучение с подкреплением, адаптивная маршрутизация

Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект № 20-19-00700).

Ссылка для цитирования: Грибанов Т.А., Фильченков А.А., Азаров А.А., Шалыто А.А. Мультиагентная адаптивная маршрутизация агентами-клонами на основе многоголового внутреннего внимания с использованием обучения с подкреплением // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1178–1186. doi: 10.17586/2226-1494-2022-22-6-1178-1186

Multi-agent adaptive routing by multi-head-attention-based twin agents using reinforcement learning

Timofey A. Gribanov¹, Andrey A. Filchenkov²✉, Artur A. Azarov³, Anatoly A. Shalyto⁴

^{1,2,3,4} ITMO University, Saint Petersburg, 197101, Russian Federation

³ North-West Institute of Management – branch of the Russian Presidential Academy of National Economy and Public Administration, Saint Petersburg, 199178, Russian Federation

¹ t.hrybanau@gmail.com, <https://orcid.org/0000-0002-1151-3405>

² afilchenkov@itmo.ru✉, <https://orcid.org/0000-0002-1133-8432>

³ artur-azarov@yandex.ru, <https://orcid.org/0000-0003-3240-597X>

⁴ shalyto@mail.ifmo.ru, <https://orcid.org/0000-0002-2723-2077>

Abstract

A regular condition, typical for packet routing, for the problem of cargo transportation, and for the problem of flow control, is the variability of the graph. Reinforcement learning based adaptive routing algorithms are designed to solve the routing problem with this condition. However, with significant changes in the graph, the existing routing algorithms require complete retraining. To handle this challenge, we propose a novel method based on multi-agent modeling with twin-agents for which new neural network architecture with multi-headed internal attention is proposed, pre-trained within the framework of the multi-view learning paradigm. An agent in such a paradigm uses a vertex as an input, twins of the main agent are placed at the vertices of the graph and select a neighbor to which the object should be transferred. We carried out a comparative analysis with the existing *DQN-LE-routing* multi-agent routing algorithm on two stages: pre-training and simulation. In both cases, launches were considered by changing the topology during testing or simulation. Experiments have shown that the proposed adaptability enhancement method provides global adaptability by increasing delivery time only by 14.5 % after global changes occur. The proposed method can be used to solve routing problems with complex path evaluation functions and dynamically changing graph topologies, for example, in transport logistics and for managing conveyor belts in production.

Keywords

routing, multi-agent learning, reinforcement learning, adaptive routing

Acknowledgements

The study was supported by the grant from the Russian Science Foundation (project no. 20-19-00700).

For citation: Gribanov T.A., Filchenkov A.A., Azarov A.A., Shalyto A.A. Multi-agent adaptive routing by multi-head-attention-based twin agents using reinforcement learning. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1178–1186 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1178-1186

Введение

Задача маршрутизации состоит в поиске оптимального пути между двумя вершинами в графе [1]. Для решения данной классической оптимизационной постановки применяются алгоритмы дискретной математики, которые можно разделить на два типа: дистанционно-векторные алгоритмы (*distance-vector*) [2] и алгоритмы состояния канала связи (*link-state*) [3], использующие алгоритмы кратчайших путей.

Более сложные постановки учитывают специфику предметных областей или конкретных условий, в которых необходимо осуществлять маршрутизацию. Регулярным условием, характерным для пакетной маршрутизации, а также задач транспортировки грузов и управления потоками, является изменчивость графа, на котором осуществлена маршрутизация. Адаптивная маршрутизация [4, 5] — направление исследований, в котором предполагается находить решения для задачи маршрутизации, устойчивые к такого рода изменениям. Алгоритмы адаптивной маршрутизации основаны на машинном обучении и преимущественно используют мультиагентный подход на основе обучения с подкреплением [6, 7]. Отметим, что появляются научные работы, основанные на других принципах машинного обучения, например — на использовании трансформеров [8].

Изменения в топологии можно условно разделить на локальные и глобальные. Локальные — удаление

одной вершины или одного ребра, глобальные — существенное изменение всей топологии графа, например, его увеличение или уменьшение в два раза. Так как текущие алгоритмы маршрутизации с использованием обучения с подкреплением в своей основе содержат мультиагентное обучение [9, 10], они очень хорошо адаптируются к локальным изменениям, которые происходят в контексте одной вершины. Однако в случае глобальных изменений этим алгоритмам требуется полное переобучение, так как они адаптированы еще к старому графу.

В настоящей работе предложен алгоритм, решающий проблему глобальной адаптации с использованием обучения с подкреплением.

Алгоритмы маршрутизации

Маршрутизация и адаптивная маршрутизация.

Пусть дан граф $G = \langle V, E \rangle$, $E = \{(u, v, param) | u, v \in V\}$, где *param* — свойства ребра, такие как пропускная способность или задержка. Также могут быть известны дополнительные параметры в зависимости от конкретной области применения, в том числе и к вершинам. Задана стоимость доставки объекта между каждой парой вершин $cost(u, v, cont)$, где $u, v \in V$ и *cont* — обстоятельства передачи (например, время).

Большинство современных алгоритмов адаптивной маршрутизации направлены на работу со временем или стоимостью доставки, которые задаются пореберно, и

время (стоимость) маршрута — как сумма значений функций для каждого ребра, принадлежащего маршруту. Такие функции называются декомпозируемыми. В более сложном случае оценка решения может производиться недекомпозируемыми функциями.

Изменения в топологии выполним путем удаления или восстановления ребер. Строгого и формального разделения на локальные и глобальные изменения провести не представляется возможным, однако, назовем изменение локальным, если оно включает удаление или восстановление одного ребра на значительном промежутке времени, чтобы алгоритм успел провести адаптацию и показать релевантные результаты. Глобальные изменения — удаление, восстановление, добавление вершин и соответствующих им ребер.

Для оценки глобальных изменений введем понятие δ -окрестности графа. δ -окрестность исходного графа $G = \langle V, E \rangle$ — любой граф, полученный из исходного с помощью удаления или добавления δ вершин из или в граф. Таким образом, новый граф $G' = \langle V', E' \rangle$ будет удовлетворять следующему неравенству: $|V| - \delta \leq |V'| \leq |V| + \delta$.

Алгоритмы адаптивной маршрутизации недекомпозируемых функций. Для решения проблемы изменений в топологии большое распространение получили алгоритмы, использующие обучение с подкреплением. Родоначальником таких алгоритмов стал *Q-routing* [10], применяющий мультиагентный подход для обучения. В данном алгоритме каждой вершине соответствует отдельный независимый агент, который отвечает только за маршрутизацию через одну свою вершину. Сами агенты используют алгоритм обучения с подкреплением *Q-learning* [11], который после каждого получения награды обновляет функцию полезности Q состояния, на основе которой впоследствии будет выбрана стратегия принятия решения. В *Q-routing* каждому агенту соответствует таблица $Q(n, d)$ с оценками минимальной стоимости пути от текущей вершины до вершины d , если следующей вершиной на пути будет вершина n , которая обновляется за счет коммуникации между агентами после пересылок. Такая модель доказала свою эффективность при локальных изменениях и показала высокий уровень адаптации по сравнению с дистанционно-векторными алгоритмами и алгоритмами состояния канала связи. Однако алгоритм *Q-routing* предполагает, что оптимизируемая функция декомпозируема.

Более современные решения для аппроксимации Q -функции используют глубокие сети. Первой была предложена *Deep Q-Network (DQN)* [12]. На основе этого решения разработан алгоритм маршрутизации *DQN-routing* [13], объединяющий алгоритм *Q-routing* и идеи *DQN*, который привнес следующие изменения.

- Вместо таблицы Q использована нейронная сеть с прямым распространением, а обновления происходят в ее весах при помощи градиентного спуска.
- Оценка минимальной стоимости пути от текущей вершины до конечной заменена на математическое ожидание оценок при возможных переходах.
- Для принятия решения о следующей вершине в пути вместо простого нахождения минимума среди возможных переходов использована функция

Softmax [14], которая применена к полученным значениям.

- Использован этап предобучения с помощью обучения с учителем. Данные для предобучения сгенерированы случайным образом на имеющемся графе без изменений, а в качестве правильного ответа взята длина кратчайшего пути (алгоритм Дейкстры).

Архитектура нейросети — сеть прямого распространения с двумя слоями по 64 нейрона. В качестве функции активации выбран гиперболический тангенс. На вход нейронной сети подаются вершины в *one-hot* кодировании: текущая вершина, конечная вершина объекта и соседи текущей вершины. Также в качестве дополнительного параметра на вход подается матрица смежности графа.

Алгоритм *DQN-routing* показал приемлемые результаты для оптимизации сложных функций, оценивающих качество доставки, однако его недостатком является размер нейронной сети. Этот размер зависит от размера самого графа, так как на его выходе сформирован вектор предсказаний для всех вершин.

Следующим логичным усовершенствованием алгоритма *DQN-routing* стал *DQN-LE-routing* [15]. Все основные компоненты алгоритма *DQN-routing* оставлены, изменения носят точечный характер и служат для избавления от его недостатков: вместо *one-hot* кодирования подающихся на вход вершин использованы их графовые векторные представления на основе алгоритма *Laplacian Eigenmaps (LE)* [16]; на входе графа подаются не все соседи сразу, а только один, и аналогично на выходе получается предсказание только для одного конкретного соседа; функцией активации вместо гиперболического тангенса стала *ReLU*: $\max(x, 0)$.

Описание разработанного метода

Основные положения нового метода. Предлагаемый метод основан на трех идеях.

1. Алгоритм должен рассматривать текущую вершину как вход. Данный алгоритм, будучи обученным, позволяет добавлять новые вершины к топологии, не переобучаясь каждый раз на новой. В результате исследований возможных способов реализации данной идеи осуществлен переход на обучение модели на всем графе, а не на каждой вершине отдельно. Полученный алгоритм остается мультиагентным, однако все агенты являются клонами (*twin agents*) [17].
2. Вследствие отказа от мультиагентности, модель единого агента должна быть сложнее, чем модели агентов, использующихся в рассмотренных в разделе «Алгоритмы маршрутизации» алгоритмах, поскольку он должен быть способен оперировать информацией как о графе в целом, так и о ситуации в конкретных вершинах.
3. Обучение сложных моделей требует большей обучающей выборки, поэтому использовано множественное предобучение на графах разного размера в пределах некоторой δ -окрестности изначального графа. Данное условие дает возможность применения предобученной модели на любом случайном графе в этой δ -окрестности.

Ранее в предобучении для одного графа использовалось большое количество синтетических данных, на несколько порядков превышающее размеры самого графа. Это давало возможность обучить нейронную сеть максимально точно под конкретный граф. В новом подходе нет необходимости в точном обучении под конкретный граф, более того, оно может ухудшать результаты, так как каждый раз будет возникать проблема переобучения нейронной сети на каждом графе. Для успешной работы предобучение должно охватывать как можно больше разной информации от разных графов из δ -окрестности. По этой причине для каждого нового графа при множественном обучении генерируется небольшой набор синтетических данных, что позволяет увеличить число графов, используемых при обучении. Такая задача в машинном обучении имеет название обучение с нескольких взглядов (*few-shot learning*) [18] и заключается в том, что нейронная сеть обучается на данных, содержащих в себе ограниченное количество информации. Такой способ обучения позволяет модели быть более обобщенной, что в контексте задачи настоящей работы означает способность обобщаться по графам. Традиционные модели, обученные классическим способом, не способны различать классы, отсутствующие в обучающем наборе данных, в то время как метод обучения с нескольких взглядов позволяет нейронным сетям разделять два и более классов, которых нет в обучающих данных.

Мультиагентное управление с агентами-клонами. Для сохранения совместимости с моделью мультиагентного обучения использовано множество агентов. Все агенты, прикрепленные к вершинам, имеют второстепенный характер и привязаны к одному дополнительно введенному главному агенту, передающего логику обработки решений по маршрутизации.

Управляющая модель основана на отправлении сообщений между агентами, которые определяют их логи-

ку поведения. Основные сообщения бывают двух видов: необходимость совершить какое-то действие (переслать объект для маршрутизации); получение новой информации (награды за совершенное раннее действие). Также имеются другие служебные сообщения, которые необходимы для полноценной работы симуляции.

В управляющей системе каждый второстепенный агент хранит в себе лишь информацию о том, к какой вершине он принадлежит, и ссылку на главного агента. При получении сообщения второстепенные агенты пересылают его главному агенту, добавляя к сообщению информацию о своей вершине.

Обработка сообщений главным агентом схожа с аналогичной обработкой любым из агентов рассмотренных ранее алгоритмов. Добавим функционал для поддержки сообщений от второстепенных агентов и чтения информации о вершине, которая теперь приходит вместе с сообщением, а не хранится в самом агенте. Изменения в схеме взаимодействия агентов при исходном мультиагентном управлении и управлении на основе агентов-клонов показаны на рис. 1.

Архитектура нейронной сети. Рассмотрим вход предлагаемой нейронной сети. При одном агенте информации только о текущей и конечной вершинах недостаточно, так как нейронная сеть теперь должна делать предсказания в контексте всего графа. Для решения данной задачи необходимо поочередно подавать на вход векторные представления всех соседей вершины, в которую пришел объект. Такой же подход применен и в алгоритме *DQN-LE-routing*. Ключевое отличие — добавление матрицы смежности к входным переменным, так как необходимо, чтобы нейронная сеть была способна подстраиваться и обучаться на графах разных размеров и топологий, а информации о соседних вершинах в таком случае недостаточно.

Отметим, что добавление матрицы смежности для использования на графах разных размерностей добав-

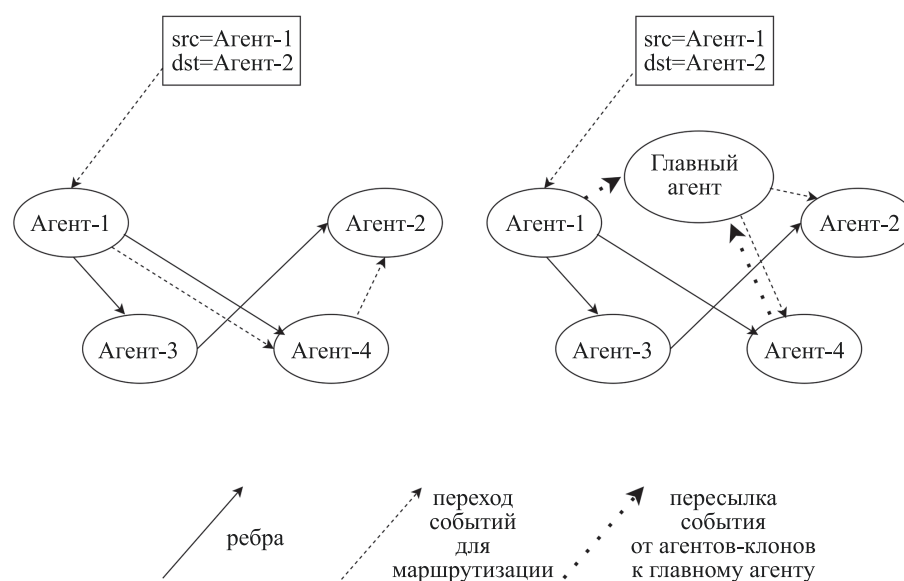


Рис. 1. Изменение схемы взаимодействия агентов в управляющей модели
src — текущая вершина, dst — вершина назначения

Fig. 1. Changes in the agent interaction schemes in the control model, src is the operating vertex, dst is the destination vertex

ляют новую проблему, суть которой состоит в том, что число значений на входе у нейронной сети не является константой. В основу решения проблемы входит понятие δ -окрестности графа. Значение δ задается и фиксируется в самом начале работы алгоритма, поэтому число вершин в наибольшем возможном графе в процессе всех изменений будет равно $n_{\max} = |V| + \delta$. Это ограничение позволяет передавать на вход нейронной сети матрицу смежности всегда определенного размера $n_{\max} \times n_{\max}$, которая будет заполняться нулевыми строками и столбцами в случае, когда текущий граф не максимального размера.

Для того чтобы увеличить сложность модели, использован механизм многоголового внутреннего внимания (*multi-head self-attention*) [19], способный к выявлению закономерности между разнесенными по времени подачи входами.

Многоголовое внимание в предлагаемой архитектуре применяется дважды. Первый слой внимания распространяется на векторные представления вершин на входе для нахождения зависимостей в них. Для второго слоя механизма внимания к результату первого слоя добавляется вектор того же размера, отвечающий за матрицу смежности.

На рис. 2 представлена схема работы механизма внимания для преобразования векторного представления текущей вершины *src*, используя информацию о векторных представлениях конечной и смежной вершин. В качестве *query*, *key* и *value* использовано одно и то же входное векторное представление соответствующей вершины.

Другим важным аспектом архитектуры нейронной сети являются слои нормализации. Нормализация пре-

дотвращает сильное изменение диапазона значений в слоях, что приводит к тому, что модель обучается быстрее и обладает лучшей способностью к обобщению [20].

Для использования второго слоя механизма внимания с помощью простой сети прямого распространения, состоящей из двух слоев, матрица смежности преобразовывается в вектор той же длины, что и векторное представление вершин. Затем этот вектор добавляется к измененным после первого слоя механизма внимания векторным представлениям вершин на входе, и полученные четыре вектора передаются на второй слой.

В связи с тем, что слой нейронной сети, представляющий сеть прямого распространения недостаточен для покрытия возросшего количества входных данных, он был расширен до четырех линейных слоев и одного *dropout*-слоя [21]. Также увеличено число нейронов в линейных слоях. Архитектура новой нейронной сети представлена на рис. 3.

Предобучения агента. Предобучение выполнено на синтетически сгенерированных данных под существующую нейронную сеть в конкретном графе. Процесс предобучения разбит на несколько этапов, в начале каждого из которых граф изменяется путем удаления или добавления δ вершин. Таким образом, за счет сэмпирования аппроксимируется предобучение на всех графах из δ -окрестности графа G . Суммарно итераций предобучения на δ -окрестностях производится достаточно много, но для каждого отдельного графа из этой окрестности их осуществляется достаточно мало. В результате нейронная сеть обучается на большом количестве разрозненных данных, что позволяет ей охватить и обучиться на всей δ -окрестности графа.

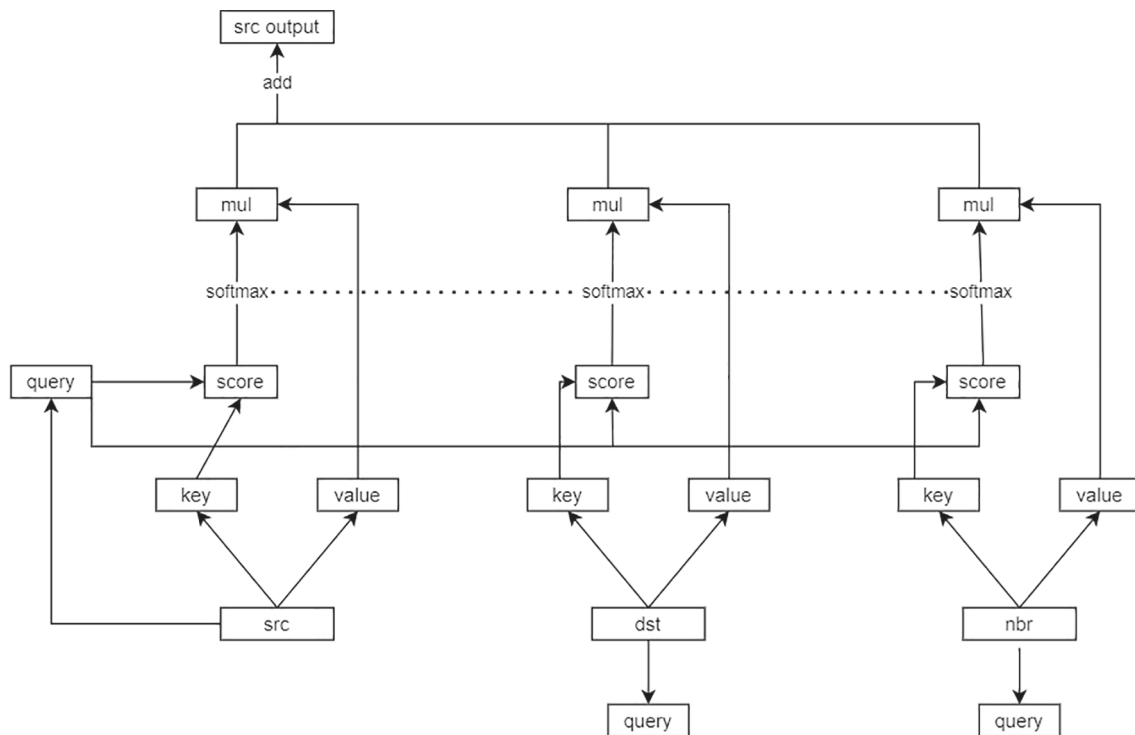


Рис. 2. Фрагмент архитектуры с внутренним вниманием для обработки входящей вершины *src*

Fig. 2. Scheme of the self-attention mechanism for processing an input vertex *src*

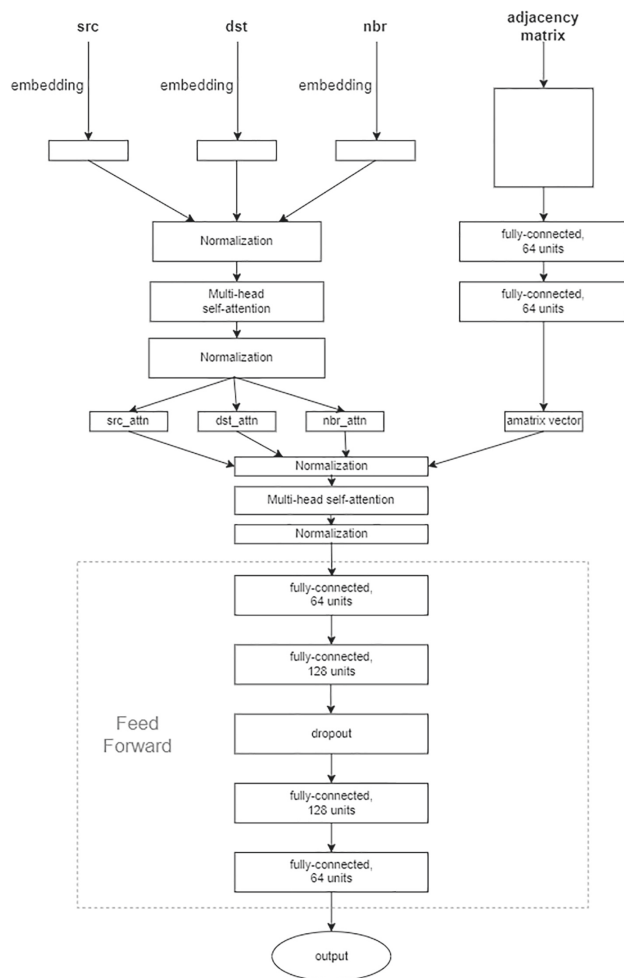


Рис. 3. Предлагаемая архитектура нейронной сети агента
Fig. 3. The proposed architecture of the neural network of the agent

Генерация синтетических данных для предобучения сопровождается дополнительным параметром, который означает, на сколько и в какую сторону нужно изменить граф. Удаление вершин происходит случайным образом и сопровождается проверкой связности полученного графа. При добавлении вершин требуется и добавление новых случайных ребер, число которых определяется также случайным образом в пределах некоторого диапазона, который задается в зависимости от размеров изначального графа. В качестве данных используются следующие значения: случайная начальная и конечная вершины; одна из вершин-соседей; наименьшая стоимость пути от стартовой вершины до конечной, проходящего через соответствующую смежную вершину; матрица смежности графа для возможности подсчета векторных представлений вершин во время предобучения.

Матрица смежности в генераторе синтетических данных передается заведомо большого размера для покрытия всей δ -окрестности. При этом во время предобучения для подсчета векторных представлений вершин матрица трансформируется в свой оригинальный вид без лишних строк и столбцов. Это необходимо для корректного подсчета векторных представлений вершин алгоритмом *LE* [16].

Описание работы агента. За предобучением следует фаза работы агента в среде. Агент использует ту же нейронную сеть, которая прошла стадию предобучения, и поэтому создает значительный прирост в показателях метрик. Так как обучению с подкреплением не приходится собирать информацию с нуля, это может привести к тому, что нейросеть «не сойдется» к оптимальному решению даже в простейших случаях [13].

В связи с измененной архитектурой нейронной сети, алгоритм обучения с подкреплением, отвечающий за интерпретируемый результат маршрутизации — определение следующей вершины на пути, состоит из следующих шагов.

Шаг 1. Нейронная сеть возвращает ожидаемую стоимость пути до конечной вершины при переходе в конкретного соседа.

Шаг 2. Запустив шаг 1 для каждого соседа, формируется список стоимостей возможных путей. Выбрав минимальный путь из полученного списка, можно было бы перейти в соответствующего соседа, но это — детерминистическая стратегия, и она не подходит для используемого подхода.

Шаг 3. К полученному списку применим стандартную функцию *Softmax* [14] для получения распределения вероятностных переходов, формируя тем самым стохастическую стратегию агента.

Шаг 4. На основе полученного распределения случайным образом выберем соседа, к которому будет совершен переход.

Экспериментальное исследование

Проведено экспериментальное исследование, которое состоит из двух частей. В первой части исследован этап предобучения, на котором проверено, что разработанный алгоритм действительно улучшает качество работы при увеличении числа подаваемых ему на вход графов. Во второй части проведено сравнение с алгоритмом *DQN-LE-routing* [15].

Для запуска процесса предобучения один из главных параметров — значение δ , в окрестности которого изменяются графы и происходит обучение. В качестве исходного графа для экспериментов выберем граф с 12 вершинами и 17 ребрами.

Для сравнения используем симуляционную систему, на которой проведено исследование алгоритма *DQN-LE-routing*. В системе смоделируем конвейерную систему, и оптимизируем функцию, которая является суммой времени доставки и затраченной на передачу электроэнергии, зависящей от маршрутизации всех сумок.

Результаты предобучения разработанного метода. В предобучении используем метод нескольких взглядов. В результате сгенерируем большое число графов разных размеров из δ -окрестности исходного графа путем удаления случайных вершин с соответствующими ребрами или добавления новых случайно сгенерированных вершин. Отметим, что на вход нейронной сети подавался ограниченный набор синтетических данных для каждого сгенерированного графа.

Значение δ увеличим итеративно от запуска к запуску, таким образом проверяя модель на все больших

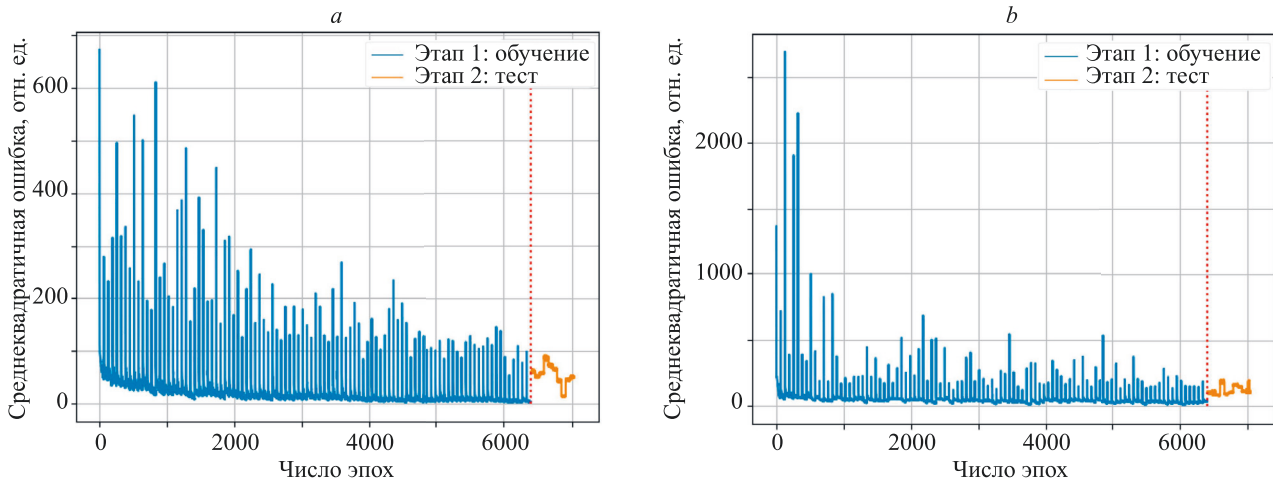


Рис. 4. Результаты предобучения со значениями $\delta = 1$ (a) и $\delta = 5$ (b)

Fig. 4. Results of pre-training with $\delta = 1$ (a) и $\delta = 5$ (b)

и больших изменениях. В качестве функции потерь выбрана среднеквадратичная ошибка. На рис. 4 показаны этапы результатов предобучения для стартового и последнего исследованных значений δ . Этап 1 — предобучение с градиентным спуском, которое длится до момента пересечения с красной пунктирной линией. Этап 2 — проверка нейронной сети.

Видно, что на графиках при этапе 1 кривые стабильно сходятся с присутствием пиков, которые соответствуют моментам изменения графа. Значения пиков убывают, вследствие чего можно сделать вывод, что нейронная сеть успешно обучается в текущей δ -окрестности, и при изменении топологии она не переобучается с нуля, о чем в том числе свидетельствуют результаты этапа 2. Отметим, что сразу после достижения пика функция ошибки стремительно идет вниз, что означает успешное обучение нейронной сети и в рамках одного графа.

Так как цель предлагаемого алгоритма — выявить его способность адаптироваться к как можно большему изменению, произведем запуск с увеличенным значением δ (рис. 4, b). Видно, что предложенный метод успешно справляется даже при изменениях топологии в два раза.

Сравнение с существующим алгоритмом. В качестве сравнения предобучения выбран алгоритм *DQN-LE-routing*, так как он показал лучшие результаты в адаптивной маршрутизации. Результаты сравнения при значении $\delta = 3$ представлены на рис. 5. Видно, что предложенный в данной работе новый метод значительно превосходит показатели алгоритма *DQN-LE-routing*.

Для сравнения результатов симуляции был взят алгоритм *DQN-LE-routing* в единственном экземпляре. После предобучения алгоритм *DQN-LE-routing* и новый метод запущены в симуляционной модели, в которую была добавлена возможность полностью изменять топологию. Результаты запусков представлены на рис. 6.

Из рис. 6, a видно, что после изменения топологии, которое произошло в середине всей симуляции, алгоритм *DQN-LE-routing* более чем в пять раз ухудшил

свое качество, после чего дообучился, однако все еще не достиг первоначальных показателей. При этом для разработанного метода изменение топологии никак не повлияло на результаты.

На рис. 6, b в процессе симуляции выполнено два изменения топологии. Заметно большое превосходство нового метода, над существующим *DQN-LE-routing*.

В рассмотренных экспериментах значение среднего времени доставки разработанного алгоритма при глобальных изменениях увеличивалось не более чем на 14,5 % в отличие от алгоритма *DQN-LE-routing*, у которого среднее время доставки увеличивалось в разы. На основе проведенных экспериментов можно заключить, что предложенный метод успешно сходится как на этапе предобучения, так и при последующей симуляции с использованием обучения с подкреплением.

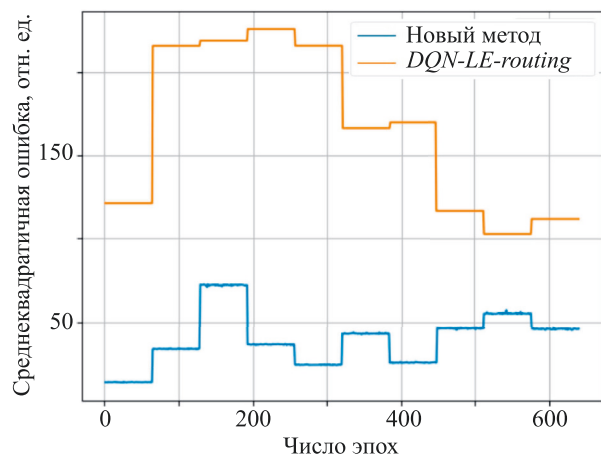


Рис. 5. Сравнение результатов предобучения нового метода и алгоритма *DQN-LE-routing* в нескольких экземплярах со значением $\delta = 3$

Fig. 5. Comparison of the pre-training results of the new method and the *DQN-LE-routing* algorithm in several instances with the value $\delta = 3$

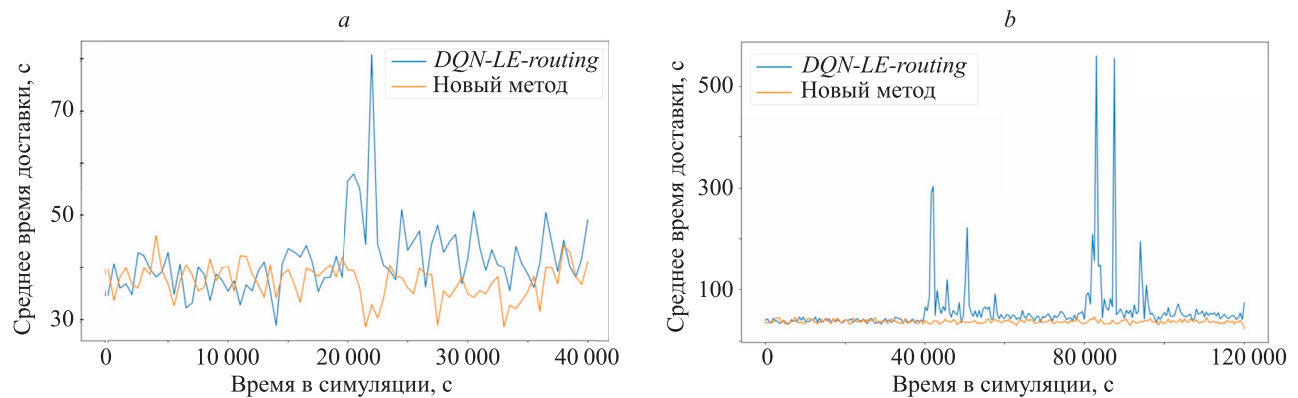


Рис. 6. Сравнение результатов работы нового метода и алгоритма *DQN-LE-routing* при однократном (а) и двукратном (б) изменении топологии

Fig. 6. Comparison of the results of the new method (New algo) and the *DQN-LE-routing* algorithm with a single (a) and two (b) topology change

Заключение

В работе предложен метод повышения адаптивности алгоритма мультиагентной маршрутизации, способный адаптироваться к любым изменениям графа, разработанный с учетом подходов, использующихся в существующих решениях в области обучения с подкреплением, и с применением нового подхода, как и в самой структуре и компонентах алгоритма, так и для обучения нейросети.

Выполнено сравнение метода с алгоритмом *DQN-LE-routing*, являющимся последним и наилучшим алго-

ритмом маршрутизации с использованием обучения с подкреплением по части архитектуры нейронной сети. По результатам экспериментов установлено, что при глобальных изменениях структуры графа разработанный метод значительно превосходит указанный алгоритм как на этапе предобучения, так и во время работы.

Предложенный метод может быть использован для решения задач маршрутизации со сложными функциями оценки пути и динамически меняющимися топологиями графов, например, в транспортной логистике и для управления конвейерными лентами на производстве.

Литература

1. Toth P., Vigo D. An overview of vehicle routing problems // *The Vehicle Routing Problem*. SIAM, 2002. P. 1–26. <https://doi.org/10.1137/1.9780898718515.ch1>
2. Vutukury S., Garcia-Luna-Aceves J.J. MDVA: A distance-vector multipath routing protocol // *Proc. 20th Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM)*. V. 1. P. 557–564. <https://doi.org/10.1109/INFOCOM.2001.916780>
3. Clausen T., Jacquet P. Optimized link state routing protocol (OLSR). 2003. N RFC3626. <https://doi.org/10.17487/RFC3626>
4. Sweda T.M., Dolinskaya I.S., Klabjan D. Adaptive routing and recharging policies for electric vehicles // *Transportation Science*. 2017. V. 51. N 4. P. 1326–1348. <https://doi.org/10.1287/trsc.2016.0724>
5. Puthal M.K., Singh V., Gaur M.S., Laxmi V. C-Routing: An adaptive hierarchical NoC routing methodology // *Proc. of the 2011 IEEE/IFIP 19th International Conference on VLSI and System-on-Chip*. 2011. P. 392–397. <https://doi.org/10.1109/VLSISoC.2011.6081616>
6. Zeng S., Xu X., Chen Y. Multi-agent reinforcement learning for adaptive routing: A hybrid method using eligibility traces // *Proc. of the 16th IEEE International Conference on Control & Automation (ICCA'20)*. 2020. P. 1332–1339. <https://doi.org/10.1109/ICCA51439.2020.9264518>
7. Ibrahim A.M., Yau K.L.A., Chong Y.W., Wu C. Applications of multi-agent deep reinforcement learning: models and algorithms // *Applied Sciences*. 2021. V. 11. N 22. P. 10870. <https://doi.org/10.3390/app112210870>
8. Bono G., Dibangoye J.S., Simonin O., Matignon L., Pereyron F. Solving multi-agent routing problems using deep attention mechanisms // *IEEE Transactions on Intelligent Transportation Systems*. 2021. V. 22. N 12. P. 7804–7813. <https://doi.org/10.1109/TITS.2020.3009289>
9. Kang Y., Wang X., Lan Z. Q-adaptive: A multi-agent reinforcement learning based routing on dragonfly network // *Proc. of the 30th International Symposium on High-Performance Parallel and*

References

1. Toth P., Vigo D. An overview of vehicle routing problems. *The Vehicle Routing Problem*. SIAM, 2002, pp. 1–26. <https://doi.org/10.1137/1.9780898718515.ch1>
2. Vutukury S., Garcia-Luna-Aceves J.J. MDVA: A distance-vector multipath routing protocol. *Proc. 20th Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM)*, vol. 1, pp. 557–564. <https://doi.org/10.1109/INFOCOM.2001.916780>
3. Clausen T., Jacquet P. *Optimized link state routing protocol (OLSR)*, 2003, no. RFC3626. <https://doi.org/10.17487/RFC3626>
4. Sweda T.M., Dolinskaya I.S., Klabjan D. Adaptive routing and recharging policies for electric vehicles. *Transportation Science*, 2017, vol. 51, no. 4, pp. 1326–1348. <https://doi.org/10.1287/trsc.2016.0724>
5. Puthal M.K., Singh V., Gaur M.S., Laxmi V. C-Routing: An adaptive hierarchical NoC routing methodology. *Proc. of the 2011 IEEE/IFIP 19th International Conference on VLSI and System-on-Chip*, 2011, pp. 392–397. <https://doi.org/10.1109/VLSISoC.2011.6081616>
6. Zeng S., Xu X., Chen Y. Multi-agent reinforcement learning for adaptive routing: A hybrid method using eligibility traces. *Proc. of the 16th IEEE International Conference on Control & Automation (ICCA'20)*, 2020, pp. 1332–1339. <https://doi.org/10.1109/ICCA51439.2020.9264518>
7. Ibrahim A.M., Yau K.L.A., Chong Y.W., Wu C. Applications of multi-agent deep reinforcement learning: models and algorithms. *Applied Sciences*, 2021, vol. 11, no. 22, pp. 10870. <https://doi.org/10.3390/app112210870>
8. Bono G., Dibangoye J.S., Simonin O., Matignon L., Pereyron F. Solving multi-agent routing problems using deep attention mechanisms. *IEEE Transactions on Intelligent Transportation Systems*, 2021, vol. 22, no. 12, pp. 7804–7813. <https://doi.org/10.1109/TITS.2020.3009289>
9. Kang Y., Wang X., Lan Z. Q-adaptive: A multi-agent reinforcement learning based routing on dragonfly network. *Proc. of the 30th International Symposium on High-Performance Parallel and*

- Distributed Computing, 2021. P. 189–200. <https://doi.org/10.1145/3431379.3460650>
10. Choi S., Yeung D.Y. Predictive Q-routing: A memory-based reinforcement learning approach to adaptive traffic control // *Advances in Neural Information Processing Systems*. 1995. V. 8. P. 945–951.
 11. Watkins C.J., Dayan P. Q-learning // *Machine Learning*. 1992. V. 8. N 3. P. 279–292. <https://doi.org/10.1023/A:1022676722315>
 12. Mnih V., Kavukcuoglu K., Silver D., Graves A., Antonoglou I., Wierstra D., Riedmiller M. Playing atari with deep reinforcement learning // *arXiv*. 2013. arXiv:1312.5602. <https://doi.org/10.48550/arXiv.1312.5602>
 13. Mukhutdinov D., Filchenkov A., Shalyto A., Vyatkin V. Multi-agent deep learning for simultaneous optimization for time and energy in distributed routing system // *Future Generation Computer Systems*. 2019. V. 94. P. 587–600. <https://doi.org/10.1016/j.future.2018.12.037>
 14. Gao B., Pavel L. On the properties of the softmax function with application in game theory and reinforcement learning // *arXiv*. 2017. arXiv:1704.00805. <https://doi.org/10.48550/arXiv.1704.00805>
 15. Мухудинов Д. Децентрализованный алгоритм управления конвейерной системой с использованием методов мультиагентного обучения с подкреплением: магистерская диссертация. СПб.: Университет ИТМО, 2019. 92 с. [Электронный ресурс]. URL: http://is.ifmo.ru/diploma-theses/2019/2_5458464771026191430.pdf (дата обращения: 01.10.2022)
 16. Belkin M., Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering // *Advances in Neural Information Processing Systems*. 2001. P. 585–591. <https://doi.org/10.7551/mitpress/1120.003.0080>
 17. Benea M.T., Florea A.M., Seghrouchni A.E.F. CAmI: An agent oriented-language for the collective development of Aml environments // *Proc. of the 20th International Conference on Control Systems and Computer Science (CSCS)*. 2015. P. 749–756. <https://doi.org/10.1109/CSCS.2015.136>
 18. Wang Y., Yao Q., Kwok J.T., Ni L.M. Generalizing from a few examples: A survey on few-shot learning // *ACM Computing Surveys*. 2020. V. 53. N 3. P. 63. <https://doi.org/10.1145/3386252>
 19. Liu J., Chen S., Wang B., Zhang J., Li N., Xu T. Attention as relation: learning supervised multi-head self-attention for relation extraction // *Proc. of the 19th International Joint Conferences on Artificial Intelligence (IJCAI)*. 2020. P. 3787–3793. <https://doi.org/10.24963/ijcai.2020/524>
 20. Sola J., Sevilla J. Importance of input data normalization for the application of neural networks to complex industrial problems // *IEEE Transactions on Nuclear Science*. 1997. V. 44. N 3. P. 1464–1468. <https://doi.org/10.1109/23.589532>
 21. Baldi P., Sadowski P.J. Understanding dropout // *Advances in Neural Information Processing Systems*. 2013. V. 26. P. 26–35.
 - Distributed Computing*, 2021, pp. 189–200. <https://doi.org/10.1145/3431379.3460650>
 10. Choi S., Yeung D.Y. Predictive Q-routing: A memory-based reinforcement learning approach to adaptive traffic control. *Advances in Neural Information Processing Systems*, 1995, vol. 8, pp. 945–951.
 11. Watkins C.J., Dayan P. Q-learning. *Machine Learning*, 1992, vol. 8, no. 3, pp. 279–292. <https://doi.org/10.1023/A:1022676722315>
 12. Mnih V., Kavukcuoglu K., Silver D., Graves A., Antonoglou I., Wierstra D., Riedmiller M. Playing atari with deep reinforcement learning. *arXiv*, 2013, arXiv:1312.5602. <https://doi.org/10.48550/arXiv.1312.5602>
 13. Mukhutdinov D., Filchenkov A., Shalyto A., Vyatkin V. Multi-agent deep learning for simultaneous optimization for time and energy in distributed routing system. *Future Generation Computer Systems*, 2019, vol. 94, pp. 587–600. <https://doi.org/10.1016/j.future.2018.12.037>
 14. Gao B., Pavel L. On the properties of the softmax function with application in game theory and reinforcement learning. *arXiv*, 2017, arXiv:1704.00805. <https://doi.org/10.48550/arXiv.1704.00805>
 15. Mukhutdinov D. *Decentralized conveyor system control algorithm using multi-agent reinforcement learning methods*. MSc Dissertation. St. Petersburg, ITMO University, 2019, 92 p. Available at: http://is.ifmo.ru/diploma-theses/2019/2_5458464771026191430.pdf (accessed: 01.10.2022). (in Russian)
 16. Belkin M., Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Advances in Neural Information Processing Systems*, 2001, pp. 585–591. <https://doi.org/10.7551/mitpress/1120.003.0080>
 17. Benea M.T., Florea A.M., Seghrouchni A.E.F. CAmI: An agent oriented-language for the collective development of Aml environments. *Proc. of the 20th International Conference on Control Systems and Computer Science (CSCS)*, 2015, pp. 749–756. <https://doi.org/10.1109/CSCS.2015.136>
 18. Wang Y., Yao Q., Kwok J.T., Ni L.M. Generalizing from a few examples: A survey on few-shot learning. *ACM Computing Surveys*, 2020, vol. 53, no. 3, pp. 63. <https://doi.org/10.1145/3386252>
 19. Liu J., Chen S., Wang B., Zhang J., Li N., Xu T. Attention as relation: learning supervised multi-head self-attention for relation extraction. *Proc. of the 19th International Joint Conferences on Artificial Intelligence (IJCAI)*, 2020, pp. 3787–3793. <https://doi.org/10.24963/ijcai.2020/524>
 20. Sola J., Sevilla J. Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Transactions on Nuclear Science*, 1997, vol. 44, no. 3, pp. 1464–1468. <https://doi.org/10.1109/23.589532>
 21. Baldi P., Sadowski P.J. Understanding dropout. *Advances in Neural Information Processing Systems*, 2013, vol. 26, pp. 26–35.

Авторы

Грибанов Тимофей Александрович — студент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-1151-3405>, t.hrybanau@gmail.com

Фильченков Андрей Александрович — кандидат физико-математических наук, инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-1133-8432>, afilchenkov@itmo.ru

Азаров Артур Александрович — кандидат технических наук, научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; заместитель директора, Северо-Западный институт управления — филиал РАНХиГС, Санкт-Петербург, 199178, Российская Федерация, <https://orcid.org/0000-0003-3240-597X>, artur-azarov@yandex.ru

Шалыто Анатолий Абрамович — доктор технических наук, профессор, главный научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-2723-2077>, shalyto@mail.ifmo.ru

Authors

Timofey A. Gribanov — Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-1151-3405>, t.hrybanau@gmail.com

Andrey A. Filchenkov — PhD (Physics & Mathematics), Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-1133-8432>, afilchenkov@itmo.ru

Artur A. Azarov — PhD, Scientific Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation; Deputy Director, North-West Institute of Management — branch of the Russian Presidential Academy of National Economy and Public Administration, Saint Petersburg, 199178, Russian Federation, <https://orcid.org/0000-0003-3240-597X>, artur-azarov@yandex.ru

Anatoly A. Shalyto — D. Sc., Professor, Chief Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-2723-2077>, shalyto@mail.ifmo.ru

Статья поступила в редакцию 28.10.2022

Одобрена после рецензирования 03.11.2022

Принята к печати 29.11.2022

Received 28.10.2022

Approved after reviewing 03.11.2022

Accepted 29.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1187-1196

УДК 004.8+ 65.011.56

Совместное обучение агентов и векторных представлений графов в задаче управления конвейерными лентами

Константин Евгеньевич Рыбкин¹, Андрей Александрович Фильченков²✉,
Артур Александрович Азаров³, Алексей Сергеевич Забашта⁴,
Анатолий Абрамович Шалыто⁵

^{1,2,3,4,5} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

³ Северо-Западный институт управления – филиал РАНХиГС, Санкт-Петербург, 199178, Российская Федерация

¹ matveich1903@gmail.com, <https://orcid.org/0000-0003-0856-4276>

² afilchenkov@itmo.ru✉, <https://orcid.org/0000-0002-1133-8432>

³ artur-azarov@yandex.ru, <https://orcid.org/0000-0003-3240-597X>

⁴ azabashata@itmo.ru, <https://orcid.org/0000-0002-7494-4309>

⁵ shalyto@mail.ifmo.ru, <https://orcid.org/0000-0002-2723-2077>

Аннотация

Предмет исследования. Рассмотрена задача маршрутизации системы конвейерных лент на основе мультиагентного подхода. В большинстве данных конвейерных систем багажных лент в аэропортах используются алгоритмы маршрутизации, основанные на ручном моделировании поведения конвейеров. Такой подход плохо масштабируем. Новые исследования в области машинного обучения предлагают решать задачу маршрутизации с помощью обучения с подкреплением. **Метод.** Сформулирован подход к совместному обучению агентов и векторных представлений графа. В рамках подхода предложен алгоритм *QSDNE*, использующий агентов *DQN* и векторные представления *SDNE*. **Основные результаты.** Выполнен сравнительный анализ разработанного алгоритма с алгоритмами мультиагентной маршрутизации без совместного обучения. На основании результатов работы алгоритма *QSDNE* сделан вывод о его эффективности для оптимизации времени доставки и энергопотребления в конвейерных системах. Алгоритм позволил сократить среднее время доставки на 6 % по сравнению с лучшим аналогом. **Практическая значимость.** Разработанный подход может быть использован для решения задач маршрутизации со сложными функциями оценки пути и динамически меняющимися топологиями графов, а предложенный алгоритм — для управления конвейерными лентами в аэропортах и в цеховых производствах.

Ключевые слова

мультиагентное обучение, обучение с подкреплением, адаптивная маршрутизация, конвейерные ленты, графовое представление

Благодарности

Исследование выполнено за счет гранта Российского научного фонда (проект № 20-19-00700).

Ссылка для цитирования: Рыбкин К.Е., Фильченков А.А., Азаров А.А., Забашта А.С., Шалыто А.А. Совместное обучение агентов и векторных представлений графов в задаче управления конвейерными лентами // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1187–1196. doi: 10.17586/2226-1494-2022-22-6-1187-1196

Joint learning of agents and graph embeddings in a conveyor belt control problem

Konstantin E. Rybkin¹, Andrey A. Filchenkov²✉, Artur A. Azarov³,
Alexey S. Zabashta⁴, Anatoly A. Shalyto⁵

^{1,2,3,4,5} ITMO University, Saint Petersburg, 197101, Russian Federation

³ North-West Institute of Management — branch of the Russian Presidential Academy of National Economy and Public Administration, Saint Petersburg, 199178, Russian Federation

¹ matveich1903@gmail.com, <https://orcid.org/0000-0003-0856-4276>

² afilchenkov@itmo.ru✉, <https://orcid.org/0000-0002-1133-8432>

³ artur-azarov@yandex.ru, <https://orcid.org/0000-0003-3240-597X>

⁴ azabashta@itmo.ru, <https://orcid.org/0000-0002-7494-4309>

⁵ shalyto@mail.ifmo.ru, <https://orcid.org/0000-0002-2723-2077>

Abstract

We focus on the problem of routing a conveyor belts system based on a multi-agent approach. Most of these airport baggage belt conveyor systems use routing algorithms based on manual simulation of conveyor behavior. This approach does not scale well, and new research in machine learning proposes to solve the routing problem using reinforcement learning. To solve this problem, we propose an approach to joint learning of agents and vector representations of a graph. Within this approach, we develop a *QSDNE* algorithm, which uses *DQN* agents and *SDNE* embeddings. A comparative analysis was carried out with multi-agent routing algorithms without joint learning. The results of the *QSDNE* algorithm showed its effectiveness in optimizing the delivery time and energy consumption in conveyor systems as it helped to reduce mean delivery time by 6 %. The proposed approach can be used to solve routing problems with complex path estimation functions and dynamically changing graph topologies, and the proposed algorithm can be used to control conveyor belts at airports and in manufacturing workshops.

Keywords

multi-agent learning, reinforcement learning, adaptive routing, conveyor belt, graph representation

Acknowledgements

The study was supported by a grant from the Russian Science Foundation (project no. 20-19-00700).

For citation: Rybkin K.E., Filchenkov A.A., Azarov A.A., Zabashta A.S., Shalyto A.A. Joint learning of agents and graph embeddings in a conveyor belt control problem. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1187–1196 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1187-1196

Введение

В 2019 году мировой пассажиропоток воздушного транспорта составил 4,5 млрд пассажиров в год [1]. Несмотря на существенное снижение этого показателя, вызванного эпидемией коронавируса, индустрия гражданской авиации быстро развивается. Например, число перевезенных пассажиров в марте 2022 года на внутренних рейсах может даже превысить аналогичный показатель в марте 2019 года [2]. 46 % всех потерь багажа происходит из-за ошибок при его транспортировке, что приводит к многомиллионным убыткам авиакомпаний ежегодно [3].

На сегодняшний день в большинстве таких конвейерных систем багажных лент в аэропортах используются алгоритмы маршрутизации, основанные на ручном моделировании поведения конвейеров на отдельных участках сети [4]. Такой подход плохо масштабируем, так как для каждой конкретной системы приходится заново реализовывать алгоритм. Новые исследования в области машинного обучения предлагают решать задачу маршрутизации с помощью обучения с подкреплением. Основным подходом является мультиагентный [5–7], при котором агенты, помещенные в вершины или транспортируемые объекты, принимают решение о том, куда далее следует перемещать объект. Такие агенты заранее обучаются стратегиям принятия решений в рамках обучения с подкреплением.

Принципиальный вопрос дизайна подобного рода решений — на каком уровне хранится информация о графе. Мультиагентность предполагает, что агенты име-

ют лишь ограниченное представление о графе. В алгоритме *DQN-LE* [8] предложено использовать векторные представления вершин графа, и, таким образом, агенты обучаются принимать решения на основе выработанного векторного представления. Следовательно, агенты работают с ограниченным представлением графа, которое строится вне зависимости от решаемой задачи.

В настоящей работе рассмотрен вопрос о том, можно ли совместить в рамках машинного обучения и мультиагентного подхода обучение стратегиям принятия решений о транспортировке и представлений графов.

Цель работы — разработка алгоритма обучения векторного представления вершин совместно с мультиагентной системой для решения задачи маршрутизации на конвейерных лентах.

Задача маршрутизации и ее решения

Задача транспортировки грузов. Задачу транспортировки грузов (задачу маршрутизации) можно описать следующим образом: необходимо составить такой оптимальный маршрут грузов из точек их отправления в точки приема таким образом, чтобы затраченное на транспортировку время и/или стоимость были минимальными.

Для решения данной задачи предложено множество подходов, например, симплекс-метод [9] или метод последовательного заполнения матрицы плана (метод северо-западного угла) [10]. Однако наиболее эффективные методы решения на сегодняшний день **основаны на теории графов**.

Существующие алгоритмы маршрутизации, использующие графовое представление, можно разделить на два класса: для статической и динамической маршрутизации. Так как рассматриваемая авторами настоящей работы система конвейеров динамическая (конвейеры могут выходить из строя и останавливаться), то основное внимание уделим **динамическим алгоритмам на графах**.

Можно выделить три основных категории таких алгоритмов. Дистанционно-векторные алгоритмы (*distance-vector*) [11], например алгоритмы в протоколах IGRP (Interior Gateway Routing Protocol) [12] и RIP (Routing Information Protocol) [13], которые основаны на алгоритме Беллмана–Форда [14]. Каждый узел графа хранит вектор расстояний до каждого соседнего узла. Для каждого приходящего объекта высчитывается оптимальный сосед, на основе хранимого в узле вектора. В случае изменения локальной топологии (например, в случае обрыва связи с одним из узлов), информация об этом событии передается всем соседям, а те, в свою очередь, могут распространить ее далее. Данный метод хорошо работает только в небольших сетях, так как при масштабировании сильно увеличивается нагрузка на сеть из-за интенсивного обмена информацией между узлами.

Алгоритмы состояния связи (*link-state*) [15], такие как OLSR (Optimized Link State Routing Protocol) [16], используют алгоритм Дейкстры [17]. В каждом узле сети хранится матрица смежности всего графа со стоимостью путей между узлами. При изменении топологии сети (обрыве или восстановлении связи) информация о событии передается сразу всем узлам. На практике, как правило, используются статические или редко обновляемые таблицы.

Третье семейство образовано алгоритмами на основе обучения с подкреплением [18], первым из которых являлся *Q-routing* [3], где каждая вершина в графе независимый агент, а средой выступает граф. Каждый агент хранит в том или ином виде информацию об оптимальном пути $Q(d, n_i)$, где i — индекс агента, d — вершина назначения, n_i — индекс соседа агента i . Значением таких функций Q будет являться оценочная стоимость перемещения объекта до пункта назначения через вершину n_i .

Алгоритм *Q-routing* записывает информацию о значениях Q в таблицы, однако современные алгоритмы [7, 8, 19] аппроксимируют функцию Q при помощи глубоких нейронных сетей. Используем данный подход, так как он показывает лучшие результаты для оптимизации конвейерных лент [8].

Алгоритм DQN-LE. Для корректной работы алгоритма нейросетевого обучения необходимо преобразовать номера вершин графа в некоторое векторное представление, в противном случае сеть может сделать ложный вывод об упорядоченности вершин в соответствии с нумерацией. В работе [8] для алгоритма *DQN-LE* использованы алгоритмы обучения векторных представлений вершин — наиболее эффективно в экспериментах показал себя алгоритм *Laplacian Eigenmaps* (*LE*) [20]. Несмотря на простоту реализации, метод *LE* хорошо работает в системах с меняющимися топологиями.

Приведем описание реализации алгоритма *DQN-LE* [8], использованного в настоящей работе. На вход в алгоритм подадим граф и запустим алгоритм предобучения агентов и графовых представлений. Предобучение важно для качественной работы алгоритмов обучения с подкреплением, иначе сходимость к точке оптимума во время исполнения алгоритма может потребовать очень большого числа итераций. Предобучение произойдет на фиксированном статичном графе. Все агенты предобучаются параллельно с векторными представлениями, причем веса нейронных сетей всех агентов и значения векторных представлений *LE*, обновляются одинаково. Полученные предобученные модели сохраним в файлы для последующего использования на стадии исполнения.

Далее запустим основную часть алгоритма — симуляцию работы конвейеров по заданному сценарию. Есть несколько типов возможных событий: поступление n новых грузов из точек входа с заданным временным интервалом; поломка одного или нескольких конвейеров и их восстановление. Принцип работы агента в основной части опишем следующим образом.

1. В маршрутизатор, к которому привязан агент, приходит груз.
2. Номера вершин, соответствующие текущему агенту, пункту назначения груза и соседям текущего агента, преобразуем в векторное представление с помощью алгоритма *LE*, получая на выходе векторы v_p , d и n_i соответственно.
3. Полученные векторы формируют матрицу размера $3k \times m$, где k — размерность векторных представлений; m — число соседей вершины. Сформированная матрица подается на вход нейронной сети, на выходе которой получим вектор размера $m \times 1$ с оценками наилучших соседей для перенаправления груза. С помощью функции *Softmax* [21] выберем наилучшего соседа.
4. Получившееся взаимодействие агента со средой (тройка: состояние–действие–награда) занесем во внутреннюю память агента и выберем некоторое случайное подмножество произошедших событий для его дообучения.
5. Выполним обновление весов нейронной сети агента, в качестве награды передадим предсказанные значения вероятности перехода в соседнюю вершину. Память агента имеет некоторый фиксированный объем, таким образом дообучение проведем только на недавно произошедших событиях.

Одновременно с работой агента происходит дообучение связанного с ним векторного представления. При изменении топологии графа собственные векторы матрицы *LE* необходимо пересчитать заново с помощью алгоритма *DQN-LE*. Таким образом, дообучение агента и векторного представления на этапе симуляции никак не взаимосвязано.

Алгоритм QSNDE

Совместное обучение агентов и векторного представления. В алгоритме *DQN-LE* векторное представление графа обучается отдельно и независимо от обу-

чения агентов, осуществляющих маршрутизацию по этому графу. Выбранный алгоритм векторного представления LE подбирает отображение из множества вершин в векторное пространство на основе теоретико-информационных принципов без влияния задачи, которая решается агентами. Агенты вынуждены работать с таким представлением.

Сформулируем подход совместного обучения агентов и векторного представления, в котором нейронные сети, осуществляющие векторное представление графов, должны обучаться по ошибкам агентов, выполняющих маршрутизацию на основе используемых ими представлений. Такой подход может учесть в векторном представлении сложную информацию из недекомпозируемых функций стоимости.

Более формально, пусть Γ — множество графов; Embeddings — семейство отображений E из $v \in V$ множества вершин графа G , $G \in \Gamma$ в некоторое векторное пространство, элементом этого семейства является алгоритм LE . Введем также параметрическое семейство $\{Q_\theta\}_{\theta \in \Theta}$ аппроксимаций стратегий агентов в мультиагентном алгоритме маршрутизации MULTIAGENT, причем каждый элемент семейства Q_θ^i принимает на вход векторное представление соответствующей агенту вершины i .

Для графа $G \in \Gamma$ и заданной меры ошибки алгоритма маршрутизации $\mathcal{L}$ задача обучения агентов маршрутизации имеет вид

$$\arg \min_{\{\theta_i\}, \theta_i \in \Theta} \mathcal{L}(\text{MULTIAGENT}((Q_{\theta_i}^i(E(v_i)))_{i=1 \dots |V|})),$$

и E выбирается независимо от $\mathcal{L}$.

Выделим параметрическое семейство отображений $\{E_\psi\}_{\psi \in \Psi}$. Задача совместного обучения агентов и векторного представления ставится следующим образом:

$$\arg \min_{\{\theta_i\}, \psi, \theta_i \in \Theta, \psi \in \Psi} \mathcal{L}(\text{MULTIAGENT}((Q_{\theta_i}^i(E_\psi(v_i)))_{i=1 \dots |V|})).$$

Поскольку Q принимает результат работы E в качестве аргумента, решение такой задачи может быть найдено за счет выбора модели (глубокой нейронной сети) для E и разработки алгоритма совместного обучения Q и E .

Выбор модели для обучаемого векторного представления. Выделим три основные группы алгоритмов представлений вершин в динамических графах [22]:

- 1) алгоритмы матричной факторизации: LE [20], $HOPE$ [23];
- 2) алгоритмы на основе случайного обхода вершин (*random walk*): $node2vec$ [24];
- 3) алгоритмы глубокого обучения: на основе автокодировщиков, *Structural Deep Network Embedding (SDNE)* [25]; с использованием рекуррентных слоев, *dyngraph2vec* [26] и *DGNN* [27]; на основе графовых нейронных сетей [28].

Несмотря на применимость всех перечисленных алгоритмов к решению задачи векторного представления вершин в конвейерной системе, именно алгоритм $SDNE$ взят за основу для реализации совместного обучения векторных представлений и агентов. Выбор основан

на неспособности других алгоритмов эффективно обмениваться информацией о соседях с агентом на этапе дообучения без существенной перестройки архитектуры исходного алгоритма. В свою очередь, в основе алгоритма $SDNE$ лежит нейронная сеть, что позволит ее дообучить на информации от агента с помощью проталкивания градиента. Отметим, что несмотря на то, что в основе методов с $LSTM$ также лежат нейронные сети, их применение к задаче маршрутизации имеет мало практического смысла, поскольку изменения топологии (поломка/восстановление конвейерной ленты) происходит не столь часто. Использование графовых нейронных сетей затратно по времени и не подходит для адаптивной маршрутизации. Приведем описание алгоритма $SDNE$.

Будем называть вершины близкими друг другу относительно метрики близости первого порядка, если они смежные. Соответственно, близостью второго порядка для пары вершин является наличие общих вершин в каждом из двух множеств их соседей больше некоторого заданного значения. Алгоритм $SDNE$ преобразует вершины графа в некоторое векторное пространство, сохраняя близость вершин относительно метрик близости первого и второго порядков. Обучение нахождению близости первого порядка является задачей обучения с учителем, поскольку имеется информация о ребрах между вершинами. С другой стороны, для обучения близости второго порядка правильных меток нет, и модель должна уметь сама находить закономерности. Таким образом, получим полууправляемую модель обучения алгоритма.

Архитектура нейронной сети представляет автокодировщик с одним или несколькими внутренними слоями (рис. 1).

Задача автокодировщика — минимизировать следующую функцию ошибки:

$$\begin{aligned} \mathcal{L}_{\text{mix}} &= \mathcal{L}_{2nd} + \alpha \mathcal{L}_{1st} + \nu \mathcal{L}_{\text{reg}} = \\ &= \sum_{i=1}^n \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2^2 \odot b_i + \alpha \sum_{i,j=1}^n s_{ij} \|E(\mathbf{x}_i) - E(\mathbf{x}_j)\|_2^2 + \\ &\quad + \nu \sum_{k=1}^K \|\hat{\mathbf{x}}_i - \mathbf{x}_i\|_2^2, \end{aligned}$$

где $\mathbf{x}_i$ — вектор смежности вершины i ; $b_i = 1$, если $s_{i,j} = 0$, иначе $b_i = \beta > 1$; $s \in \mathbf{G}_{n,n}$ — элемент матрицы смежности $\mathbf{G}_{n,n}$; $E(\mathbf{x}_i)$ — векторное представление для i -ой вершины; $\mathcal{L}_{\text{reg}}$ — функция L2-регуляризации, используемой для предотвращения переобучения; α и ν — коэффициенты функций регуляризации.

Алгоритм совместного обучения сети агента и сети векторного представления графа. Зафиксируем веса декодировщика сети и дообучим кодировщик на текущей топологии графа, подменив градиент на выходном слое кодировщика на градиент, полученный от агента, на основе которого при помощи обратного спуска обновим веса. Таким образом, кодировщик будет дообучен на актуальном состоянии своего агента. После распространения градиента обнулим память векторного представления до следующего вызова процедуры.

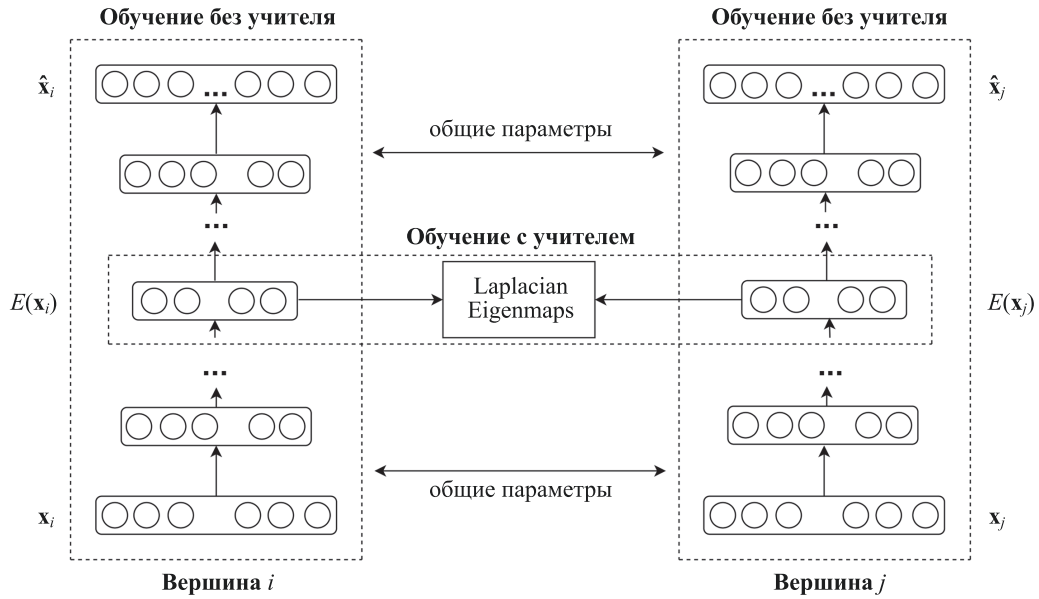


Рис. 1. Архитектура SDNE [25]

Fig. 1. SDNE architecture [25]

Опишем формально, как происходит распространение градиента от сети агента Q_θ^i к сети векторного представления E_ψ . На шаге дообучения агента Q^i вычислим значение функции потерь $\mathcal{L}(Q_\theta^i)$. С помощью метода обратного распространения ошибки получим значение градиента $\nabla_\theta = \partial \mathcal{L}(Q_\theta^i) / \partial \theta$. Опуская шаги пересчета параметров во внутренних слоях, выпишем итоговый вектор градиентов:

$$\begin{bmatrix} \frac{\partial \text{Loss}_i}{\partial E_\psi(\mathbf{x}_i)} \\ \frac{\partial \text{Loss}_i}{\partial E_\psi(\mathbf{x}^d)} \\ \frac{\partial \text{Loss}_i}{\partial E_\psi(\mathbf{x}_j)} \end{bmatrix},$$

$$\text{Loss}_i = \mathcal{L}(Q_\theta^i(E_\psi(\mathbf{x}_i), E_\psi(\mathbf{x}^d), E_\psi(\mathbf{x}_j))),$$

где $\mathbf{x}_i$ — вектор смежности для текущей вершины i ; $\mathbf{x}^d$ — вектор смежности для вершины точки назначения; $\mathbf{x}_j$ — вектор для j -го соседа текущей вершины. Обновим соответствующие строки матрицы градиентов:

$$\nabla_{\mathbf{x}_i}[j] \leftarrow \nabla_{\mathbf{x}_i}[j] + \frac{\partial \text{Loss}_i}{\partial E_\psi(\mathbf{x}_i)}.$$

Изменим параметры кодировщика, используя обратное распространение градиента ошибки и вычислим градиент:

$$\nabla_\psi = \frac{\partial E_\psi(\mathbf{x}_i)}{\partial \psi}.$$

В данном случае вместо значения функции потерь на выходной слой кодировщика подадим вектор $1 \times k$, где k — размерность векторного представления. В ре-

зультате после обратного распространения градиента сквозь скрытые слои получим:

$$\nabla E_\psi(\mathbf{v}_i) \leftarrow \frac{\partial E_\psi(\mathbf{v}_i)}{\partial \mathbf{v}_i}.$$

Рассчитаем градиент по $\mathbf{G}_{n,n}$:

$$\frac{\partial \mathcal{L}(Q_\theta^i)}{\partial \mathbf{G}_{n,n}} \otimes \frac{\partial \text{Loss}_i}{\partial E_\psi(\mathbf{x}_i)} \times \frac{\partial E_\psi(\mathbf{x}_i)}{\partial \mathbf{x}_i},$$

где $\otimes$ — произведение Кронекера.

Архитектура и описание алгоритма QSDNE. На вход алгоритма подадим матрицу смежности графа $\mathbf{G}$, $\mathbf{G}_{n,n}$. Архитектура QSDNE содержит архитектуру агента, совпадающую с алгоритмом DQN-LE, и архитектуру векторного представления графа, совпадающую с SDNE (рис. 2).

Этап предобучения QSDNE полностью совпадает с предобучением в DQN-LE.

На этапе дообучения используем алгоритм совместного обучения весов сетей агента и векторного представления графа. Для этого введем внутреннюю память векторных представлений $\mathbf{M}_{n,k}$, где n — число вершин графа. На шаге дообучения агента Q запишем в память получившийся частичный градиент по входным переменным, который имеет вид $(i_{grad}, d_{grad}, n_{i,grad})$. Добавим градиенты к соответствующим строкам $\mathbf{M}_{n,k}$. При совершении заданного числа итераций дообучения агента, запускается процедура обратного распространения градиента по сети SDNE.

При изменении топологии графа векторные представления дообучаются так же, как и в DQN-LE, однако скорость дообучения декодировщика будет выше, чем у кодировщика. Это необходимо по той причине, что веса кодировщика дополнительно обновлялись вместе с обновлением весов сети агента, поэтому для лучшей схожести декодировщик должен дообучаться быстрее.

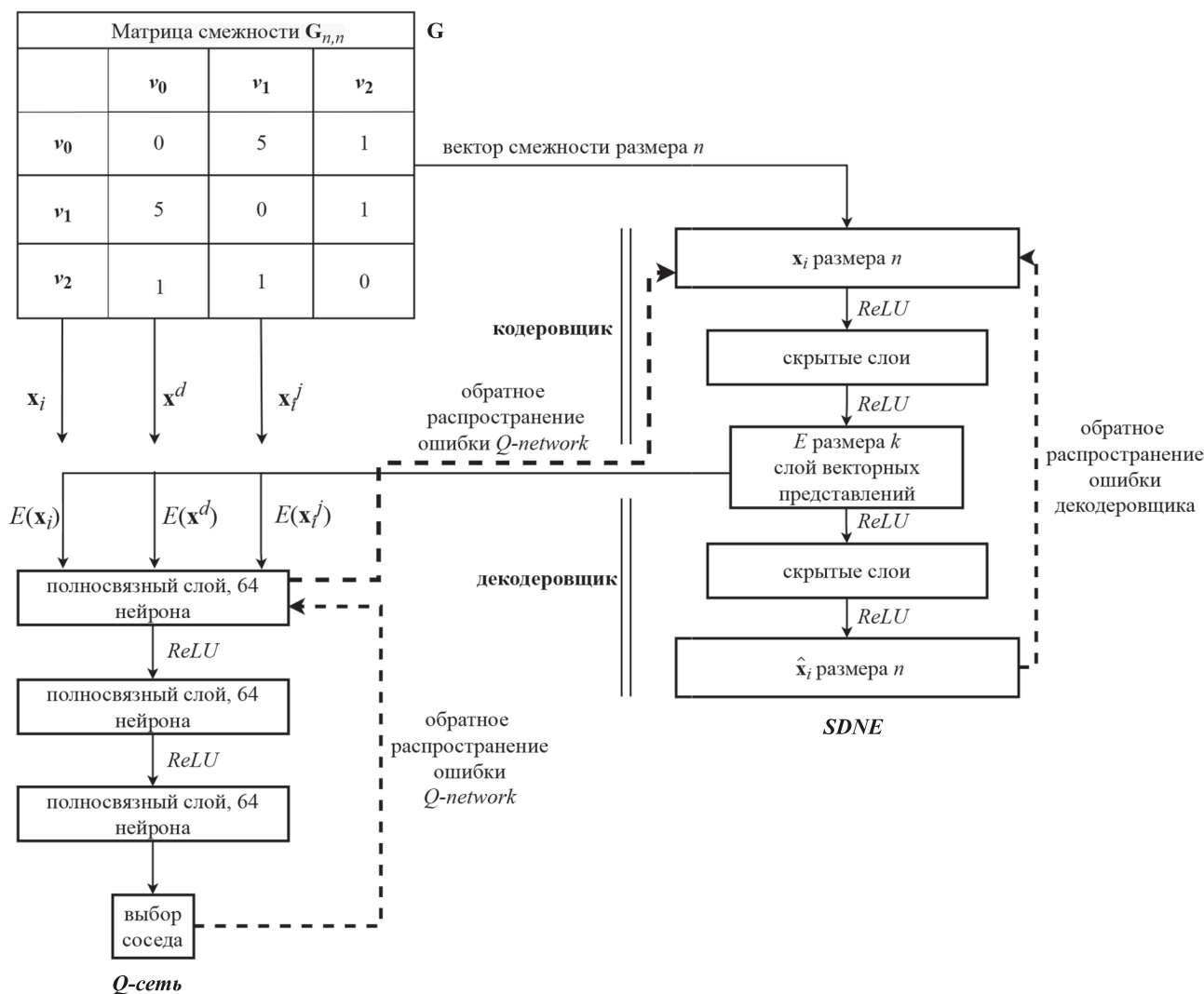


Рис. 2. Архитектура QSDNE.

G — граф, подаваемый на вход; SDNE — архитектура одного блока автокодировщика (рис. 1); Q-сеть — архитектура агента DQN

Fig. 2. QSDNE architecture

G is input graph. SDNE is the architecture of one SDNE autoencoder block (Fig. 1).

Q-network is the architecture of an agent's Q-network (DQN)

Экспериментальное исследование алгоритма QSDNE

Симуляционная среда для тестирования. Тестирование алгоритма выполним в симуляционной среде DQN-routing, что мотивировано рядом причин. Во-первых, среда позволяет симулировать работу конвейеров максимально приближенно к реальным условиям. Во-вторых, среда дает возможность выполнить сравнение в удобном виде с существующими подходами к обучению векторных представлений.

Процесс симуляции работы конвейерных систем предназначен для работы с дискретно-событийными моделями — системой конвейеров для транспортировки грузов. Приоритетная задача в работе конвейеров — предотвращение столкновений багажа. Принимающие маршрутизаторы исходят из первоочередной задачи оптимизации числа столкновений (в идеале равному нулю).

Симуляционная среда позволяет рассчитывать основные метрики эффективности доставки по конвейерной ленте: время доставки груза, измеряемое в секундах, и затрачиваемую на это электроэнергию, измеряемую в условных единицах симуляционной системы. Последняя метрика является недекомпозируемой, поскольку учитывает затраты на перемещение конвейерной ленты, на которой могут одновременно находиться сразу несколько перемещаемых объектов. Также система учитывает число столкновений перемещаемых объектов.

Поместим агенты в маршрутизаторы, которые стоят в точках сочленений конвейеров и отвечают за изменение пути груза до пункта назначения. Места маршрутизаторов в конвейере зададим расстоянием от начала конвейера до точки сочленения. Выполним преобразование конвейерной системы в граф, при этом маршрутизаторы и пункты отправки/приема грузов являются

вершинами графа, а части конвейеров — его ребрами. Пример такого графа представлен на рис. 3.

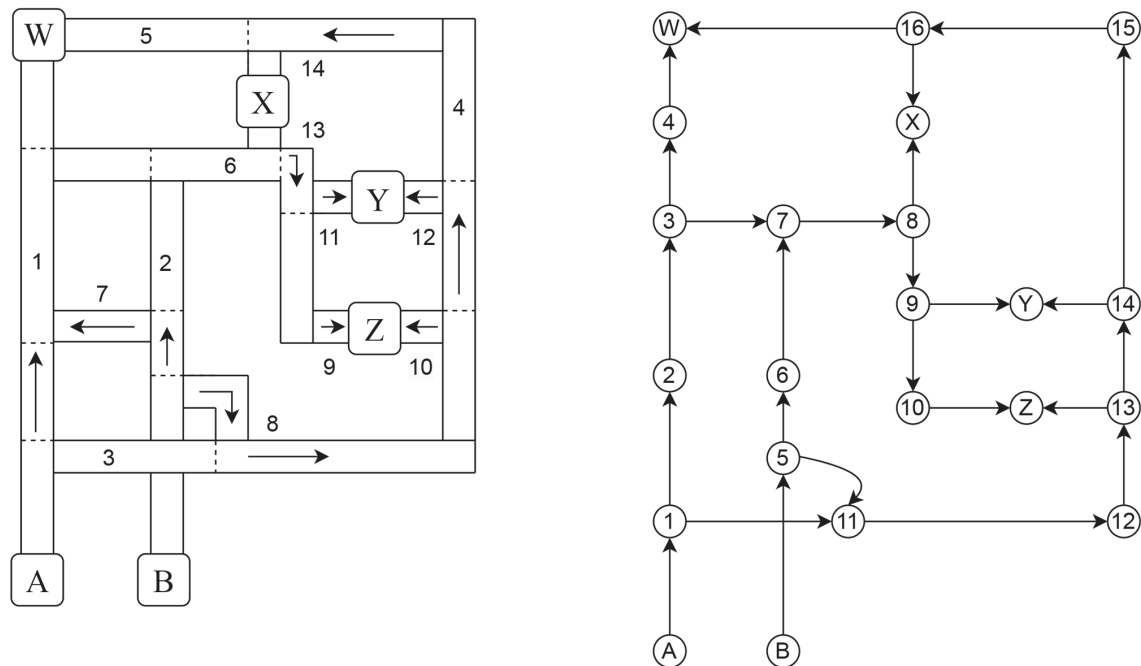


Рис. 3. Пример конвейерной системы и ее представление в виде графа [7].

А и В — начало конвейера; W, X, Y, Z — точки сочленения

Fig. 3. An example of a conveyor system and its representation as a graph [7], A and B and conveyor entry points, W, X, Y, Z are diverters

Технические особенности реализации. Представим разработанный модуль генерации данных для предобучения модели *QSDNE*, который собирает данные для предобучения на заданном графе. Генератор находит индексы ненулевых элементов в матрице смежности и выдает на их основе список пар соответствующих им вершин. Полученные вершины передаются на вход алгоритма *QSDNE*.

Перед началом сравнения произведем отбор гиперпараметров модели *QSDNE* с использованием библиотеки *hyperopt* [29]. Отбор выполним на стадии предобучения модели без использования реальных сценариев поведения. Число итераций метода *QSDNE* примем равным 500, при этом оптимальный коэффициент обучения — 0,02. Кроме гиперпараметров, подберем оптимальное число слоев в автокодировщике. Автокодировщик имеет по одному скрытому слою в 15 нейронов каждый в кодировщике и декодировщике. Размер выходного векторного представления равен 10. На этапе предобучения кодировщик и декодировщик обучены с одинаковой скоростью.

Сравнение *QSDNE* с другими алгоритмами. Так как применение векторных представлений неразрывно связано с оптимизацией решения задачи маршрутизации, то и сравнение алгоритмов векторного представления вершин произведем в контексте оптимизации метрик работы симуляционной системы. Выполним сравнение по следующим метрикам: среднее время доставки груза, среднее энергопотребление всей системы и число столкновений сумок. Для уменьшения

дисперсии произведем по 20 запусков на каждом сценарии.

Проведем сравнение четырех алгоритмов: *LE* [20], *node2vec* [24], *SDNE* [25] и *QSDNE*. Отметим, что векторные представления дообучались только в случае алгоритма *QSDNE*, а алгоритм *LE* соответствует *DQN-LE*.

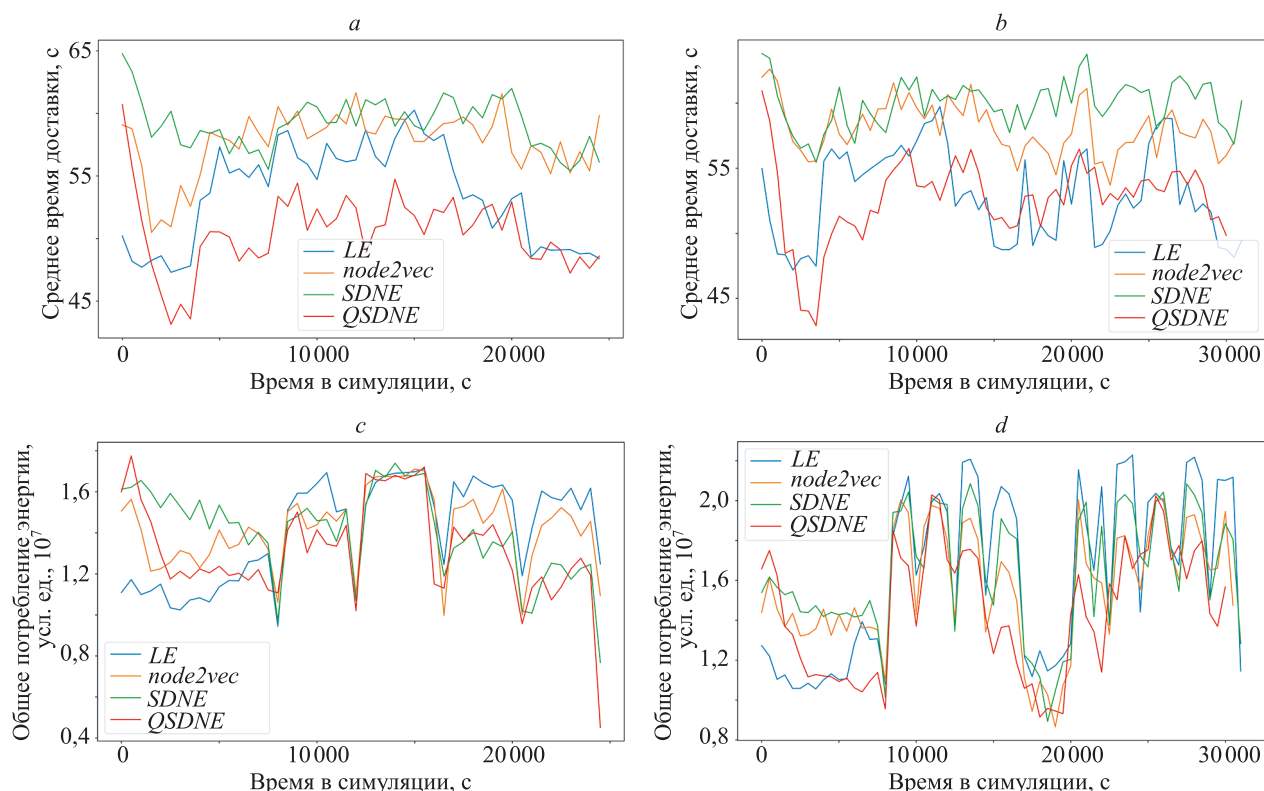
Рассмотрим два сценария записи системы. Первый сценарий (базовый) предполагает две полочки разных конвейеров и последующее их восстановление. Во втором сценарии (усложненном) увеличена интенсивность изменения топологии графа: при меньшем количестве загружаемого на ленты багажа, события поломки и восстановления конвейеров стали происходить чаще. В этом сценарии тестировалось, как эффективно алгоритм *QSDNE* способен приспосабливаться к изменениям окружения по сравнению с остальными подходами. Результаты работы алгоритмов представлены в таблице и на рис. 4. Полученные результаты по среднему времени доставки статистически значимы при параметре *p-value*, который равен 10^{-6} .

На рис. 4, а видно, что алгоритм *QSDNE* большую часть работы имеет лучшие результаты, чем реализация *LE*. Из рис. 4, б следует, что график *QSDNE* ведет себя ровнее, а значит, быстрее происходит адаптация к изменениям в графе. Несмотря на то, что результаты не являются статистически значимыми, они позволяют говорить, что *QSDNE* не менее эффективен на данном сценарии, чем метод *LE*. Тем не менее, с учетом вышеизложенного, его использование может быть более востребовано на реальных примерах.

Таблица. Сравнение времени доставки в базовом и усложненном сценариях

Table. Comparison of delivery time in basic and advanced scenarios

Алгоритм векторного представления графа	Среднее время доставки, с		Минимальное время доставки, с		Максимальное время доставки, с		Число столкновений
	базовый	усложненный	базовый	усложненный	базовый	усложненный	
<i>LE</i>	53,49	53,04	40,56	40,46	74,64	74,73	0
<i>node2vec</i>	57,67	58,06	41,19	41,24	78,89	79,77	0
<i>SDNE</i>	59,19	59,99	41,22	41,21	79,68	80,88	0
<i>QSDNE</i>	50,49	52,63	39,43	41,29	75,75	76,97	0

Рис. 4. Сравнение *QSDNE* с другими алгоритмами на базовом и усложненном сценариях по параметрам: время доставки (a, b) и энергия (c, d)Fig. 4. Comparison of *QSDNE* with other algorithms in the basic and advanced scenarios in terms of parameters: delivery time (a, b) and energy (c, d)

Отдельно рассмотрим графики метрики — энергопотребления (рис. 4, c, d). Можно заметить, что *QSDNE* эффективнее алгоритма *LE* и стандартной реализации *SDNE*.

Несмотря на то, что полученные результаты по усложненному сценарию для *QSDNE* не показали существенную разницу в энергопотреблении с алгоритмом *node2vec*, нельзя сказать, что *node2vec* является лучшим для использования. Это доказано результатами статистического теста (по энергопотреблению *node2vec* не обходит статистически *QSDNE*) и с помощью сравнения алгоритмов по среднему времени доставки. При этом алгоритм *node2vec* показал один из худших результатов, тогда как *QSDNE* уменьшил среднее время доставки на 6 % по сравнению с лучшим результатом, который был достигнут алгоритмом *LE*.

Заключение

В работе предложен подход совместного обучения агентов и векторных представлений графа в парадигме мультиагентной маршрутизации на основе глубокого обучения с подкреплением, и предложен алгоритм *QSDNE*, реализующий этот подход для решения задачи маршрутизации в системе конвейерных лент.

Результаты работы алгоритма *QSDNE* позволили говорить о его эффективности для оптимизации времени доставки и энергопотребления в конвейерных системах.

Разработанный подход может быть использован для решения задач маршрутизации со сложными функциями оценки пути и динамически меняющимися топологиями графов. Предложенный алгоритм может применяться для управления конвейерными лентами в аэропортах и в цеховых производствах.

Литература

1. The World of Air Transport in 2019. ICAO Annual Report, 2019 [Электронный ресурс]. URL: <https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx> (дата обращения: 01.10.2022).
2. COVID-19 Air Travel Recovery, OAG, 2022 [Электронный ресурс]. URL: <https://www.oag.com/coronavirus-airline-schedules-data>. (дата обращения: 01.10.2022).
3. 2019 SITA Baggage IT Insights report, SITA, 2019 [Электронный ресурс]. URL: <https://www.sita.aero/resources/surveys-reports/baggage-it-insights-2019>. (дата обращения: 06.09.2022)
4. Sørensen R.A., Nielsen M., Karstoft H. Routing in congested baggage handling systems using deep reinforcement learning // Integrated Computer-Aided Engineering. 2020. V. 27. N 2. P. 139–152. <https://doi.org/10.3233/ICA-190613>
5. Kang Y., Wang X., Lan Z. Q-adaptive: A multi-agent reinforcement learning based routing on dragonfly network // Proc. of the 30th International Symposium on High-Performance Parallel and Distributed Computing (HPDC). 2021. P. 189–200. <https://doi.org/10.1145/3431379.3460650>
6. Choi S., Yeung D.Y. Predictive Q-routing: A memory-based reinforcement learning approach to adaptive traffic control // Advances in Neural Information Processing Systems. 1995. V. 8. P. 945–951.
7. Mukhutdinov D., Filchenkov A., Shalyto A., Vyatkin V. Multi-agent deep learning for simultaneous optimization for time and energy in distributed routing system // Future Generation Computer Systems. 2019. V. 94. P. 587–600. <https://doi.org/10.1016/j.future.2018.12.037>
8. Мухудинов Д. Децентрализованный алгоритм управления конвейерной системой с использованием методов мультиагентного обучения с подкреплением: магистерская диссертация. СПб.: Университет ИТМО, 2019, 92 с. [Электронный ресурс]. URL: http://is.ifmo.ru/diploma-theses/2019/2_5458464771026191430.pdf (дата обращения: 01.10.2022).
9. Dantzig G.B. Origins of the simplex method // A history of scientific computing. 1990. P. 141–151. <https://doi.org/10.1145/87252.88081>
10. Кузнецов А.В., Холод Н.И., Костевич Л.С. Руководство к решению задач по математическому программированию. Минск: Высшая школа, 1978. 256 с.
11. Vutukury S., Garcia-Luna-Aceves J.J. MDVA: A distance-vector multipath routing protocol // Proc. of the 20th Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM). 2001. V. 1. P. 557–564. <https://doi.org/10.1109/INFOCOM.2001.916780>
12. Albrightson R., Garcia-Luna-Aceves J.J., Boyle J. EIGRP — A fast routing protocol based on distance vectors. 1994.
13. Verma A., Bhardwaj N. A review on routing information protocol (RIP) and open shortest path first (OSPF) routing protocol // International Journal of Future Generation Communication and Networking. 2016. V. 9. N 4. P. 161–170. <https://doi.org/10.14257/ijfgcn.2016.9.4.13>
14. Bang-Jensen J., Gutin G.Z. Digraphs: Theory, Algorithms and Applications. Springer Science & Business Media, 2009. 795 p.
15. Clausen T., Jacquet P. Optimized link state routing protocol (OLSR). 2003. N RFC3626. <https://doi.org/10.17487/RFC3626>
16. Jacquet P., Muhlethaler P., Clausen T., Laouiti A., Qayyum A., Viennot L. Optimized link state routing protocol for ad hoc networks // Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century. 2001. P. 62–68. <https://doi.org/10.1109/INMIC.2001.995315>
17. Noto M., Sato H. A method for the shortest path search by extended Dijkstra algorithm // Smc 2000 conference proceedings. 2000 IEEE International Conference on Systems, Man and Cybernetics. 2000. V. 3. P. 2316–2320. <https://doi.org/10.1109/icsmc.2000.886462>
18. Mammeri Z. Reinforcement learning based routing in networks: Review and classification of approaches // IEEE Access. 2019. V. 7. P. 55916–55950. <https://doi.org/10.1109/ACCESS.2019.2913776>
19. Yao Z., Wang Y., Qiu X. DQN-based energy-efficient routing algorithm in software-defined data centers // International Journal of Distributed Sensor Networks. 2020. V. 16. N 6. <https://doi.org/10.1177/1550147720935775>
20. Belkin M., Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering // Advances in Neural Information Processing Systems. 2002.
21. Gao B., Pavel L. On the properties of the softmax function with application in game theory and reinforcement learning // arXiv. 2017. arXiv:1704.00805. <https://doi.org/10.48550/arXiv.1704.00805>

References

1. *The World of Air Transport in 2019. ICAO Annual Report, 2019.* Available at: <https://www.icao.int/annual-report-2019/Pages/the-world-of-air-transport-in-2019.aspx> (accessed: 01.10.2022).
2. *COVID-19 Air Travel Recovery, OAG, 2022.* Available at: <https://www.oag.com/coronavirus-airline-schedules-data>. (accessed: 01.10.2022).
3. *2019 SITA Baggage IT Insights report, SITA, 2019.* Available at: <https://www.sita.aero/resources/surveys-reports/baggage-it-insights-2019>. (accessed: 06.09.2022)
4. Sørensen R.A., Nielsen M., Karstoft H. Routing in congested baggage handling systems using deep reinforcement learning. *Integrated Computer-Aided Engineering*, 2020, vol. 27, no. 2, pp. 139–152. <https://doi.org/10.3233/ICA-190613>
5. Kang Y., Wang X., Lan Z. Q-adaptive: A multi-agent reinforcement learning based routing on dragonfly network. *Proc. of the 30th International Symposium on High-Performance Parallel and Distributed Computing (HPDC)*, 2021, pp. 189–200. <https://doi.org/10.1145/3431379.3460650>
6. Choi S., Yeung D.Y. Predictive Q-routing: A memory-based reinforcement learning approach to adaptive traffic control. *Advances in Neural Information Processing Systems*, 1995, vol. 8, pp. 945–951.
7. Mukhutdinov D., Filchenkov A., Shalyto A., Vyatkin V. Multi-agent deep learning for simultaneous optimization for time and energy in distributed routing system. *Future Generation Computer Systems*, 2019, vol. 94, pp. 587–600. <https://doi.org/10.1016/j.future.2018.12.037>
8. Mukhudinov D. *Decentralized conveyor system control algorithm using multi-agent reinforcement learning methods.* MSc Dissertation. St. Petersburg, ITMO University, 2019, 92 p. Available at: http://is.ifmo.ru/diploma-theses/2019/2_5458464771026191430.pdf (accessed: 01.10.2022). (in Russian)
9. Dantzig G.B. Origins of the simplex method. *A history of scientific computing*, 1990, pp. 141–151. <https://doi.org/10.1145/87252.88081>
10. Kuznetsov A.V., Kholod N.I., Kostevich L.S. *Guide to Solving Mathematical Programming Problems*. Minsk, Vyshejschaja shkola Publ., 1978, 256 p. (in Russian)
11. Vutukury S., Garcia-Luna-Aceves J.J. MDVA: A distance-vector multipath routing protocol. *Proc. of the 20th Annual Joint Conference on the IEEE Computer and Communications Societies (INFOCOM)*, 2001, vol. 1, pp. 557–564. <https://doi.org/10.1109/INFOCOM.2001.916780>
12. Albrightson R., Garcia-Luna-Aceves J.J., Boyle J. *EIGRP — A fast routing protocol based on distance vectors*, 1994.
13. Verma A., Bhardwaj N. A review on routing information protocol (RIP) and open shortest path first (OSPF) routing protocol. *International Journal of Future Generation Communication and Networking*, 2016, vol. 9, no. 4, pp. 161–170. <https://doi.org/10.14257/ijfgcn.2016.9.4.13>
14. Bang-Jensen J., Gutin G.Z. *Digraphs: Theory, Algorithms and Applications*. Springer Science & Business Media, 2009, 795 p.
15. Clausen T., Jacquet P. *Optimized link state routing protocol (OLSR)*. 2003, no. RFC3626. <https://doi.org/10.17487/RFC3626>
16. Jacquet P., Muhlethaler P., Clausen T., Laouiti A., Qayyum A., Viennot L. Optimized link state routing protocol for ad hoc networks. *Proceedings. IEEE International Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century*, 2001, pp. 62–68. <https://doi.org/10.1109/INMIC.2001.995315>
17. Noto M., Sato H. A method for the shortest path search by extended Dijkstra algorithm. *Smc 2000 conference proceedings. 2000 IEEE International Conference on Systems, Man and Cybernetics*, 2000, vol. 3, pp. 2316–2320. <https://doi.org/10.1109/icsmc.2000.886462>
18. Mammeri Z. Reinforcement learning based routing in networks: Review and classification of approaches. *IEEE Access*, 2019, vol. 7, pp. 55916–55950. <https://doi.org/10.1109/ACCESS.2019.2913776>
19. Yao Z., Wang Y., Qiu X. DQN-based energy-efficient routing algorithm in software-defined data centers. *International Journal of Distributed Sensor Networks*, 2020, vol. 16, no. 6. <https://doi.org/10.1177/1550147720935775>
20. Belkin M., Niyogi P. Laplacian eigenmaps and spectral techniques for embedding and clustering. *Advances in Neural Information Processing Systems*, 2002.
21. Gao B., Pavel L. On the properties of the softmax function with application in game theory and reinforcement learning. *arXiv*, 2017, arXiv:1704.00805. <https://doi.org/10.48550/arXiv.1704.00805>

22. Barros C.D., Mendonça M.R., Vieira A.B., Ziviani A. A survey on embedding dynamic graphs // *ACM Computing Surveys*, 2021. V. 55. N 1. P. 1–37. <https://doi.org/10.1145/3483595>
23. Ou M., Cui P., Pei J., Zhang Z., Zhu W. Asymmetric transitivity preserving graph embedding // *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. P. 1105–1114. <https://doi.org/10.1145/2939672.2939751>
24. Grover A., Leskovec J. Node2vec: Scalable feature learning for networks // *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. P. 855–864. <https://doi.org/10.1145/2939672.2939754>
25. Wang D., Cui P., Zhu W. Structural deep network embedding // *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. P. 1225–1234. <https://doi.org/10.1145/2939672.2939753>
26. Goyal P., Chhetri S.R., Canedo A. dyngraph2vec: Capturing network dynamics using dynamic graph representation learning // *Knowledge-Based Systems*, 2020. V. 187. P. 104816. <https://doi.org/10.1016/j.knosys.2019.06.024>
27. Ma Y., Guo Z., Ren Z., Tang J., Yin D. Streaming graph neural networks // *Proc. of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020. P. 719–728. <https://doi.org/10.1145/3397271.3401092>
28. Kool W., Van Hoof H., Welling M. Attention, learn to solve routing problems! // *Proc. of the 7th International Conference on Learning Representations (ICLR)*, 2019.
29. Komer B., Bergstra J., Eliasmith C. Hyperopt-sklearn // *Automated Machine Learning*, 2019. P. 97–111. https://doi.org/10.1007/978-3-030-05318-5_5
22. Barros C.D., Mendonça M.R., Vieira A.B., Ziviani A. A survey on embedding dynamic graphs. *ACM Computing Surveys*, 2021, vol. 55, no. 1, pp. 1–37. <https://doi.org/10.1145/3483595>
23. Ou M., Cui P., Pei J., Zhang Z., Zhu W. Asymmetric transitivity preserving graph embedding. *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1105–1114. <https://doi.org/10.1145/2939672.2939751>
24. Grover A., Leskovec J. Node2vec: Scalable feature learning for networks. *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 855–864. <https://doi.org/10.1145/2939672.2939754>
25. Wang D., Cui P., Zhu W. Structural deep network embedding. *Proc. of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 1225–1234. <https://doi.org/10.1145/2939672.2939753>
26. Goyal P., Chhetri S.R., Canedo A. dyngraph2vec: Capturing network dynamics using dynamic graph representation learning. *Knowledge-Based Systems*, 2020, vol. 187, pp. 104816. <https://doi.org/10.1016/j.knosys.2019.06.024>
27. Ma Y., Guo Z., Ren Z., Tang J., Yin D. Streaming graph neural networks. *Proc. of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020, pp. 719–728. <https://doi.org/10.1145/3397271.3401092>
28. Kool W., Van Hoof H., Welling M. Attention, learn to solve routing problems! *Proc. of the 7th International Conference on Learning Representations (ICLR)*, 2019.
29. Komer B., Bergstra J., Eliasmith C. Hyperopt-sklearn. *Automated Machine Learning*, 2019, pp. 97–111. https://doi.org/10.1007/978-3-030-05318-5_5

Авторы

Рыбкин Константин Евгеньевич — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0003-0856-4276>, matveich1903@gmail.com

Фильченков Андрей Александрович — кандидат физико-математических наук, инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-1133-8432>, afilchenkov@itmo.ru

Азаров Артур Александрович — кандидат технических наук, научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация; заместитель директора, Северо-Западный институт управления — филиал РАНХиГС, Санкт-Петербург, 199178, Российская Федерация, <https://orcid.org/0000-0003-3240-597X>, artur-azarov@yandex.ru

Забашта Алексей Сергеевич — кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-7494-4309>, azabashta@itmo.ru

Шалыто Анатолий Абрамович — доктор технических наук, профессор, главный научный сотрудник, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-2723-2077>, shalyto@mail.ifmo.ru

Authors

Konstantin E. Rybkin — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0003-0856-4276>, matveich1903@gmail.com

Andrey A. Filchenkov — PhD (Physics & Mathematics), Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-1133-8432>, afilchenkov@itmo.ru

Artur A. Azarov — PhD, Scientific Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation; Deputy Director, North-West Institute of Management - branch of the Russian Presidential Academy of National Economy and Public Administration, Saint Petersburg, 199178, Russian Federation, <https://orcid.org/0000-0003-3240-597X>, artur-azarov@yandex.ru

Alexey S. Zabashta — PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-7494-4309>, azabashta@itmo.ru

Anatoly A. Shalyto — D. Sc., Professor, Chief Researcher, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-2723-2077>, shalyto@mail.ifmo.ru

Статья поступила в редакцию 30.10.2022

Одобрена после рецензирования 03.11.2022

Принята к печати 29.11.2022

Received 30.10.2022

Approved after reviewing 03.11.2022

Accepted 29.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

МАТЕМАТИЧЕСКОЕ И КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ MODELING AND SIMULATION

doi: 10.17586/2226-1494-2022-22-6-1197-1204

УДК 535.3, 519.85

Моделирование процессов переноса излучения в газожидкостных пенах

Анна Андреевна Исаева¹✉, Елена Андреевна Исаева², Алексей Викторович Пантюков³

^{1,2,3} Саратовский государственный технический университет имени Гагарина Ю.А., Саратов, 410054, Российская Федерация

¹ isanna.1987@mail.ru✉, <https://orcid.org/0000-0002-2925-2733>

² 27isaevaeva@mail.ru, <https://orcid.org/0000-0003-0407-312X>

³ alex.pantukoff@yandex.ru, <https://orcid.org/0000-0002-6422-501X>

Аннотация

Предмет исследования. Представлены результаты численного моделирования процесса переноса излучения в газожидкостных пенах. Для исследования переноса зондирующего излучения в газожидкостных пеноподобных средах на различных этапах старения предложен метод численного моделирования Монте-Карло. Для учета переотражения на границах раздела двух фаз применена методика на основе формул Френеля. **Метод.** В качестве модельной среды рассмотрены ячеистые структуры Кельвина, имитирующие газовые пузырьки в жидкой матрице в процессе старения (coarsening), представляющем переход пеноподобной среды от «влажной» к «сухой». Подобный переход для изолированной системы обусловлен медленным стеканием жидкости по стенкам газовых ячеек и каналам Плато–Гиббса вследствие влияния гравитационной силы. В процессе эволюции происходит уменьшение объемной доли жидкой фазы и увеличение среднего размера газовых ячеек пеноподобной среды. Трехмерная ячеистая структура на разных этапах эволюции представлена как система плотноупакованных упорядоченных сфер или тетраэдров различных геометрических размеров в зависимости от длительности процесса старения. **Основные результаты.** На основе численного моделирования методом Монте-Карло и формул Френеля получены значения коэффициентов пропускания и обратного рассеяния зондирующего излучения для газожидкостной пены. Полученные результаты хорошо коррелируют со спектрометрическими измерениями вспененной жидкости производства Gillette (Gillette shaving cream). Приведен анализ влияния фактора анизотропии среды на среднюю оптическую длину зондирующего излучения. **Практическая значимость.** Разработка теоретических подходов к диагностике двухфазных пеноподобных материалов позволит оптимизировать синтез вспененных материалов с заданными реологическими и структурными свойствами и повысить эффективность оценки функциональных параметров подобных материалов.

Ключевые слова

газожидкостные пены, ячейка Кельвина, коэффициент пропускания, метод Монте-Карло, формулы Френеля

Благодарности

Работа выполнена при финансовой поддержке гранта Российского научного фонда (проект № 21-79-00051).

Ссылка для цитирования: Исаева А.А., Исаева Е.А., Пантюков А.В. Моделирование процессов переноса излучения в газожидкостных пенах // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1197–1204. doi: 10.17586/2226-1494-2022-22-6-1197-1204

Simulation of radiative transfer in gas-liquid foams

Anna A. Isaeva¹✉, Elena A. Isaeva², Aleksey V. Pantukov³

^{1,2,3} Yuri Gagarin State Technical University of Saratov, Saratov, 410054, Russian Federation

¹ isanna.1987@mail.ru✉, <https://orcid.org/0000-0002-2925-2733>

² 27isaevaeva@mail.ru, <https://orcid.org/0000-0003-0407-312X>

³ alex.pantukoff@yandex.ru, <https://orcid.org/0000-0002-6422-501X>

Abstract

The results of the radiation light transfer in the gas-liquid foams are presented. To study the probing light transfer in gas-liquid foam-like media at different stages of aging, a Monte Carlo numerical simulation method is proposed. To

take into account the re-reflections at the phase interfaces, the approach based on the Fresnel formula is applied. Kelvin cells structures are considered as a model medium, imitating gas bubbles in a liquid matrix during aging (coarsening), which represents the transition of a foam-like medium from “wet” to “dry”. Such transition is caused by the slow liquid flow along the gas cells walls and Plateau-Gibbs channels due to the influence of the gravitational force in an isolated system. During the evolution process, the volume fraction of the liquid phase decreases and the average size of the gas cells of the foam-like medium increases. The three-dimensional cellular structure at different evolution stages is represented as a system of close-packed ordered spheres or tetracaedrons of various geometric sizes depending on the aging process duration. The Monte Carlo modeling of the radiation light transfer in scattering two-phase foam-like media, taking into account the redistribution at the interfaces of two phases by use of Fresnel formulas at different coursing stages, is presented. The transmittance and backscattering coefficients of a probing light are obtained using the Monte Carlo numerical simulations and Fresnel formulas for the gas-liquid foams. The obtained results are well correlated with the spectrometric measurements of the foamed liquid produced by Gillete (Gillete shaving cream). The influence of the anisotropy factor on the optical length of the probing light was carried out. The development of the theoretical approaches to the diagnostics of the two-phase foam-like materials makes it possible to synthesize foamed materials with established rheological and structural properties to increase and improve the efficiency of evaluating of the functional characteristics for such materials.

Keywords

gas-liquid foams, Kelvin cell, transmission coefficient, Monte Carlo simulation, Fresnel formulas

Acknowledgements

This work was supported by the Russian Science Foundation (project No. 21-79-00051).

For citation: Isaeva A.A., Isaeva E.A., Pantukov A.V. Simulation of radiative transfer in gas-liquid foams. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1197–1204 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1197-1204

Введение

Двухфазные пеноподобные структуры представляют большой интерес для науки, промышленности и биомедицины [1–3]. Скаффолд-технологии активно используются в биомедицине для таких задач, как выращивание трехмерных клеточных конструкций, обеспечения жизнедеятельности клеток в процессе формирования определенных типов живых тканей в тканевой инженерии [4–9]. Скаффолды могут быть разработаны на основе синтетических или природных материалов, в частности, биоактивной керамики, стекол, полимеров или наноккомпозитов [9–13]. Один из перспективных методов изготовления скаффолдов основан на вспенивании предварительно пластифицированных биосовместимых полимеров с помощью газа или сверхкритической жидкости [14, 15]. Структурные параметры 3D-скаффолдов определяют функциональные свойства биоматериалов. Микроскопические поры (размером более 50 мкм) влияют на форму и размер формируемой ткани, мезоскопические поры (размером 1–50 мкм) задают функции клеток в процессе их синтеза, а наноскопические поры (размером менее 1 мкм) определяют эффективность диффузии питательных веществ [16].

Разработка эффективных методов оптической диагностики для оценки структурных параметров синтезированных высокопористых матриц, например распределения пор по размерам и форме, а также степени их взаимосвязи, актуальная и важная задача.

Пены подразделяются на сухие (объемная доля жидкого компонента менее 0,1–0,15) и влажные (объемная доля жидкого компонента более 0,1–0,15). Экспериментальные исследования процессов переноса электромагнитного излучения во вспененных жидкостях и влияние таких структурных параметров, как средний размер газовых пузырей и объемная доля жидкой фазы в пене, проводились многими группами исследователей. В частности, группа ученых во гла-

ве с Д. Дурианом исследовала структурные и функциональные особенности вспененных жидкостей в процессе их старения методами спектроскопии диффузионного пропускания и диффузионно-волновой спектроскопии [17]. Динамика дренирования жидкой фазы во вспененных жидкостях описана рядом уравнений, формирующих единую модель [18–21]. Для развития экспериментальных методик исследования влияния структурных особенностей газожидкостных пен на их оптические свойства необходима адекватная математическая модель рассеяния электромагнитного излучения в подобных средах.

Описание метода и основные результаты моделирования

Газожидкостные пены можно рассмотреть, как ячеистую структуру, состоящую из газовых ячеек с перегородками из пленок жидкости, при этом форма перегородок и характер их пересечения описываются законами Плато.

Процесс старения пены (coarsening) и, соответственно, переход от «влажной» к «сухой» пене (wet to dry transition) сопровождается трансформацией формы ячеек от сферической к многогранной, увеличением среднего размера газовой ячейки и убыванием объемной доли жидкой матрицы (рис. 1). Этот процесс обусловлен частичным испарением жидкой фазы в случае неизолированной, открытой системы и ее дренажом на дно кюветы, вызванным гравитационными эффектами. Схлопывание и коалесценция газовых пузырей приводит к перестройкам внутри ячеистой структуры.

В процессе эволюции системы изменяется механизм рассеяния света. На начальных этапах старения пены рассеяние излучения происходит на газовых пузырях в жидкой матрице, а на последних этапах — на структурных элементах сухой пены, так называемых каналах Плато–Гиббса [22]. В процессе численного моделирова-

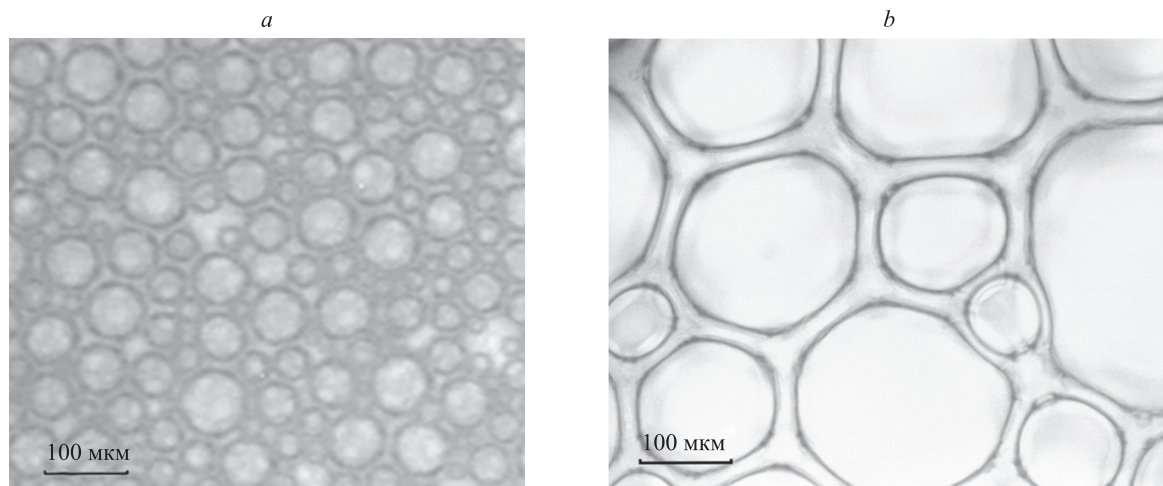


Рис. 1. Фотографии вспененной жидкости для разного времени старения пены: 500 с (а) и 60 000 с (b)
Fig. 1. Photographs of the foamed liquid at the different foam aging time: 500 s (a); 60000 s (b)

ния в качестве пеноподобной среды рассмотрим монодисперсную сотообразную модель плотноупакованных элементарных ячеек, имеющих форму усеченных октаэдров равного объема. Такая ячеистая структура впервые была упомянута в 1887 г. как структура Кельвина.

Уравнение, описывающее трансформацию ячеистой структуры пены от «влажной» к «сухой» и, соответственно, переход от сферической формы к многогранной, имеет вид

$$\left[h_{cube}(x, y, z) + \frac{1}{a^p} h_{octa}(x, y, z) + \frac{1}{b^p} h_{dodeca}(x, y, z) \right]^{1/p} = r, \quad (1)$$

где a и b — описывают геометрические свойства многогранников, вписанных в сферу (например, число сторон); r — радиус сферы, из которой формируется многогранник; p — определяет полиэдральность (степень близости многогранников к сфере).

Если $p \rightarrow 2$, то форма ячейки сферическая, если $p \rightarrow \infty$, то ячейка имеет форму многогранника [23].

Уравнение трансформации сферы в многогранник:

$$h_{cube}(x, y, z) = |x|^p + |y|^p + |z|^p. \quad (2)$$

Для $p = 2$ уравнение (2) задает сферу радиуса r , для $p \rightarrow \infty$ это уравнение описывает куб с длиной ребра $2r$.

Уравнение трансформации сферы в правильный октаэдр:

$$h_{octa}(x, y, z) = |x + y + z|^p + |-x + y + z|^p + |x - y + z|^p + |x + y - z|^p.$$

Уравнение ромбододекаэдрической формы сферы:

$$h_{dodeca}(x, y, z) = |x + y|^p + |x - y|^p + |y + z|^p + |y - z|^p + |x + z|^p + |x - z|^p.$$

Ячейка Кельвина в форме усеченного октаэдра описывается уравнением (1), где $a = b \approx 1,5355$ (рис. 2).

Выполним моделирование распространения рассеянного средой зондирующего излучения на этапе «влажной» газожидкостной пены в среде с помощью метода Монте-Карло, и на этапе «сухой» пены в рамках геометрической оптики с использованием формул Френеля. В качестве критического параметра, характеризующего переход «влажная–сухая» пена, выбран средний радиус газовых пузырей $r_{кр} = 10$ мкм.

На первых этапах эволюции модель «влажной» газожидкостной пены представляет собой систему газовых пузырей в жидкости. Среда имеет следующие параметры: бесконечную ширину, толщину 3000 мкм, коэффициент рассеяния $\mu_s = 0,177$ мкм⁻¹, малый коэффициент поглощения $\mu_a = 10^{-11}$ мкм⁻¹.

Из падающего однонаправленного пакета фотонов выбрана и отмечена одна парциальная компонента или один фотон. Фотон, распространяясь в среде, испытывает акты рассеяния, при этом траектория движения фотона задана направляющими косинусами и оптической длиной пути фотона (расстоянием, проходимым фотоном до последующего акта рассеяния).

На первом шаге генерируем направляющие косинусы, которые определяются на основе угла рассеяния ϑ в виде:

$$\cos \vartheta = \frac{1}{2g} \left(1 + g^2 - \left(\frac{1 - g^2}{1 - g + 2g\chi} \right)^2 \right),$$

где χ — случайная величина, равномерно распределенная на интервале $[0, 1]$; g — параметр анизотропии.

На следующем шаге случайным образом зададим оптическую длину пути фотона, с помощью набора равномерно распределенных случайных чисел.

Информация о длине пути фотона и направляющих косинусах сохраним в матрицу и запишем в файл. Далее сгенерируем новое значение угла и длины пробега фотона. Повторяем процедуру до тех пор, пока фотон не достигнет заданной глубины внутри среды или не вылетит из среды в направлении, обратном падающему излучению. Выберем следующий фотон из набора и повторим всю процедуру.

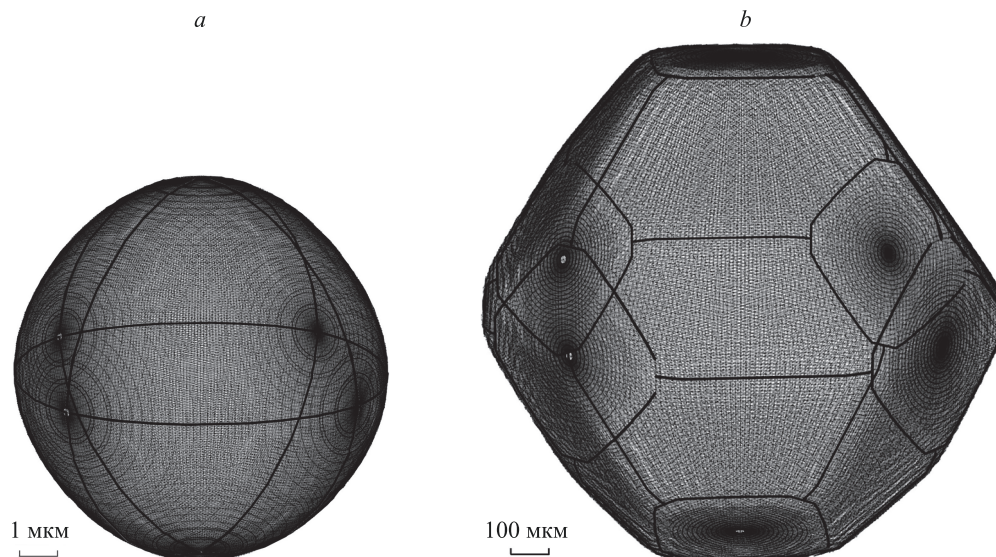


Рис. 2. Сферическая и многогранная формы ячеек Кельвина, имитирующих газовые ячейки пеноподобной среды на разных этапах ее старения: $p = 5$ (a); $p = 50$ (b)

Fig. 2. The spherical and the polyhedral shapes of Kelvin cells imitating the gas cells of a foam-like medium at the different aging stages: $p = 5$ (a); $p = 50$ (b)

На последующих этапах эволюции пены модель «сухой» газожидкостной пены представляет систему плотноупакованных ячеек Кельвина одинакового размера.

Показатели преломления газовых ячеек и перегородок из жидкой пленки равны $n_1 = 1$ и $n_2 = 1,34$ соответственно. Теоретические значения среднего размера газовых пузырей и объемной доли жидкой фазы, используемые при моделировании процесса старения пеноподобной среды, выбраны исходя из экспериментальных исследований процессов эволюции пены производства Gillette (Gillette shaving cream) [24]. Объемная доля жидкости рассчитана по толщине слоя дренированной жидкости на дне кюветы в процессе эволюции пены от «влажной» к «сухой». На начальных этапах эволюции пены доля жидкости — $0,1 \pm 0,01$. Объемная доля жидкой фазы и средний размер газовых ячеек рассматриваемой пеноподобной среды представлен в таблице. Максимальное время старения ячейки структуры в процессе моделирования составило 180 000 с.

Событие отражения или преломления на границе раздела «жидкостная пленка–газовая ячейка» опреде-

лено случайным образом при сравнении коэффициента отражения, рассчитанного по формулам Френеля, со случайной величиной равномерно распределенной на интервале $[0, 1]$. Рассчитаем коэффициент обратного рассеяния для случая отличного от нормального падения:

$$R_{bs} = \frac{1}{2} \left(\frac{\sin^2(\alpha_i - \alpha_t)}{\sin^2(\alpha_i + \alpha_t)} + \frac{\tan^2(\alpha_i - \alpha_t)}{\tan^2(\alpha_i + \alpha_t)} \right),$$

где α_i и α_t — углы падения и преломления.

В ходе моделирования пакет фотонов распространялся в ячейистой среде. Парциальные составляющие, испытавшие многократное взаимодействие с каналами Плато–Гиббса (границами раздела «жидкостная пленка–газовая ячейка») и с узлами каналов Плато–Гиббса (точки пересечения жидкостных пленок), и достигшие заданной глубины, вносили вклад в коэффициент пропускания, а составляющие, рассеянные назад — в коэффициент обратного рассеяния.

В численном эксперименте выполним моделирование рассеяния зондирующего излучения на структурных элементах газожидкостной пены с различными

Таблица. Объемная доля жидкой фазы и средний радиус газовых ячеек пеноподобной среды

Table. The volume fraction of the liquid phase and the mean radius of the foam-like medium gas bubbles

Средний радиус газовых пузырей, мкм		Объемная доля жидкой фазы	
Моделирование	Теоретические значения	Моделирование	Теоретические значения
8,7	8,0	0,10	0,10
13,5	14,0	0,10	0,09
26,2	25,0	0,09	0,09
74,8	75,0	0,08	0,08
140,0	140,0	0,06	0,06

значениями параметра p и радиусов ячеек r . Световое поле, рассеянное исследуемой средой и формирующее сигнал, представляет суперпозицию парциальных составляющих, которые распространены в среде на различную глубину и, соответственно, имеют различные значения оптических длин путей. При этом парциальные составляющие содержат информацию о значениях характеристик подвижности рассеивающих структур. Интенсивности прошедшего и рассеянного излучений содержат информацию о реорганизации газовых ячеек, происходящей в процессе старения пеноподобной среды. В ходе моделирования проведена оценка коэффициентов пропускания и обратного рассеяния излучения в зависимости от характеристик, задающих геометрию ячейистой структуры газожидкостной пены.

Коэффициенты обратного рассеяния R_{bs} и пропускания T , зондирующего пеноподобную среду излучения, имеют вид

$$R_{bs} = \frac{I_{bs}}{I_0}, T = \frac{I_{tr}}{I_0}, \quad (3)$$

где интенсивности излучения: I_{bs} — обратнорассеянного пеноподобной средой; I_{tr} — прошедшего на заданную глубину в пеноподобной среде; I_0 — зондирующего.

Зависимости коэффициентов, полученные по формуле (3) с помощью моделирования методом Монте-Карло и учетом переотражения на границах раздела двух фаз с помощью формул Френеля для различных характеристик ячейистой структуры системы, представлены на рис. 3.

Полученные численные значения коэффициентов пропускания для модельной ячейистой среды соответствуют значениям коэффициента пропускания пены производства Gillette (Gillette shaving cream), полученным экспериментально с использованием спектрометра Ocean Optics QE65000 [24]. Средний радиус газовых ячеек, используемый при моделировании, изменялся от 9 до 140 мкм, при этом полученное пропускание ячейистой структуры изменялось в пределах от 0,04 до 0,21, а коэффициент пропускания пены производства Gillette изменялся в пределах от 0,07 до 0,23. При

анализе рассеянных компонент зондирующего излучения наблюдались увеличение доли рассеянного вперед излучения и уменьшение доли обратно рассеянного излучения. Переход от «влажной» к «сухой» пене привел к трансформации формы ячейки Кельвина от сферической к многогранной и, соответственно, увеличению значения p . При этом происходит увеличение коэффициента пропускания среды от 0,04 до 0,19.

Газовые пузыри в ячейистой структуре плотно упакованы и обладают высокой корреляцией. Рассеяние в подобных структурах носит нетипичный характер, в частности при увеличении времени старения пены происходит переход от рассеяния на газовых ячейках (на ранних этапах старения) к рассеянию на структурных элементах пены (на поздних этапах старения). Такие изменения влияют на оптические параметры: транспортную длину l_{tr} (расстояние, на котором происходит изменение режима рассеяния от направленного к диффузному); длину рассеяния l (среднее расстояние в среде, проходимое электромагнитной волной от первого акта рассеяния до последующего); анизотропию рассеяния света g (средний косинус угла рассеяния электромагнитного излучения).

Выполнен анализ влияния фактора анизотропии рассеяния на среднее значение оптической длины пути зондирующего излучения (рис. 4).

В процессе эволюции пеноподобной структуры увеличивается объемная доля газовых пузырей и изменяется плотность рассеивающих центров, и, как следствие, кратность рассеяния. В результате происходит увеличение оптической длины пути рассеянного излучения (рис. 4).

Разработанная модель среды позволила описать изменяющийся механизм рассеяния, заключающийся в переходе от рассеяния света на системе газовых пузырей в жидкости к рассеянию на структурных элементах «сухой» пены — каналах Плато–Гиббса [22].

Моделирование переноса излучения в газожидкостной пене методом Монте-Карло на малых временах эволюции системы и переход к геометрической оптике на больших временах эволюции демонстрирует способность выбранного подхода к описанию физических процессов, протекающих в газожидкостных пенах.

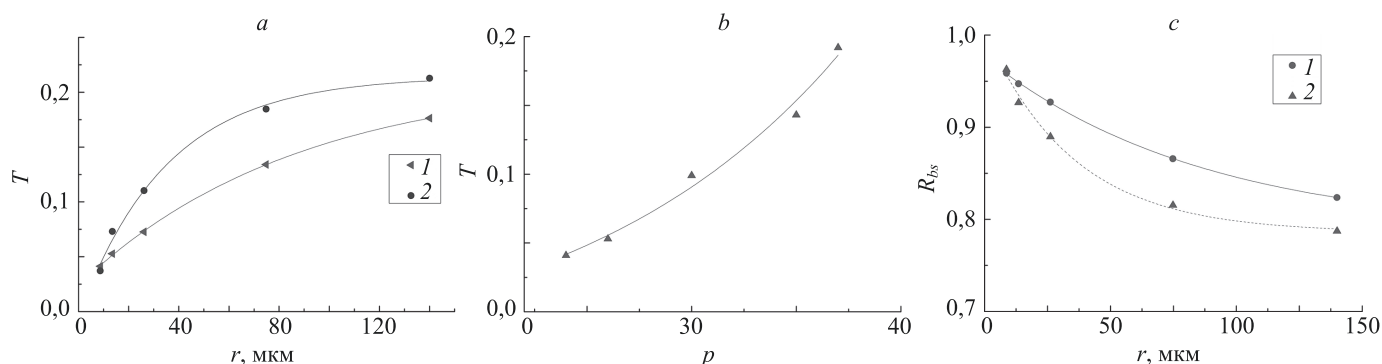


Рис. 3. Коэффициенты пропускания (T) (a, b) и обратного рассеяния (R_{bs}) (c) излучения для различных характеристик ячейистой структуры системы при значениях параметра: $p = 24$ (кривая 1); $p = 50$ (кривая 2)

Fig. 3. The transmission (T) (a, b) and backscattering coefficients (R_{bs}) (c) of the radiation for various characteristics of the parameter value: $p = 24$ (curve 1); $p = 50$ (curve 2)

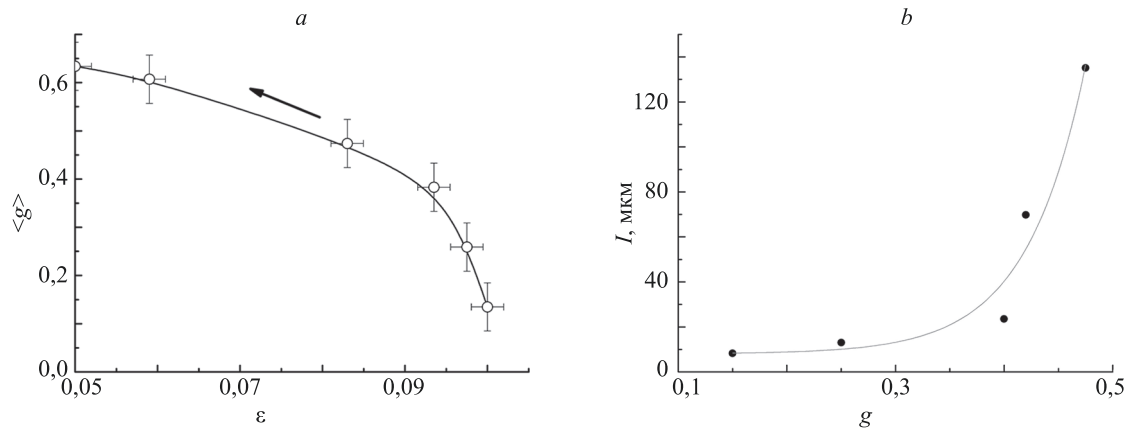


Рис. 4. Восстановленные зависимости по экспериментальным данным: параметра анизотропии рассеяния от объемной доли жидкой фазы в процессе эволюции пеноподобной среды [24] (а); среднего значения оптической длины пути зондирующего излучения от анизотропии рассеяния среды (б).

Стрелка — эволюция системы по мере ее старения

Fig. 4. The experimentally reconstructed dependence: of the scattering anisotropy parameter on the liquid phase volume fraction [24] (a); of the average optical path length of the probing radiation on the medium anisotropy parameter (b).

The arrow shows the evolution of aging system

Закключение

В работе рассмотрен подход с использованием формул Френеля и численного метода Монте-Карло для исследования переноса излучения в ячеистых структурах, имитирующих пеноподобные среды на различных этапах старения.

Увеличение коэффициента пропускания зондирующего излучения, наблюдаемое при переходе от «влажной» к «сухой» пене, обусловлено не только увеличением геометрических размеров газовых ячеек Кельвина в жидкой фазе, но и трансформацией ячеек от сферической к многогранной форме.

Коэффициенты пропускания излучения, полученные в результате моделирования и изменяющиеся в

пределах 0,04–0,21, демонстрируют хорошую сходимость с экспериментальными спектрометрическими измерениями коэффициента пропускания вспененной жидкости 0,07–0,23. Таким образом, для интерпретации экспериментальных исследований газожидкостных пен на ранних этапах старения может быть использована трехмерная ячеистая структура Кельвина со сферической формой, на поздних этапах старения — система ячеек Кельвина многогранной формы.

Полученные результаты могут быть использованы для развития когерентно-оптических методов морфофункциональной диагностики подобных систем в режиме реального времени.

Литература

1. Li N., Fu Y., Lu Q., Xiao C. Microstructure and performance of a porous polymer membrane with a copper nano-layer using vapor-induced phase separation combined with magnetron sputtering // *Polymers*. 2017. V. 9. N 10. P. 524–527. <https://doi.org/10.3390/polym9100524>
2. Poh P.S.P., Chhaya M.P., Wunner F.M., De-Juan-Pardo E.M., Schilling A.F., Schantz J.-T., van Griensven M., Huttmacher D.W. Polylactides in additive biomanufacturing // *Advanced Drug Delivery Reviews*. 2016. V. 107. P. 228–246. <https://doi.org/10.1016/j.addr.2016.07.006>
3. Zimnyakov D.A., Yuvchenko S.A., Isaeva A.A., Isaeva E.A., Tsylin D.V. Growth/collapse kinetics of the surface bubbles in fresh constrained foams: transition to self-similar evolution // *Colloids and Surfaces A: Physicochemical and Engineering Aspects*. 2019. V. 579. P. 123693. <https://doi.org/10.1016/j.colsurfa.2019.123693>
4. Hollister S.J. Porous scaffold design for tissue engineering // *Nature Materials*. 2005. V. 4. N 7. P. 518–524. <https://doi.org/10.1038/nmat1421>
5. Садовой М.А., Ларионов П.М., Самохин А.Г., Рожнова О.М. Клеточные матрицы (скаффолды) для целей регенерации кости: современное состояние проблемы // *Хирургия позвоночника*. 2014. № 2. С. 79–86. <https://doi.org/10.14531/ss2014.2.79-86>

References

1. Li N., Fu Y., Lu Q., Xiao C. Microstructure and performance of a porous polymer membrane with a copper nano-layer using vapor-induced phase separation combined with magnetron sputtering. *Polymers*, 2017, vol. 9, no. 10, pp. 524–527. <https://doi.org/10.3390/polym9100524>
2. Poh P.S.P., Chhaya M.P., Wunner F.M., De-Juan-Pardo E.M., Schilling A.F., Schantz J.-T., van Griensven M., Huttmacher D.W. Polylactides in additive biomanufacturing. *Advanced Drug Delivery Reviews*, 2016, vol. 107, pp. 228–246. <https://doi.org/10.1016/j.addr.2016.07.006>
3. Zimnyakov D.A., Yuvchenko S.A., Isaeva A.A., Isaeva E.A., Tsylin D.V. Growth/collapse kinetics of the surface bubbles in fresh constrained foams: transition to self-similar evolution. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 2019, vol. 579, pp. 123693. <https://doi.org/10.1016/j.colsurfa.2019.123693>
4. Hollister S.J. Porous scaffold design for tissue engineering. *Nature Materials*, 2005, vol. 4, no. 7, pp. 518–524. <https://doi.org/10.1038/nmat1421>
5. Sadovoy M.A., Larionov P.M., Samokhin A.G., Rozhnova O.M. Cellular matrices (scaffolds) for bone regeneration: state of the art. *Journal of Spine Surgery*, 2014, no. 2, pp. 79–86. (in Russian). <https://doi.org/10.14531/ss2014.2.79-86>

6. Ngadiman N.H.A., Yusof N.M., Idris A., Fallahiazouard E., Kurniawan D. Novel processing technique to produce three dimensional polyvinyl alcohol/maghemite nanofiber scaffold suitable for hard tissues // *Polymers*. 2018. V. 10. N 4. P. 353. <https://doi.org/10.3390/polym10040353>
7. Seo S.J., Mahapatra C., Singh R.K., Knowles J.C., Kim H.-W. Strategies for osteochondral repair: Focus on scaffolds // *Journal of Tissue Engineering*. 2014. V. 5. P. 1–5. <https://doi.org/10.1177/2041731414541850>
8. Hokmabad V.R., Davaran S., Ramazani A., Salehi R. Design and fabrication of porous biodegradable scaffolds: a strategy for tissue engineering // *Journal of Biomaterials Science, Polymer Edition*. 2017. V. 28. N 16. P. 1797–1825. <https://doi.org/10.1080/09205063.2017.1354674>
9. Langer R., Vacanti J.P. Tissue engineering // *Science*. 1993. V. 260. N 5110. P. 920–926. <https://doi.org/10.1126/science.8493529>
10. Celikkin N., Rinoldi C., Costantini M., Trombetta M., Rainer A., Świąszkowski W. Naturally derived proteins and glycosaminoglycan scaffolds for tissue engineering applications // *Materials Science and Engineering: C*. 2017. V. 78. P. 1277–1299. <https://doi.org/10.1016/j.msec.2017.04.016>
11. Laube T., Weisser J., Berger S., Börner S., Bischoff S., Schubert H., Gajda M., Bräuer R., Schnabelrauch M. In situ foamable, degradable polyurethane as biomaterial for soft tissue repair // *Materials Science and Engineering: C*. 2017. V. 78. P. 163–174. <https://doi.org/10.1016/j.msec.2017.04.061>
12. O'Brien F.J. Biomaterials & scaffolds for tissue engineering // *Materials Today*. 2011. V. 14. N 3. P. 88–95. [https://doi.org/10.1016/S1369-7021\(11\)70058-X](https://doi.org/10.1016/S1369-7021(11)70058-X)
13. Maitz M. Applications of synthetic polymers in clinical medicine // *Biosurface and Biotribology*. 2015. V. 1. N 3. P. 161–176. <https://doi.org/10.1016/j.bsbt.2015.08.002>
14. Dehghani F., Annabi N. Engineering porous scaffolds using gas-based techniques // *Current Opinion in Biotechnology*. 2011. V. 22. N 5. P. 661–666. <https://doi.org/10.1016/j.copbio.2011.04.005>
15. Quirk R.A., France R.M., Shakesheff K.M., Howdle S.M. Supercritical fluid technologies and tissue engineering scaffolds // *Current Opinion in Solid State and Materials Science*. 2004. V. 8. N 3-4. P. 313–321. <https://doi.org/10.1016/j.cossms.2003.12.004>
16. Loh Q.L., Choong C. Three-dimensional scaffolds for tissue engineering applications: Role of porosity and pore size // *Tissue Engineering Part B: Reviews*. 2013. V. 19. N 6. P. 485–502. <https://doi.org/10.1089/ten.teb.2012.0437>
17. Durian D.J., Weitz D.A., Pine D.J. Multiple light-scattering probes of foam structure and dynamics // *Science*. 1991. V. 252. N 5006. P. 686–688. <http://doi.org/10.1126/science.252.5006.686>
18. Koehler S.A., Stone H.A., Brenner M.P., Eggers J. Dynamics of foam drainage // *Physical Review E*. 1998. V. 58. N 2. P. 2097–2106. <https://doi.org/10.1103/PhysRevE.58.2097>
19. Kraynik A.M. Foam drainage: Internal Report. 1983. 83-0844.
20. Schmiedeberg M., Miri M.F., Stark H. Photon channelling in foams // *European Physical Journal E*. 2005. V. 18. N 1. P. 123–131. <http://doi.org/10.1140/epje/i2005-10034-6>
21. Koehler S.A., Hilgenfeldt S., Stone H.A. A Generalized view of foam drainage: experiment and theory // *Langmuir*. 2000. V. 16. N 15. P. 6327–6341. <https://doi.org/10.1021/la9913147>
22. Cantat I., Cohen-Addad S., Elias F., Graner F., Höhler R., Pitois O., Rouyer F., Saint-Jalmes A. *Foams: Structure and Dynamics*. Oxford: Oxford University Press, 2013. 278 p.
23. Onaka S. Superspheres: Intermediate shapes between spheres and polyhedra // *Symmetry*. 2012. V. 4. N 3. P. 336–343. <https://doi.org/10.3390/sym4030336>
24. Зимняков Д.А., Ювченко С.А., Исаева А.А., Исаева Е.А., Ушакова О.В. Анизотропия рассеяния света вспененными жидкостями // *Оптика и спектроскопия*. 2018. Т. 125. № 5. С. 699–707. <https://doi.org/10.21883/OS.2018.11.46846.99-18>
6. Ngadiman N.H.A., Yusof N.M., Idris A., Fallahiazouard E., Kurniawan D. Novel processing technique to produce three dimensional polyvinyl alcohol/maghemite nanofiber scaffold suitable for hard tissues. *Polymers*, 2018, vol. 10, no. 4, pp. 353. <https://doi.org/10.3390/polym10040353>
7. Seo S.J., Mahapatra C., Singh R.K., Knowles J.C., Kim H.-W. Strategies for osteochondral repair: Focus on scaffolds. *Journal of Tissue Engineering*, 2014, vol. 5, pp. 1–5. <https://doi.org/10.1177/2041731414541850>
8. Hokmabad V.R., Davaran S., Ramazani A., Salehi R. Design and fabrication of porous biodegradable scaffolds: a strategy for tissue engineering. *Journal of Biomaterials Science, Polymer Edition*, 2017, vol. 28, no. 16, pp. 1797–1825. <https://doi.org/10.1080/09205063.2017.1354674>
9. Langer R., Vacanti J.P. Tissue engineering. *Science*, 1993, vol. 260, no. 5110, pp. 920–926. <https://doi.org/10.1126/science.8493529>
10. Celikkin N., Rinoldi C., Costantini M., Trombetta M., Rainer A., Świąszkowski W. Naturally derived proteins and glycosaminoglycan scaffolds for tissue engineering applications. *Materials Science and Engineering: C*, 2017, vol. 78, pp. 1277–1299. <https://doi.org/10.1016/j.msec.2017.04.016>
11. Laube T., Weisser J., Berger S., Börner S., Bischoff S., Schubert H., Gajda M., Bräuer R., Schnabelrauch M. In situ foamable, degradable polyurethane as biomaterial for soft tissue repair. *Materials Science and Engineering: C*, 2017, vol. 78, pp. 163–174. <https://doi.org/10.1016/j.msec.2017.04.061>
12. O'Brien F.J. Biomaterials & scaffolds for tissue engineering. *Materials Today*, 2011, vol. 14, no. 3, pp. 88–95. [https://doi.org/10.1016/S1369-7021\(11\)70058-X](https://doi.org/10.1016/S1369-7021(11)70058-X)
13. Maitz M. Applications of synthetic polymers in clinical medicine. *Biosurface and Biotribology*, 2015, vol. 1, no. 3, pp. 161–176. <https://doi.org/10.1016/j.bsbt.2015.08.002>
14. Dehghani F., Annabi N. Engineering porous scaffolds using gas-based techniques. *Current Opinion in Biotechnology*, 2011, vol. 22, no. 5, pp. 661–666. <https://doi.org/10.1016/j.copbio.2011.04.005>
15. Quirk R.A., France R.M., Shakesheff K.M., Howdle S.M. Supercritical fluid technologies and tissue engineering scaffolds. *Current Opinion in Solid State and Materials Science*, 2004, vol. 8, no. 3-4, pp. 313–321. <https://doi.org/10.1016/j.cossms.2003.12.004>
16. Loh Q.L., Choong C. Three-dimensional scaffolds for tissue engineering applications: Role of porosity and pore size. *Tissue Engineering Part B: Reviews*, 2013, vol. 19, no. 6, pp. 485–502. <https://doi.org/10.1089/ten.teb.2012.0437>
17. Durian D.J., Weitz D.A., Pine D.J. Multiple light-scattering probes of foam structure and dynamics. *Science*, 1991, vol. 252, no. 5006, pp. 686–688. <http://doi.org/10.1126/science.252.5006.686>
18. Koehler S.A., Stone H.A., Brenner M.P., Eggers J. Dynamics of foam drainage. *Physical Review E*, 1998, vol. 58, no. 2, pp. 2097–2106. <https://doi.org/10.1103/PhysRevE.58.2097>
19. Kraynik A.M. *Foam drainage*. Internal Report, 1983, 83-0844.
20. Schmiedeberg M., Miri M.F., Stark H. Photon channelling in foams. *European Physical Journal E*, 2005, vol. 18, no. 1, pp. 123–131. <http://doi.org/10.1140/epje/i2005-10034-6>
21. Koehler S.A., Hilgenfeldt S., Stone H.A. A Generalized view of foam drainage: experiment and theory. *Langmuir*, 2000, vol. 16, no. 15, pp. 6327–6341. <https://doi.org/10.1021/la9913147>
22. Cantat I., Cohen-Addad S., Elias F., Graner F., Höhler R., Pitois O., Rouyer F., Saint-Jalmes A. *Foams: Structure and Dynamics*. Oxford, Oxford University Press, 2013, 278 p.
23. Onaka S. Superspheres: Intermediate shapes between spheres and polyhedra. *Symmetry*, 2012, vol. 4, no. 3, pp. 336–343. <https://doi.org/10.3390/sym4030336>
24. Zimnyakov D.A., Yuvchenko S.A., Isaeva A.A., Isaeva E.A., Ushakova O.V. Anisotropy of light scattering by foamed liquids. *Optics and Spectroscopy*, 2018, vol. 125, no. 5, pp. 795–802. <https://doi.org/10.1134/S0030400X18110371>

Авторы

Исаева Анна Андреевна — кандидат физико-математических наук, доцент, доцент, Саратовский государственный технический университет имени Гагарина Ю.А., Саратов, 410054, Российская Федерация, <https://orcid.org/0000-0002-2925-2733>, isanna.1987@mail.ru

Authors

Anna A. Isaeva — PhD (Physics & Mathematics), Associate Professor, Associate Professor, Yuri Gagarin State Technical University of Saratov, Saratov, 410054, Russian Federation, <https://orcid.org/0000-0002-2925-2733>, isanna.1987@mail.ru

Исаева Елена Андреевна — кандидат физико-математических наук, доцент, доцент, Саратовский государственный технический университет имени Гагарина Ю.А., Саратов, 410054, Российская Федерация, [sc 54884743700](https://orcid.org/0000-0003-0407-312X), <https://orcid.org/0000-0003-0407-312X>, 27isaevaea@mail.ru

Пантюков Алексей Викторович — аспирант, Саратовский государственный технический университет имени Гагарина Ю.А., Саратов, 410054, Российская Федерация, <https://orcid.org/0000-0002-6422-501X>, alex.pantukoff@yandex.ru

Elena A. Isaeva — PhD (Physics & Mathematics), Associate Professor, Associate Professor, Yuri Gagarin State Technical University of Saratov, Saratov, 410054, Russian Federation, [sc 54884743700](https://orcid.org/0000-0003-0407-312X), <https://orcid.org/0000-0003-0407-312X>, 27isaevaea@mail.ru

Aleksey V. Pantukov — PhD Student, Yuri Gagarin State Technical University of Saratov, Saratov, 410054, Russian Federation, <https://orcid.org/0000-0002-6422-501X>, alex.pantukoff@yandex.ru

Статья поступила в редакцию 15.04.2022
Одобрена после рецензирования 03.10.2022
Принята к печати 16.11.2022

Received 15.04.2022
Approved after reviewing 03.10.2022
Accepted 16.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1205-1215

УДК 620.179.17

Статистическая оценка влияния величины сигнал/помеха на погрешность измерения параметров акустической эмиссии

Алексей Владимирович Федоров¹, Ельдос Алтай²✉, Ксения Андреевна Степанова³,
Дмитрий Олегович Кузиванов⁴

^{1,2,3,4} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ avfedorov@itmo.ru, <https://orcid.org/0000-0003-0612-922X>

² aeldos@inbox.ru✉, <https://orcid.org/0000-0002-3736-0291>

³ ledy.xs93@yandex.ru, <https://orcid.org/0000-0003-1811-3807>

⁴ kuzivanovdmitry@gmail.com, <https://orcid.org/0000-0001-9661-9614>

Аннотация

Предмет исследования. Современные акустико-эмиссионные диагностические системы и комплексы являются чувствительным инструментом обнаружения развивающихся дефектов при мониторинге технического состояния объектов в условиях эксплуатационных нагрузок на ранних стадиях. Существенное ограничение применения метода акустической эмиссии — сложность выделения сигналов на фоне акустических и электромагнитных помех. Влияние помех при регистрации акустической эмиссии существенно затрудняет интерпретацию ее параметров, характеризующих техническое состояние объекта контроля. Для повышения значения величины сигнал/помеха и достоверности полученных результатов контроля при количественной оценке параметров эмиссии используют методы фильтрации. Рассмотрено влияние величины сигнал/помеха на погрешность измерения параметров акустической эмиссии, выработанных при компенсации помех с помощью полиномиального метода фильтрации. **Метод.** В основу статистической модели определения влияния величины сигнал/помеха на погрешность измерения параметров акустической эмиссии положен метод машинного обучения — линейной регрессии. Зависимость погрешности измерения от величины сигнал/помеха аппроксимирована методом наименьших квадратов и визуализирована с помощью скатерограммы. **Основные результаты.** Выявлено, что при применении фильтра Баттерворта величина относительной погрешности измерений параметров акустической эмиссии не превышает 3 %, что на порядки ниже значений, полученных для фильтра Бесселя и вейвлет-фильтра на основе материнской функции Добеши 8-го порядка. Установлена высокая обратная не случайная корреляционная связь ($r > 0,9$), обусловленная снижением значений относительной погрешности измерений параметров эмиссии и повышением величины сигнал/помеха. Разработанная статистическая модель описывает влияние величины сигнал/помеха на значение относительной погрешности при оценке параметров акустической эмиссии. Функционирование предложенной модели подтверждено вычислением коэффициента детерминации и проверки его статистической значимости. **Практическая значимость.** Показано, что применение фильтра Баттерворта для компенсации помех существенно повышает информативность результатов измерений параметров акустической эмиссии. Разработанная статистическая модель может быть использована при создании новых или усовершенствовании существующих диагностических комплексов и систем обработки данных для повышения достоверности результатов акустического контроля.

Ключевые слова

статистическая обработка, повышение точности измерения, акустическая эмиссия, величина сигнал/помеха, акустический контроль, помехи, фильтр Баттерворта, вейвлет-фильтр, фильтр Бесселя

Ссылка для цитирования: Федоров А.В., Алтай Е., Степанова К.А., Кузиванов Д.О. Статистическая оценка влияния величины сигнал/помеха на погрешность измерения параметров акустической эмиссии // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1205–1215. doi: 10.17586/2226-1494-2022-22-6-1205-1215

The effect of signal-to-noise ratio value on the error in measuring acoustic emission parameters: statistical assessment

Alexey V. Fedorov¹, Yeldos Altay²✉, Ksenia A. Stepanova³, Dmitry O. Kuzivanov⁴

^{1,2,3,4} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ avfedorov@itmo.ru, <https://orcid.org/0000-0003-0612-922X>

² aeldos@inbox.ru✉, <https://orcid.org/0000-0002-3736-0291>

³ ledy.xs93@yandex.ru, <https://orcid.org/0000-0003-1811-3807>

⁴ kuzivanovdmitry@gmail.com, <https://orcid.org/0000-0001-9661-9614>

Abstract

Modern acoustic emission diagnostic systems and complexes are a sensitive tool for detecting developing defects at an early stage when monitoring the technical condition of objects under operational loads. A significant limitation of the application acoustic emission method is the difficulty in isolating signals against the background of acoustic and electromagnetic interference. The effect of interference during acoustic emission recording significantly complicates the interpretation of parameters that characterize the technical condition of the test object. To increase the value signal-to-noise ratio and increase the reliability of the results of acoustic emission testing in the quantitative assessment of parameters, filtering methods are used. The subject of this study is the study of the effect of signal-to-noise ratio value on the measurement error acoustic emission parameters formatted during noise compensation using the polynomial filtering method. The basis of the statistical model characterizing the effect of signal-to-noise ratio value on the measurement error acoustic emission parameters is based on the machine learning method — linear regression. The dependence of the measurement error on the signal-to-noise ratio value was approximated by the least-squares method and visualized using a scattergram. It was found that when using the Butterworth filter, the relative measurement error acoustic emission parameters do not exceed 3 %, which are orders of magnitude lower than the values obtained for the Bessel filter and Daubechies mother functions 8 based on wavelet filter. A high inverse non-random correlation was established ($r > 0.9$), due to a decrease in the values of the relative measurement error emission parameters and an increase in the signal-to-noise ratio value. The developed statistical model describes the effect of the signal-to-noise ratio value on the value relative error in estimating the acoustic emission parameters. The adequacy of the developed model was confirmed by calculating the coefficient of determination and checking its statistical significance. It is shown that the use of Butterworth filter to compensate for interference significantly increases the information content of the results of measurements of acoustic emission parameters. The developed statistical model can be used in the development of new or improvement of existing complexes and systems for processing acoustic emission data to improve the reliability of the results of acoustic testing.

Keywords

statistical processing, measurement accuracy improvement, acoustic emission, signal-to-noise ratio, non-destructive testing, noise, Butterworth filter, wavelet filter, Bessel filter

For citation: Fedorov A. V., Altay Y., Stepanova K. A., Kuzivanov D. O. The effect of signal-to-noise ratio value on the error in measuring acoustic emission parameters: statistical assessment. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1205–1215 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1205-1215

Введение

В акустико-эмиссионной диагностике для уменьшения погрешности измерений параметров сигналов акустической эмиссии (АЭ) важной научно-технической задачей является разработка помехоустойчивых методов обработки акусто-эмиссионной информации. Актуальность обусловлена тем, что помехоустойчивые методы во многом позволяют противостоять искажающим действиям помех, вызванных влиянием внешних факторов. К внешним факторам при проведении контроля АЭ и эксплуатации объекта следует отнести помехи, возникающие при работе силовых электрических установок, наводящие электрические импульсы в измерительный тракт, при сварке, в ходе механообработки и фрезеровании объекта контроля, от нагружающего устройства, при кавитации, при истечении жидкости и газа и др. [1–6].

Среди перечисленных помех ослабление электрической помехи импульсного происхождения — нетривиальная задача [4, 5]. Это связано с тем, что помеха имеет не только сумму детерминированных компонент, но и вариабельность в значениях амплитуды каждой компоненты. Влияние данной помехи снижает точность

измерения информационных составляющих и приводит к невозможности корректной интерпретации событий АЭ. Для уменьшения искажений и ослабления влияния помехи на составляющие сигнала АЭ применяют различные методы фильтрации, устойчивые к действиям помех. Однако при разработке методов фильтрации помех АЭ важно сохранение формы информационной составляющей, несущей диагностическую информацию результатов акустического контроля. Таким образом, среди известных методов обработки сигналов актуальность приобретает поиск методов фильтрации, обеспечивающих минимальность искажения формы информационной составляющей сигнала АЭ на выходе фильтров.

Обзор методов фильтрации сигналов АЭ

В работе [7] для уменьшения погрешности измерения и повышения значения величины сигнал/помеха выполнен анализ методов обработки сигнала АЭ, который показал, что широкое применение нашли способы вейвлет-фильтрации, эмпирической модовой декомпозиции, кластерного анализа и пространственно-временной обработки. В ходе анализа особое внимание

уделено поиску особенностей методов. Во-первых, простых и высокоточных методов по математической реализации. Во-вторых, особенностям системы обработки, например отсутствию необходимости использования схемы фильтрации высоких порядков, влияющих на форму сигнала, выбором математического базиса разложения сигнала и специальных опорных каналов, требующих отдельной формы сигнала помехи для ее фильтрации. В результате анализа методов обработки данных АЭ для высокоточной фильтрации, максимального ослабления помех при минимальных искажениях формы сигналов АЭ, установлена применимость методов полиномиальной фильтрации, представляющих собой оптимальный класс фильтров [5]. Минимальное искажение формы сигнала АЭ необходимо для дальнейшей оценки их информативных параметров при определении технического состояния объекта контроля [8]. Наличие искажений увеличивает погрешность оценки параметров АЭ и, как следствие, достоверность результатов контроля.

Применимость методов полиномиальной фильтрации обусловлена возможностью вариативной настройки характеристик фильтрации в зависимости от параметров сигналов АЭ для выделения информационных составляющих из зашумленной смеси сигнала и помехи. При синтезе фильтров данного класса принято использовать полином Баттерворта, реже полиномы Бесселя, Чебышева [9] и, в ряде случаев, полином Ньютона [10]. Это связано с тем, что фильтры, аппроксимированные полиномом Баттерворта в полосах пропускания сигнала и подавления помехи, не имеют пульсаций, частотная характеристика и коэффициент усиления фильтра не зависят от порядка (степени) полинома и являются равномерными, монотонными и стабильными при сравнении с перечисленными фильтрами. Во время обработки сигналов фильтры с полиномом Баттерворта обеспечивают выработку наименьших значений собственной ошибки, обусловленной их характеристиками. Отметим, что корни полиномов Баттерворта [5] и Ньютона [11, 12] с биномиальной структурой [10] являются общепринятым видом размещения в круговую структуру корней передаточной функции, что делает систему фильтрации устойчивой и более предпочтительной для практического применения при обработке сигналов.

В работе [13] для улучшения помехозащищенности тракта регистрации сигналов АЭ впервые предложено использовать аналоговую топологию фильтров Баттерворта на операционных усилителях. Топология реализована в виде автоматически перестраиваемых четырех двухзвенных активных фильтров нижних и верхних частот на основе операционных усилителей [13]. Такая реализация фильтров позволила улучшить помехозащищенность системы регистрации сигналов. Однако, в силу изменения параметров и структуры сигналов АЭ, а также влияющих помех в результате аналоговой фильтрации, происходит недостаточное шумоподавление. В связи этим решаются задачи цифровых методов обработки сигналов АЭ для снижения влияния помех на выходе средств аналоговой фильтрации [5]. Наличие не ослабленных помех существенно

снижает достоверность и информативность анализа сигналов АЭ. В связи с этим важной задачей является применение методов цифровой фильтрации.

В работах [5, 7, 8] выявлено, что реализация цифровых фильтров, аппроксимированных полиномами Баттерворта для фильтрации помех, является менее проработанной для решения задачи цифровой обработки сигналов АЭ. Проведен сравнительный анализ и дана оценка эффективности методов цифровой фильтрации, а именно полиномиальных фильтров верхних и нижних частот Баттерворта, Бесселя для уменьшения погрешности измерения параметров АЭ и повышения значения величины сигнал/помеха. Установлено, что сочетание методов спектрального анализа и вариативной настройки параметров фильтров позволяют максимизировать величину сигнал/помеха [5, 14] и снизить погрешность измерения амплитудных и временных параметров сигнала [8]. В работах [8, 14] показано, что среди методов полиномиальной фильтрации высокую точность обработки сигнала АЭ при минимальных искажениях информативных составляющих сигнала акустической эмиссии способны обеспечить фильтры, аппроксимированные полиномом Баттерворта. Среди фильтров Баттерворта высокие значения величины сигнал/помеха и наименьшее значение погрешности измерения параметров АЭ наблюдалось для полиномиального фильтра высокой частоты. Это позволило обосновать выбор данного метода для оценки взаимосвязи информационных составляющих сигнала и помех при многоканальной регистрации данных АЭ [5, 7].

В отличие от [5, 7, 8, 14], в настоящей работе рассмотрена статистическая оценка влияния величины сигнал/помеха на погрешность измерения диагностических параметров АЭ, выработанных на выходе методов полиномиальной цифровой фильтрации Баттерворта, которая сравнивается с зарекомендовавшими себя фильтрами Вейвлета [15] и Бесселя [8]. Для полиномиальных методов цифровой фильтрации применено понятие «фильтр вырабатывает», согласно введенной терминологии в монографии Б.Р. Левина [16] при определении возможностей фильтра для обработки сигналов АЭ. Отметим, что в работе [16] понятие «фильтр вырабатывает» было применено в контексте согласованной фильтрации сигналов. Вместо общепринятого отношения сигнал/шум применено понятие «сигнал/помеха», так как рассматриваются результаты влияния помехи на погрешность измерения параметров АЭ.

Цель исследования состоит в разработке статистической модели регрессионной зависимости величины сигнал/помеха от погрешности измерения параметров АЭ на выходе методов полиномиальной фильтрации. Модель, характеризующая влияния величины сигнал/помеха на погрешность измерения параметров АЭ разработана на основе метода машинного обучения — линейной регрессии. Выбор метода для разработки модели обоснован его работоспособностью и высокой чувствительностью к анализируемым данным [17], а также отсутствием требований по специальному обучению выборок АЭ данных, полученных с объекта контроля.

Постановка задачи

Рассмотрим, по аналогии с работой [14], задачу обработки сигнала АЭ вида:

$$x(q) = s(q) + \xi(q), \quad (1)$$

$$\xi(q) = \sum_{i=1}^m A_i \sin\left(2\pi \frac{f_i}{f_d} k + \varphi_i\right), \quad (2)$$

где q — отсчеты измерений; $s(q)$ — сигнал АЭ; $\xi(q)$ — помеха; m — число гармоник помехи; A_i — амплитуда i -ой гармоники помехи; f_i — частота идентифицированной помехи [5, 7] в диапазоне от 400 до 1200 кГц; f_d — частота дискретизации помехи 4 МГц; k — порядковый номер отсчетов помехи; φ_i — начальная фаза сигнала i -ой гармоники помехи.

При обработке сигнала АЭ необходимо решить задачу выделения составляющих $s(q)$ путем фильтрации анализируемого сигнала $x(q)$, регистрируемого в дискретные моменты времени $q = q_1, \dots, q_n$, от помехи $\xi(q)$.

Данная задача решена при допущении, что амплитуда влияющей помехи (2) известна, что является типовым допущением при решении аналогичных задач для представления результатов обработки.

Материал и метод исследования

Материал исследования. Исходным материалом исследований при разработке модели регрессионной зависимости служат значения отношения сигнал/помеха и относительной погрешности, измеренные при реализации аддитивной смеси (1) сформированного сигнала АЭ [8] и помехи (2) [14] на выходе методов цифровой фильтрации. Согласно допущению, рассмотренному в разделе «Постановка задачи», для реализации различного набора входных данных АЭ аддитивная смесь сигнала сформирована при значениях зашумленности: –10 дБ, –5 дБ, 0 дБ, 5 дБ, 10 дБ [14].

Метод исследования. Синтез цифровых фильтров с бесконечной импульсной характеристикой, а именно фильтров верхних частот Баттерворта и Бесселя, осуществлена в нормированном диапазоне при помощи фильтров-аналогов с использованием билинейного преобразования [5, 14]. Во избежание вносимого искажающего воздействия фильтрами высоких порядков на параметры сигнала АЭ и уменьшения громоздкости вычисления передаточной функции фильтров, были вы-

браны полиномы второго порядка ($n = 2$) [5, 7, 8]. Для рассматриваемых бесконечных импульсных характеристик фильтров выбрана частота среза сформированной информационной составляющей сигнала АЭ, равная 240 кГц при частоте дискретизации 4 МГц [8, 14]. На основании данных частот рассчитана передаточная функция цифровых фильтров Баттерворта и Бесселя $n = 2$. Рассчитанные передаточные функции непрерывных фильтров высоких частот и их билинейное преобразование представлены в табл. 1.

Для обеспечения минимальных искажений информационной составляющей обрабатываемого сигнала использовалась двунаправленная фильтрация. Отметим, что в отличие от однонаправленной фильтрации сигнала в виде «вход-выход», под двунаправленной фильтрацией понимается обработка сигнала «в обе стороны». При такой фильтрации сигналы АЭ обрабатываются в прямом, а затем в обратном направлениях [7]. Возникающие искажения, вносимые системой фильтрации, взаимно компенсируются. Структурная схема двунаправленной реализации фильтров высокой частоты, а также их свертка в частотной области для обработки сигналов АЭ представлены в работе [14].

Вейвлет-фильтрация сигнала выполнена на основе мягкой пороговой обработки (soft threshold) при выборе материнской вейвлет-функции Добеши 8-го порядка гладкости и 8-го уровня разложения [15]. Данные параметры вейвлет-фильтрации для оценки и сравнения результативности обработки выбраны из работы [15]. Следует отметить, что вейвлет-функция Добеши выбранного порядка и уровня разложения сигнала в [15] применены исходя из способности данного метода для выделения информативных параметров сигнала АЭ при низких значениях величины сигнал/помеха.

Согласно рекомендации [18], относительная погрешность измерений (δ) оценена для следующих параметров АЭ: среднее квадратическое отклонение (σ_U), среднеквадратическое значение сигнала ($U_{СКЗ}$), амплитуда сигнала (U), энергетический параметр MARSE (Measured Area of the Rectified Signal Envelope — измеренная площадь огибающей сигнала). Выбор перечисленных параметров АЭ обоснован тем, что они при контроле технического состояния инструмента являются чувствительными к выявлению связей со свойствами источника АЭ [18]. Статистические параметры характеризуют: отклонение отсчетов амплитуды анализируемого сигнала относительно среднего

Таблица 1. Передаточная функция рассчитанных фильтров верхних частот

Table 1. Transfer function of calculated high-pass filter

Полином фильтра, $A(s)$	Передаточные функции	
	непрерывная, $W(s)$	дискретная, $W(z)$
Баттерворта $s^2 + 1,414\Omega_C s + \Omega_C^2$	$\frac{s^2}{s^2 + 0,176s + 0,015}$	$\frac{0,915z^2 - 1,831z + 0,915}{z^2 - 1,824z + 0,832}$
Бесселя $s^2 + 1,732\Omega_C s + \Omega_C^2$	$\frac{s^2}{s^2 + 0,216s + 0,015}$	$\frac{0,899z^2 - 1,798z + 0,899}{z^2 - 1,791z + 0,805}$

Примечание. s — комплексная переменная непрерывной передаточной функции фильтра; Ω_C — частота среза фильтра; z — комплексная переменная дискретной передаточной функции фильтра.

значения (σ_U), мощности или действующего значения сигнала (U_{CK3}) и максимальной амплитуды электрического напряжения сигнала (U). Параметр MARSE характеризует вклад амплитуды колебаний электрического напряжения в общую энергию сигнала АЭ, рассчитываемую под площадью или квадратом огибающей импульса. Огибающая сигнала АЭ — специальная линия, объединяющая вершины амплитуды положительной полуволны анализируемого сигнала [18]. Расчет перечисленных амплитудных параметров осуществлен при использовании встроенных команд программной среды MATLAB. При вычислении относительной погрешности перечисленных параметров под исходной формой сигнала (эталоном) принята форма экспериментально-тестового сигнала АЭ, сформированная в

работе [8]. Схема обработки и вычисления параметров показана на рис. 1.

Статистическая модель, характеризующая влияния величины сигнал/помеха на значение относительной погрешности при оценке параметров АЭ, разработана с помощью линейной регрессии. Зависимость погрешности измерения от величины сигнал/помеха аппроксимирована методом наименьших квадратов и визуализирована при помощи скатерограммы, изображающей распределение двух переменных между величинами сигнал/помеха и относительной погрешности измерения параметров АЭ (предиктор). Выбор величины сигнал/помеха для регрессионной зависимости обоснована тем, что данный критерий наиболее полно характеризует устойчивость системы обработки сигнала

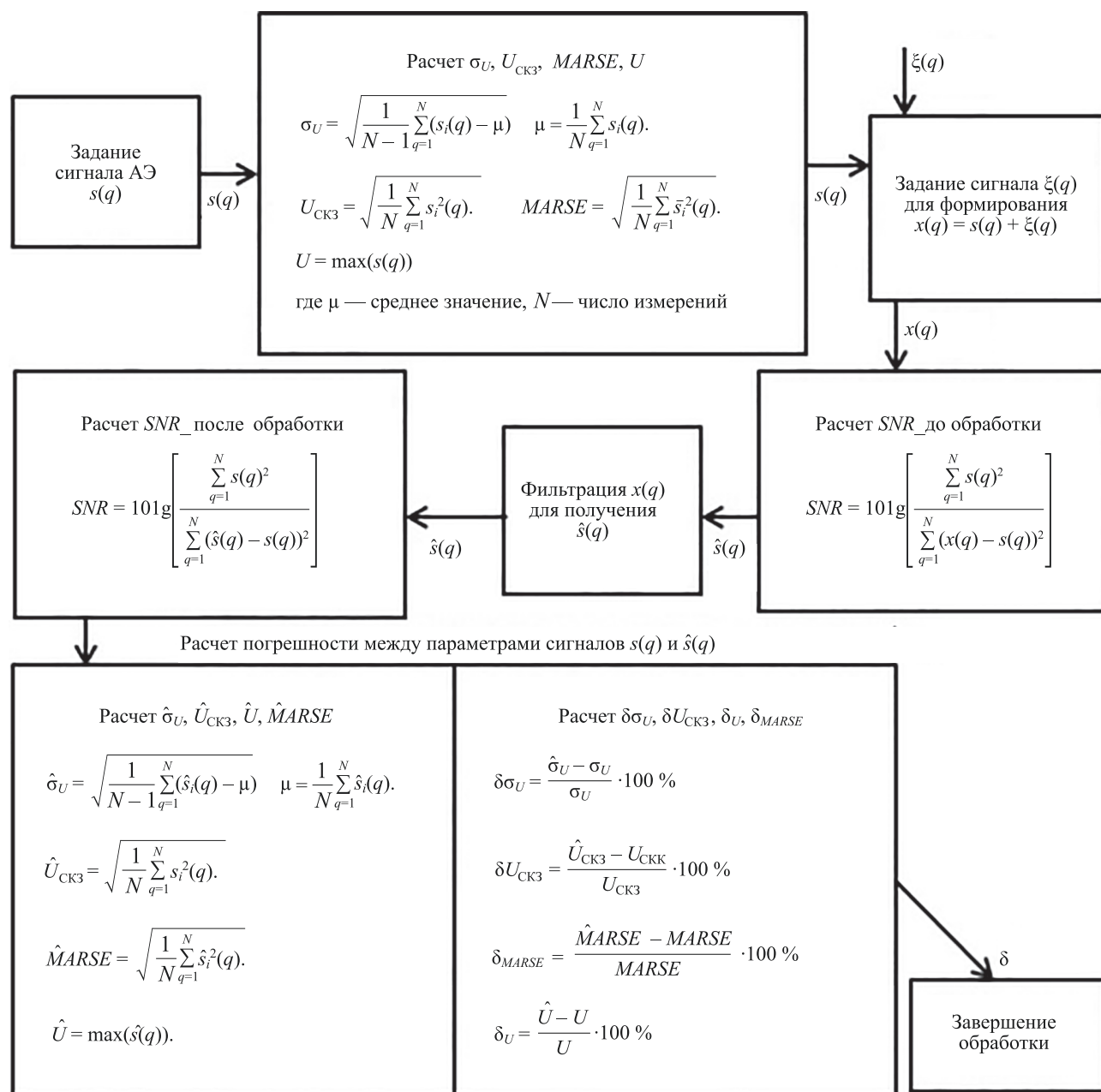


Рис. 1. Схема обработки сигнала акустической эмиссии

Fig. 1. Acoustic emission signal processing scheme

к искажающей помехе для уменьшения погрешности измерения параметров АЭ.

Связь величины сигнал/помеха и относительной погрешностью оценена с помощью коэффициента парной корреляции Пирсона r . Сила корреляционной связи определена по оценочной шкале Чедокка: $r_{0,4-0,7}$ — умеренная связь, $r_{0,7-0,9}$ — высокая связь, $r_{0,9-0,99}$ — весьма высокая связь [19]. Адекватность построенной регрессионной модели оценена коэффициентом детерминации R^2 . Корреляционная связь считалась значимой и неслучайной при уровне значимости $p < 0,05$.

Результаты исследований

Выполним оценку влияния величины сигнал/помеха на погрешность измерения параметров АЭ после фильтрации зашумленных сигналов для значений: -10 дБ, -5 дБ, 0 дБ, 5 дБ, 10 дБ [14]. На рис. 2 для фильтра Баттерворта показаны зависимости между величинами сигнал/помеха до и после обработки сигналов АЭ в виде скатерограммы.

Из представленных на рис. 2 зависимостей следует, что в силу уменьшения значения сигнал/помеха зашумленного сигнала на -1 дБ повышаются значения величины сигнал/помеха после обработки. Результат оценки парной корреляционной взаимосвязи показал, что между измерениями величины сигнал/помеха до и после обработки присутствует статистически значимая высокая обратная корреляционная взаимосвязь, а именно: $r = -0,972$, $R^2 = 0,946$ при $p = 0,005$. Данное утверждение характеризует уменьшение влияния составляющей помехи на сигнал АЭ за счет повышения величины сигнал/помеха на выходе фильтра. Это позволило выбрать значения сигнал/помеха для дальнейшей оценки, построения зависимости и разработки статистической модели влияния обсуждаемой величины на погрешность измерения параметров АЭ. Скатерограммы регрессионных зависимостей представлены на рис. 3.

Результаты статистического моделирования свидетельствуют о том, что с увеличением значений величины сигнал/помеха уменьшаются значения погрешности измерения параметров АЭ на выходе фильтра [5]. Величина погрешности не превышает 3 %. Вместе с тем между значениями величины сигнал/помеха и относительной погрешностью существует обратная корреляционная связь, характеризующая убывания значений погрешности измерений параметров АЭ. Запишем статистическую модель, характеризующую данную связь, с помощью уравнения линейной регрессии, ее функциональность оценим коэффициентом детерминации R^2 , а силу связи корреляцией r по значению значимости p модели.

Для метода фильтрации [5], основанного на фильтре Баттерворта, получены следующие значения:

$$y_{\delta\sigma} = -3,584 \cdot 10^{-5} \cdot \text{СП} + 2,904, R^2 = 0,964, p = 0,002;$$

$$y_{\delta U} = -4,625 \cdot 10^{-5} \cdot \text{СП} + 2,244 \\ \text{при } r = -0,871, R^2 = 0,759, p = 0,050;$$

$$y_{\delta\text{СКЗ}} = -3,584 \cdot 10^{-5} \cdot \text{СП} + 2,904 \\ \text{при } r = -0,981, R^2 = 0,964, p = 0,002;$$

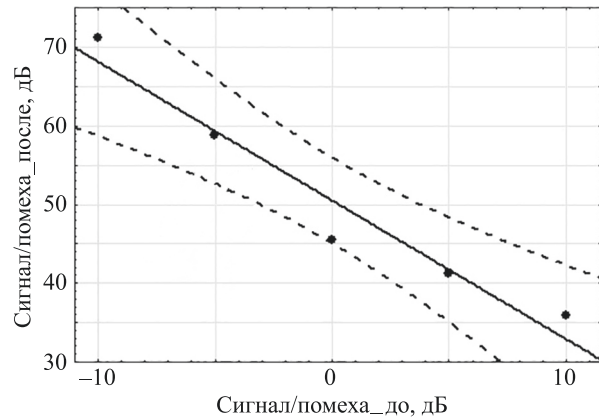


Рис. 2. Зависимости величин сигнал/помеха до и после обработки

● измеренные данные; — линия регрессии;
--- доверительный интервал

Fig. 2. Signal-to-noise ratio dependence before and after processing

Curve designations: ● measured data; — is the regression line;
--- confidence interval

$$y_{\delta\text{MARSE}} = -1,5 \cdot 10^{-5} \cdot \text{СП} + 2,902 \\ \text{при } r = -0,963, R^2 = 0,928, p = 0,008,$$

где СП — значение показателя сигнал/помеха.

Результат регрессионного анализа подтвердил влияние величины сигнал/помеха на погрешность измерений параметров АЭ. Факт данного влияния подтверждается статистически значимой сильной обратной корреляцией между величиной сигнал/помеха и относительной погрешностью. Минимальное ($R^2 = 0,759$) и максимальное ($R^2 = 0,964$) значения коэффициента детерминации (для метода [5]) ближе к единице, которая характеризует точность построенной статистической модели. Статистическая модель считается точной, когда величина коэффициента детерминации ближе к единице [7]. В данном исследовании значения значимости при $p = 0,05$ или $p < 0,05$ определяют неслучайность корреляционных связей. Следует также отметить, что параметры АЭ, измеренные при высоких значениях величины сигнал/помеха, могут быть необходимыми для использования при проведении предиктивной аналитики технического состояния объекта контроля [20].

Обсуждение

Амплитудные параметры сигналов демонстрируют наилучшую корреляционную связь со свойствами источника АЭ, наделяемыми объект контроля, и описывают энергетическое содержание акустических явлений [18]. Однако на точность и достоверность оценки амплитудных параметров АЭ сильно влияют помехи, снижая величину сигнал/помеха и увеличивая погрешность оценки параметров АЭ.

Рассматриваемые данные исследования демонстрируют статистически значимое влияние величины сигнал/помеха на относительную погрешность при оценке

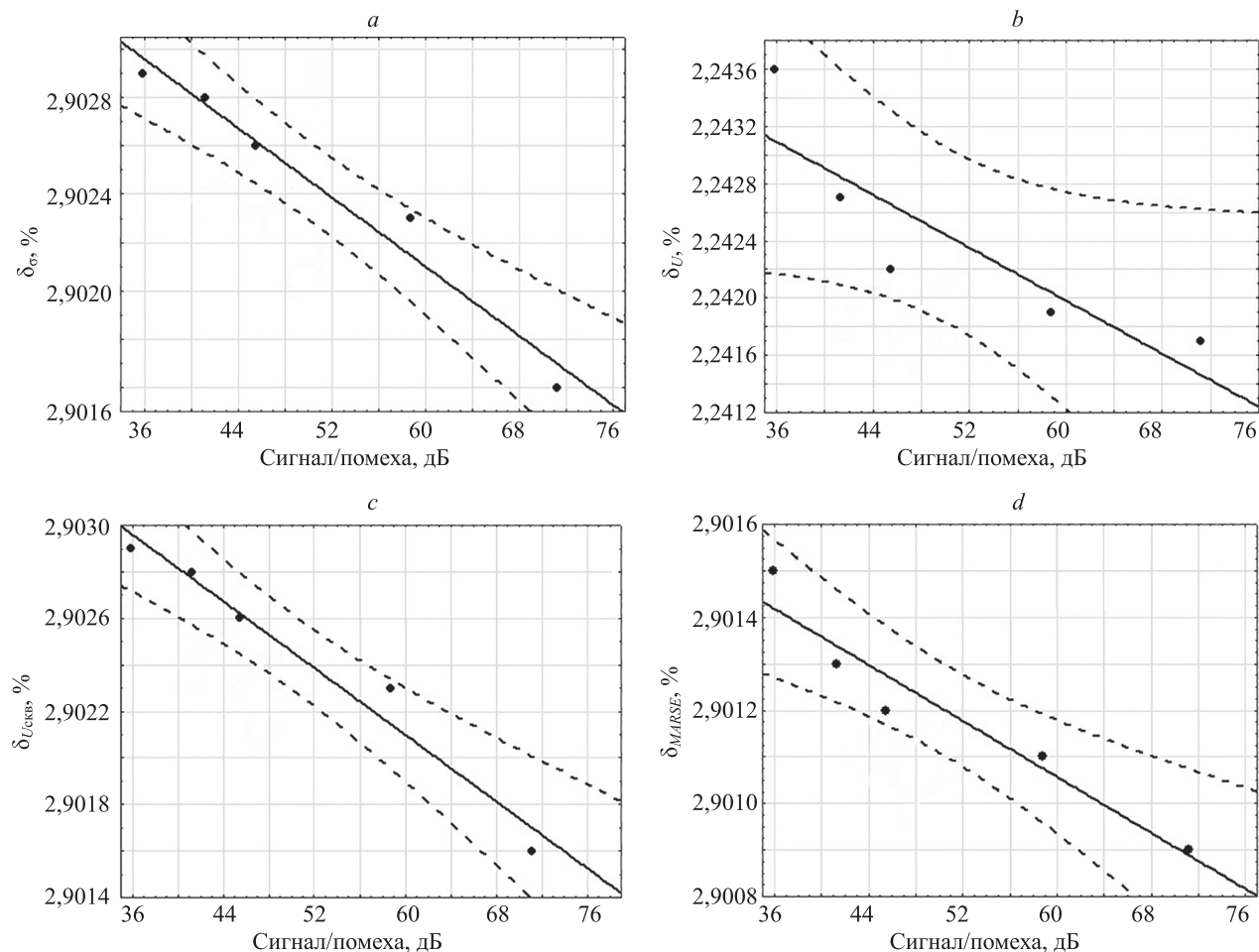


Рис. 3. Результаты погрешности измерения параметров акустической эмиссии в зависимости от величины сигнал/помеха для фильтра Баттерворта [5]: σ_U (a); U (b); U_{CK3} (c); $MARSE$ (d)

● измеренные данные; — линия регрессии; --- доверительный интервал

Fig. 3. Error of measurement of the acoustic emission parameters in dependence value signal-to-noise ratio for Butterworth filters [5]: σ_U (a); U (b); U_{CK3} (c); $MARSE$ (d)

Curve designations: ● measured data; — is the regression line; --- confidence interval

параметров АЭ. Величина относительной погрешности измерений параметров АЭ не превысила 3 % при значении сигнал/помеха 71 дБ. Данный результат обнаружен для фильтра, аппроксимированного полиномом Баттерворта [5]. Для подтверждения влияния величины сигнал/помеха на относительную погрешность рассматриваемые результаты были подвергнуты сравнению с имеющимися методами фильтрации помех АЭ. Для сравнения выбраны фильтры Бесселя [8] и вейвлета [15], статистические зависимости которых показаны в табл. 2.

Из зависимостей, представленных в табл. 2, видно, что аналогично фильтру Баттерворта в силу увеличения величины сигнал/помеха убывает значение относительной погрешности. Однако, если сравнить величину относительной погрешности между собой, фильтр, аппроксимированный полиномом Бесселя (табл. 2) на выходе системы обработки сигнала, наделает погрешность до 82 % при 25 дБ, в то время как для метода вейвлет-фильтрации данная величина составило 83 % при 8 дБ. В работе [21, С. 60] показано, что величина относительной погрешности в определя-

емых параметрах сигнала АЭ может быть неустойчива и находиться в диапазоне от 29,22 % до 82,72 %, соответственно.

Полученные результаты оценки в настоящей работе показали, что величины относительной погрешности для фильтров Бесселя и вейвлета находятся внутри обсуждаемого диапазона погрешностей, согласно работе [21, С. 60], тогда как для разработанного метода [5] они составили не более 3 %.

Высокое значение величины относительной погрешности связано с тем, что во время обработки фильтры Бесселя [8] и вейвлета [15] способны не только задерживать помехи, но и сглаживать, и искажать амплитудные составляющие экспериментально-тестовой формы сигнала, что также наблюдалось в результатах исследований [14]. Вейвлет-фильтр помимо сглаживания информационных составляющих АЭ, размывает форму сигнала, что было обнаружено в работах [22, 23]. Заметим, что ранее отмечалось повышение результативности обработки сложно структурированных сигналов при использовании фильтров: Баттерворта в работе [24], Бесселя в [25] и вейвлет в [26]. При различных

Таблица 2. Сравнение погрешностей измерения параметров акустической эмиссии в зависимости от величин сигнал/помеха для фильтров Бесселя [8] и вейвлет

Table 2. Comparison of measurement errors of acoustic emission parameters depending on the signal/noise values for the Bessel filter [8] and the wavelet filter

Параметр	Фильтр Бесселя	Вейвлет-фильтр
σ_U		
U		
$U_{СКЗ}$		
MARSE		

Примечание. Обозначения кривых: • измеренные данные; — линия регрессии; --- доверительный интервал

наборах входных данных измерений повышение результативности обработки сигнала для фильтра Баттерворта также было обнаружено в [26].

Между значениями величин сигнал/помеха и относительных погрешностей для фильтров Бесселя [8] и вейвлета [15] существует обратная связь.

Приведем показатели связи для метода фильтрации, основанного на фильтре Бесселя [14]:

$$y_{\delta\sigma} = -0,013 \cdot \text{СП} + 81,912 \\ \text{при } r = -0,810, R^2 = 0,657, p = 0,009;$$

$$y_{\delta U} = -9 \cdot 10^{-4} \cdot \text{СП} + 79,173 \\ \text{при } r = -0,928, R^2 = 0,862, p = 0,022;$$

$$y_{\delta \text{СКЗ}} = -0,013 \cdot \text{СП} + 81,919 \\ \text{при } r = -0,818, R^2 = 0,669, p = 0,090;$$

$$y_{\delta \text{MARSE}} = -0,007 \cdot \text{СП} + 81,781 \\ \text{при } r = -0,738, R^2 = 0,546, p = 0,153,$$

и для вейвлет-фильтрации [15]:

$$y_{\delta\sigma} = -0,029 \cdot \text{СП} + 80,467 \\ \text{при } r = -0,847, R^2 = 0,718, p = 0,069;$$

$$y_{\delta U} = -0,048 \cdot \text{СП} + 82,702 \\ \text{при } r = -0,687, R^2 = 0,473, p = 0,199;$$

$$y_{\delta \text{СКЗ}} = -0,029 \cdot \text{СП} + 80,465 \\ \text{при } r = -0,848, R^2 = 0,719, p = 0,069;$$

$$y_{\delta \text{MARSE}} = -0,037 \cdot \text{СП} + 80,343 \\ \text{при } r = -0,867, R^2 = 0,753, p = 0,056.$$

В теории статистической обработки измерений принято считать существование высокой связи полученной зависимости, если величина коэффициента корреляции $r > 0,700$.

В настоящей работе значения коэффициента корреляции для каждой построенной зависимости составили диапазоны: для метода обработки [5] $r = -0,871$ до $r = -0,981$ при $p < 0,05$; для фильтра Бесселя [8] $r = -0,738$ до $r = -0,928$ при $p < 0,05$ и $p > 0,05$; для метода фильтрации [15] $r = -0,687$ до $r = -0,867$ при $p > 0,05$. Высокая статистически значимая взаимосвязь показателей сигнал/помеха и относительной погрешности для метода [5] объясняется повышением

результативности обработки сигналов АЭ по точности и помехоустойчивости [14]. В связи со слабой результативностью фильтров вейвлет [15] и Бесселя [8] по точности и помехоустойчивости (табл. 2) значения коэффициента корреляции, характеризующего силу связи построенной зависимости, намного ниже, чем для метода, основанного на фильтре Баттерворта [5].

Коэффициент детерминации R^2 регрессионной модели для предложенного метода [5] имеет значение $R^2 > 0,759$, объясняющее сходимость данных измерений с линией регрессии при сравнении с фильтрами Бесселя $R^2 > 0,546$ [8] и вейвлета $R^2 > 0,473$ [15], полученных при помощи линейной модели.

Проведенные исследования показали, что снижение погрешности измерений параметров АЭ способен обеспечить метод обработки, основанный на использовании фильтра Баттерворта. Это обусловлено высокой помехоустойчивостью данного метода обработки [14]. Новизна исследования заключается в том, что для повышения достоверности оценки параметров АЭ на основе применения фильтров разработана статистическая модель, описывающая влияние величины сигнал/помеха на погрешность измерений. Ранее оценка данного влияния в существующих научных работах не была достаточно рассмотрена.

Дальнейшим исследованием может быть статистическая оценка параметров АЭ, характеризующих вторичные диагностические показатели АЭ и ее значимости для натурных измерений при реализации цифрового фильтра Баттерворта.

Заключение

В работе представлен результат статистической оценки влияния величины сигнал/помеха на погрешность измерения параметров акустической эмиссии при цифровой обработке сигналов фильтрами Баттерворта, Бесселя и вейвлет. Между величинами сигнал/помеха и относительной погрешности измерения параметров установлена обратная высокая статистически значимая корреляционная взаимосвязь. Установлено, что при высоких значениях величины сигнал/помеха фильтр Баттерворта способен обеспечить относительную погрешность измерения диагностических параметров акустической эмиссии, не превышающую 3 % при отношении сигнал/помеха 71 дБ.

Литература

1. He Y., Li M., Meng Z., Chen S., Huang S., Hu Y., Zou X. An overview of acoustic emission inspection and monitoring technology in the key components of renewable energy systems // *Mechanical Systems and Signal Processing*. 2021. V. 148. P. 107146. <https://doi.org/10.1016/j.ymssp.2020.107146>
2. Kharrat M.A., Ramasso E., Placet V., Boubakar M.L. A signal processing approach for enhanced acoustic emission data analysis in high activity systems: Application to organic matrix composites // *Mechanical Systems and Signal Processing*. 2016. V. 70-71. P. 1038–1055. <https://doi.org/10.1016/j.ymssp.2015.08.028>
3. Il K.K., Hwan R.U., Pil C.B. An appropriate thresholding method of wavelet denoising for dropping ambient noise // *International Journal of Wavelets, Multiresolution and Information Processing*. 2018. V. 16. N 3. P. 1850012. <https://doi.org/10.1142/S0219691318500121>

References

1. He Y., Li M., Meng Z., Chen S., Huang S., Hu Y., Zou X. An overview of acoustic emission inspection and monitoring technology in the key components of renewable energy systems. *Mechanical Systems and Signal Processing*, 2021, vol. 148, pp. 107146. <https://doi.org/10.1016/j.ymssp.2020.107146>
2. Kharrat M.A., Ramasso E., Placet V., Boubakar M.L. A signal processing approach for enhanced acoustic emission data analysis in high activity systems: Application to organic matrix composites. *Mechanical Systems and Signal Processing*, 2016, vol. 70-71, pp. 1038–1055. <https://doi.org/10.1016/j.ymssp.2015.08.028>
3. Il K.K., Hwan R.U., Pil C.B. An appropriate thresholding method of wavelet denoising for dropping ambient noise. *International Journal of Wavelets, Multiresolution and Information Processing*, 2018, vol. 16, no. 3, pp. 1850012. <https://doi.org/10.1142/S0219691318500121>

4. Barat V., Borodin Y., Kuzmin A. Intelligent AE signal filtering methods // *Journal of Acoustic Emission*. 2010. V. 28. P. 109–119.
5. Altay Y.A., Fedorov A.V., Stepanova K.A. Acoustic emission signal processing based on polynomial filtering method // *Proc. of the 2022 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*. 2022. P. 1320–1326. <https://doi.org/10.1109/ElConRus54750.2022.9755729>
6. Барат В.А. Развитие метода акустической эмиссии за счет автоматизации обработки данных, повышения помехоустойчивости и достоверности обнаружения трещиноподобных дефектов металлоконструкций: автореферат диссертации на соискание ученой степени доктора технических наук. М., 2019. 40 с.
7. Алтай Е., Федоров А.В., Степанова К.А. Оценка взаимосвязи информационных составляющих и помех сигналов акустической эмиссии // *Контроль. Диагностика*. 2022. Т. 25. № 6. С. 38–47. <https://doi.org/10.14489/td.2022.06.pp.038-047>
8. Алтай Е., Федоров А.В., Степанова К.А. Оценка влияния методов фильтрации на погрешность измерения параметров сигнала акустической эмиссии // *Международная конференция по мягким вычислениям и измерениям*. 2022. Т. 1. С. 24–27.
9. Paarman L.D. *Design and Analysis of Analog Filters: A Signal Processing Perspective*. NY: Kluwer Academic Publishers, 2001. 440 p.
10. Somefun O., Akingbade K., Dahunsi F. Uniformly damped binomial filters: five-percent maximum overshoot optimal response design // *Circuits, Systems, and Signal Processing*. 2022. V. 41. N 6. P. 3282–3305. <https://doi.org/10.1007/s00034-021-01931-2>
11. Быстров С.В., Вундер Н.А., Ушаков А.В. Решение проблемы сигнальной неопределенности при аналитическом конструировании последовательного компенсатора в задаче управления пьезопроводом // *Научно-технический вестник информационных технологий, механики и оптики*. 2016. № 3. С. 451–459. <https://doi.org/10.17586/2226-1494-2016-16-3-451-459>
12. Быстров С.В., Вундер Н.А., Синетова М.М., Ушаков А.В. Аналитическое конструирование последовательного компенсатора для систем с запаздыванием на основе модификации типовых полиномиальных моделей // *Труды СПИИРАН*. 2017. № 3(52). С. 115–136. <https://doi.org/10.15622/sp.52.6>
13. Сергеев А.Н., Степанова Л.Н., Кабанов С.И., Чернова В.В. Диагностический модуль акустико-эмиссионной системы с автоматической фильтрацией помех // *Датчики и системы*. 2020. № 5. С. 3–14. <https://doi.org/10.25728/datsys.2020.5.1>
14. Алтай Е., Федоров А.В., Степанова К.А., Кузванов Д.О. Оценка эффективности методов обработки сигналов акустической эмиссии при реализации полиномиальных цифровых фильтров // *Омский научный вестник*. 2022. № 3. С. 128–134. <https://doi.org/10.25206/1813-8225-2022-183-128-134>
15. Kharrat M., Ramasso E., Placet V., Baubakar M.L. A signal processing approach for enhanced Acoustic Emission data analysis in high activity systems: Application to organic matrix composites // *Mechanical Systems and Signal Processing*. 2016. V. 70-71. P. 1038–1055. <https://doi.org/10.1016/j.ymssp.2015.08.028>
16. Левин Б.Р. Теоретические основы статистической радиотехники. М.: Советское радио, 1968. 504 с.
17. Zakharov L.A., Martyshev D.A., Ponomareva I.N. Predicting dynamic formation pressure using artificial intelligence methods // *Journal of Mining Institute*. 2022. V. 253. N 1. P. 23–32. <https://doi.org/10.31897/PMI.2022.11>
18. Бехер С.А., Бобров А.Л. Основы неразрушающего контроля методом акустической эмиссии. Новосибирск: СГУПС, 2013. 145 с.
19. Салин В.Н., Чурилова Э.Ю. Практикум по курсу «Статистика» (в системе STATISTICA). М.: Перспектива, 2002. 188 с.
20. Elforjani M., Shanbr S. Prognosis of bearing acoustic emission signals using supervised machine learning // *IEEE Transactions on Industrial Electronics*. 2018. V. 65. N 7. P. 5864–5871. <https://doi.org/10.1109/TIE.2017.2767551>
21. Овчарук В.Н., Турисев Ю.А. Регистрация и обработка акустико-эмиссионной информации в многоканальных системах. Хабаровск: ТОГУ, 2017. 116 с.
22. Rakshit M., Das S. An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter // *Biomedical Signal Processing and Control*. 2018. V. 40. P. 140–148. <https://doi.org/10.1016/j.bspc.2017.09.020>
23. Altay Y.A., Kremlev A.S. Signal-to-noise ratio and mean square error improving algorithms based on newton filters for measurement ECG data processing // *Proc. of the 2021 IEEE Conference of Russian*
4. Barat V., Borodin Y., Kuzmin A. Intelligent AE signal filtering methods. *Journal of Acoustic Emission*, 2010, vol. 28, pp. 109–119.
5. Altay Y.A., Fedorov A.V., Stepanova K.A. Acoustic emission signal processing based on polynomial filtering method. *Proc. of the 2022 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2022, pp. 1320–1326. <https://doi.org/10.1109/ElConRus54750.2022.9755729>
6. Barat V.A. Development of the acoustic emission method by automating data processing, increasing noise immunity and detection fidelity of the crack-like defects in metal structures. *Dissertation abstract for the degree of doctor of technical sciences*. Moscow, 2019, 40 p. (in Russian)
7. Altay Ye., Fedorov A.V., Stepanova K.A. Estimation of relationship between information components and noise of acoustic emission signals. *Diagnostics*, 2022, vol. 25, no. 6, pp. 38–47. (in Russian). <https://doi.org/10.14489/td.2022.06.pp.038-047>
8. Altay Y., Fedorov A.V., Stepanova K.A. Assessment the effect of filtering methods on the measurement error of acoustic emission signal parameters. *International Conference on Soft Computing and Measurements*, 2022, vol. 1, pp. 24–27. (in Russian)
9. Paarman L.D. *Design and Analysis of Analog Filters: A Signal Processing Perspective*. NY: Kluwer Academic Publishers, 2001, 440 p.
10. Somefun O., Akingbade K., Dahunsi F. Uniformly damped binomial filters: five-percent maximum overshoot optimal response design. *Circuits, Systems, and Signal Processing*, 2022, vol. 41, no. 6, pp. 3282–3305. <https://doi.org/10.1007/s00034-021-01931-2>
11. Bystrov S.V., Vunder N.A., Ushakov A.V. Solution of signal uncertainty problem at analytical design of consecutive compensator in piezo actuator control. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2016, no. 3, pp. 451–459. (in Russian). <https://doi.org/10.17586/2226-1494-2016-16-3-451-459>
12. Bystrov S.V., Vunder N.A., Sinetova M.M., Ushakov A.V. Analytical design of consecutive compensator for control systems with delay based on modification of typical polynomial models. *SPIIRAS Proceedings*, 2017, no. 3(52), pp. 115–136. (in Russian). <https://doi.org/10.15622/sp.52.6>
13. Seryeznov A.N., Stepanova L.N., Kabanov S.I., Chernova V.V. Diagnostic module of acoustic emission system with automatic noise filtering. *Sensors & Systems*, 2020, no. 5, pp. 3–14. (in Russian). <https://doi.org/10.25728/datsys.2020.5.1>
14. Altay Y., Fedorov A.V., Stepanova K.A., Kuzivanov D.O. Estimating efficiency of acoustic emission signal processing methods in implementation of polynomial digital filters. *Omsk Scientific Bulletin*, 2022, no. 3, pp. 128–134. (in Russian). <https://doi.org/10.25206/1813-8225-2022-183-128-134>
15. Kharrat M., Ramasso E., Placet V., Baubakar M.L. A signal processing approach for enhanced Acoustic Emission data analysis in high activity systems: Application to organic matrix composites. *Mechanical Systems and Signal Processing*, 2016, vol. 70-71, pp. 1038–1055. <https://doi.org/10.1016/j.ymssp.2015.08.028>
16. Levin B.R. *Theoretical Background Of Statistical Radio Engineering*. Moscow, Sovetskoe radio Publ., 1968, 504 p. (in Russian)
17. Zakharov L.A., Martyshev D.A., Ponomareva I.N. Predicting dynamic formation pressure using artificial intelligence methods. *Journal of Mining Institute*, 2022, vol. 253, no. 1, pp. 23–32. <https://doi.org/10.31897/PMI.2022.11>
18. Bekher S.A., Bobrov A.L. *Fundamentals of Nondestructive Testing by the Method of Acoustic Emission*. Novosibirsk, STU, 2013, 145 p. (in Russian)
19. Salin V.N., Churilova E.Iu. *Practical Course on “Statistics”*. Moscow, Perspektiva Publ., 2002, 188 p. (in Russian)
20. Elforjani M., Shanbr S. Prognosis of bearing acoustic emission signals using supervised machine learning. *IEEE Transactions on Industrial Electronics*, 2018, vol. 65, no. 7, pp. 5864–5871. <https://doi.org/10.1109/TIE.2017.2767551>
21. Ovcharuk V.N., Turisev Iu.A. *Registration and Processing of Acoustic-Emission Information in Multichannel Systems*. Khabarovsk, PNU, 2017, 116 p. (in Russian)
22. Rakshit M., Das S. An efficient ECG denoising methodology using empirical mode decomposition and adaptive switching mean filter. *Biomedical Signal Processing and Control*, 2018, vol. 40, pp. 140–148. <https://doi.org/10.1016/j.bspc.2017.09.020>
23. Altay Y.A., Kremlev A.S. Signal-to-noise ratio and mean square error improving algorithms based on newton filters for measurement ECG

- Young Researchers in Electrical and Electronic Engineering (ElConRus). 2021. P. 1590–1595. <https://doi.org/10.1109/ElConRus51938.2021.9396391>
24. Altay Y.A., Kremlev A.S., Zimenko K.A., Margun A.A. The effect of filter parameters on the accuracy of ECG signal measurement // *Biomedical Engineering*. 2019. V. 53. N 3. P. 176–180. <https://doi.org/10.1007/s10527-019-09903-2>
 25. Avdeeva D.K., Kazakov V.Y., Natalinova N.M., Ivanov M.L., Yuzhakova M.A., Turushev N.V. The simulation results of the high-pass and low-pass filter effect on the quality of micropotential recordings on the electrocardiogram // *European Journal of Physical and Health Education*. 2014. V. 6. P. 1–10.
 26. Malghan P.G., Hota M.K. A review on ECG filtering techniques for rhythm analysis // *Research on Biomedical Engineering*. 2020. V. 36. N 2. P. 171–186. <https://doi.org/10.1007/s42600-020-00057-9>
- data processing. *Proc. of the 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus)*, 2021, pp. 1590–1595. <https://doi.org/10.1109/ElConRus51938.2021.9396391>
24. Altay Y.A., Kremlev A.S., Zimenko K.A., Margun A.A. The effect of filter parameters on the accuracy of ECG signal measurement. *Biomedical Engineering*, 2019, vol. 53, no. 3, pp. 176–180. <https://doi.org/10.1007/s10527-019-09903-2>
 25. Avdeeva D.K., Kazakov V.Y., Natalinova N.M., Ivanov M.L., Yuzhakova M.A., Turushev N.V. The simulation results of the high-pass and low-pass filter effect on the quality of micropotential recordings on the electrocardiogram. *European Journal of Physical and Health Education*, 2014, vol. 6, pp. 1–10.
 26. Malghan P.G., Hota M.K. A review on ECG filtering techniques for rhythm analysis. *Research on Biomedical Engineering*, 2020, vol. 36, no. 2, pp. 171–186. <https://doi.org/10.1007/s42600-020-00057-9>

Авторы

Федоров Алексей Владимирович — доктор технических наук, доцент, профессор, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57219346304](https://orcid.org/0000-0003-0612-922X), <https://orcid.org/0000-0003-0612-922X>, avfedorov@itmo.ru

Алтай Ельдос — аспирант, инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57194240500](https://orcid.org/0000-0002-3736-0291), <https://orcid.org/0000-0002-3736-0291>, aeldos@inbox.ru

Степанова Ксения Андреевна — кандидат технических наук, ассистент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57212027443](https://orcid.org/0000-0003-1811-3807), <https://orcid.org/0000-0003-1811-3807>, ledy.xs93@yandex.ru

Кузиванов Дмитрий Олегович — инженер, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0001-9661-9614>, kuzivanovdmitry@gmail.com

Authors

Alexey V. Fedorov — D. Sc., Associate Professor, Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57219346304](https://orcid.org/0000-0003-0612-922X), <https://orcid.org/0000-0003-0612-922X>, avfedorov@itmo.ru

Yeldos Altay — PhD Student, Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57194240500](https://orcid.org/0000-0002-3736-0291), <https://orcid.org/0000-0002-3736-0291>, aeldos@inbox.ru

Ksenia A. Stepanova — PhD, Assistant, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 57212027443](https://orcid.org/0000-0003-1811-3807), <https://orcid.org/0000-0003-1811-3807>, ledy.xs93@yandex.ru

Dmitry O. Kuzivanov — Engineer, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0001-9661-9614>, kuzivanovdmitry@gmail.com

Статья поступила в редакцию 06.07.2022
Одобрена после рецензирования 19.09.2022
Принята к печати 12.11.2022

Received 06.07.2022
Approved after reviewing 19.09.2022
Accepted 12.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1216-1225

УДК 53.072

Моделирование процесса стационарного термоотражения для измерения теплопроводности материалов

Анастасия Сергеевна Тукмакова^{1✉}, Пётр Сергеевич Демченко²,
Иван Леонидович Тхоржевский³, Анна Владимировна Новотельнова⁴,
Михаил Константинович Ходзицкий⁵

^{1,2,3,5} ООО «ТИДЕКС», Санкт-Петербург, 194292, Российская Федерация

^{1,2,3,4} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ astukmakova@itmo.ru✉, <https://orcid.org/0000-0002-9123-8983>

² petr.s.demchenko@gmail.com, <https://orcid.org/0000-0003-4948-2023>

³ iltkhorzhevskiy@itmo.ru, <https://orcid.org/0000-0002-9781-0610>

⁴ novotelnova@yandex.ru, <https://orcid.org/0000-0003-0073-2415>

⁵ khodzitskiy@yandex.ru, <https://orcid.org/0000-0001-7261-8350>

Аннотация

Предмет исследования. Измерение теплопроводности методом стационарного термоотражения основано на эффекте нагрева исследуемого образца лазерным излучением. Мощность отраженного от образца излучения определяется с помощью дополнительного зондирующего лазера. Изменение коэффициента отражения исследуемого материала при нагреве пропорционально изменению температуры образца. Поглощенная материалом мощность излучения выступает в качестве объемного источника теплоты. Величина теплопроводности рассчитана в соответствии с законом теплопроводности Фурье. При этом установка стационарного термоотражения требует калибровки, то есть проведения измерений теплопроводности эталонных образцов. Величина калибровочного коэффициента, в свою очередь, зависит от характеристик используемых в установке лазеров. В работе предложена численная модель нагрева образца, позволяющая учитывать форму и диаметр пучка падающего излучения, распределение мощности излучения по площади пучка, коэффициент поглощения и отражения материала. Предложена методика определения коэффициентов отражения и поглощения образцов, включающая экспериментальные измерения и расчеты с учетом эффекта Фабри–Перо. **Метод.** Исследования выполнены для образцов германия, кремния, арсенида галлия и ситалла. Облучение проведено диодным одномодовым лазером в непрерывном режиме с длиной волны 980 нм. Распределение мощности падающего излучения по площади пучка аппроксимировано в соответствии с функцией Гаусса в программе OriginPro. Выполнено измерение мощности излучения, прошедшего через образцы и отраженного от них. Температура образцов при облучении определена при помощи тепловизора. Коэффициенты отражения и поглощения образцов исследуемых материалов получены по результатам измерения мощности излучения с использованием математической модели взаимодействия плоскополяризованного ТЕ электромагнитного излучения с материалом. Сравнение результатов расчета с известными данными для образцов германия, кремния и ситалла показало их соответствие. Для арсенида галлия отмечено расхождение результатов расчета с данными, полученными другими исследователями. С целью изучения образцов арсенида галлия применена модель, учитывающая эффект Фабри–Перо, а оптические свойства определены численно, путем поиска минимума модулей передаточных функций прошедшего и отраженного излучений в программе MATLAB. Модель электромагнитного нагрева исследуемых образцов реализована в программной среде COMSOL Multiphysics. **Основные результаты.** Предложена методика определения коэффициентов отражения и поглощения материалов, исследуемых методом стационарного термоотражения. Представленная модель позволяет учесть форму, ширину пучка и распределение мощности излучения, а также величину поглощенной мощности излучения для каждого образца. Различие расчетных значений температуры образца с результатами измерений не превышает 9 %. **Практическая значимость.** Модель может быть применена для измерения теплопроводности объемных и тонкопленочных материалов с малоизученными свойствами.

Ключевые слова

стационарное термоотражение, коэффициент теплопроводности, оптические методы измерения теплопроводности, моделирование электромагнитного нагрева

Благодарности

Работа выполнена при поддержке гранта РФФИ № 22-22-00597 «Исследование теплопроводности тонкопленочных термоэлектриков оптическими методами».

Ссылка для цитирования: Тукмакова А.С., Демченко П.С., Тхоржевский И.Л., Новотельнова А.В., Ходзицкий М.К. Моделирование процесса стационарного термоотражения для измерения теплопроводности материалов // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1216–1225. doi: 10.17586/2226-1494-2022-22-6-1216-1225

Simulating the process of steady-state thermorefectance for measuring the thermal conductivity of materials

Anastasiia S. Tukmakova¹✉, Petr S. Demchenko², Ivan L. Tkhorzhevskiy³, Anna V. Novotelnova⁴, Mikhail K. Khodzitskiy⁵

^{1,2,3,5} ООО Tidex, Saint Petersburg, 194292, Russian Federation

^{1,2,3,4} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ astukmakova@itmo.ru✉, <https://orcid.org/0000-0002-9123-8983>

² petr.s.demchenko@gmail.com, <https://orcid.org/0000-0003-4948-2023>

³ iltkhorzhevskiy@itmo.ru, <https://orcid.org/0000-0002-9781-0610>

⁴ novotelnova@yandex.ru, <https://orcid.org/0000-0003-0073-2415>

⁵ khodzitskiy@yandex.ru, <https://orcid.org/0000-0001-7261-8350>

Abstract

The measurement of thermal conductivity by the method of steady-state thermorefectance is based on the effect of heating the sample by laser radiation. The power of reflected radiation from the sample is measured using an additional probe laser. The change in the reflection coefficient of the studied material due to the heating is proportional to the change of the sample temperature. The radiation power absorbed by the material acts as a volumetric heat source. The value of thermal conductivity is calculated according to the Fourier's law of thermal conductivity. A steady-state thermorefectance setup requires calibration, i.e., thermal conductivity measurement of reference samples. The value of the calibration coefficient, in turn, depends on the characteristics of the lasers used in the setup. In this work, a numerical model of sample heating is proposed which allows considering the shape and diameter of the incident radiation beam, the distribution of radiation power over the beam area, and the absorption and reflection coefficients of the material. A methodology for samples reflection and absorption coefficients determination, including measurements and calculations considering the Fabry-Perot effect, is proposed. The study was performed for the samples of germanium, silicon, gallium arsenide, and sital. The irradiation was performed with a continuous wave single-mode diode laser with a wavelength of 980 nm. The incident radiation power distribution over the beam area was approximated according to the Gaussian function in the OriginPro software. The radiation power that passed through the samples and reflected from them was measured. The temperature of the samples during irradiation was measured with a thermal imager. The reflection and absorption coefficients of the samples were determined from the results of radiation power measurements using a mathematical model of the interaction of plane-polarized TE electromagnetic radiation with the material. Comparison of the calculation results with the literature data for germanium, silicon, and sital samples showed their correspondence. For gallium arsenide there was a discrepancy between the calculation results and the literature data. For gallium arsenide samples, a model considering the Fabry-Perot effect was used, and the optical properties were determined numerically by searching for the minimum modules of the transfer functions of the transmitted and reflected radiation in the Matlab program. The model of electromagnetic heating of the investigated samples was implemented in the COMSOL Multiphysics software. A method for the determination of the reflection and absorption coefficients of materials investigated by the steady-state thermorefectance is proposed. The proposed model allows considering beam shape, width and distribution of the radiation power as well as the value of the absorbed radiation power for each sample. The difference between the calculated values of the sample temperature and the measurement results does not exceed 9 %. The model can be applied to measure the thermal conductivity of bulk and thin-film materials with unknown properties.

Keywords

steady-state thermorefectance, thermal conductivity coefficient, optical methods for thermal conductivity measurement, electromagnetic heating modelling

Acknowledgements

This work was supported by the Russian Science Foundation grant No. 22-22-00597 “Study of thermal conductivity of thin-film thermoelectrics by optical “methods”.

For citation: Tukmakova A.S., Demchenko P.S., Tkhorzhevskiy I.L., Novotelnova A.V., Khodzitskiy M.K. Simulating the process of steady-state thermorefectance for measuring the thermal conductivity of materials. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1216–1225 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1216-1225

Введение

Измерение теплопроводности материалов оптическими методами имеет ряд преимуществ, таких как отсутствие контактов и контактного сопротивления, высокая скорость, возможность проводить измерение нескольких образцов одновременно, минимизация вероятности повреждения образца. Наличие всех этих преимуществ делает оптические методы наиболее подходящими для измерения теплопроводности тонких пленок — объектов с толщиной от 10 нм до 2,5 мкм, и толстых пленок — объектов с толщиной от 2 до 100 мкм [1]. Измерение теплопроводности тонких пленок — оригинальная задача, для решения которой невозможно использовать классические методы измерений, разработанные для объемных образцов. Для подобных измерений в последнее время часто применяются нестационарные оптические методы: метод лазерной вспышки [2], термоотражение в частотной [3] и временной [4] областях. Однако данные методы не позволяют непосредственно получить величину теплопроводности. Как правило, в результате измерений возможно получить зависимость температуры тела от времени воздействия. Зная данную зависимость, можно рассчитать величину температуропроводности и далее, после проведения дополнительных измерений теплоемкости и плотности материала, — коэффициент теплопроводности. Обработка результатов измерения также связана с необходимостью точного определения фазового сдвига сигнала. Все это вносит трудности в обработку результатов, увеличивает время исследования, приводит к технологической сложности установок и повышению неопределенности результатов измерений.

Метод стационарного термоотражения (СТО) (steady-state thermoreflectance [5]) — оптический метод измерения теплопроводности, который, при этом, относится к стационарным. СТО обладает главным преимуществом стационарных методов — позволяет непосредственно определить величину теплопроводности, обрабатывая результаты эксперимента при помощи закона Фурье [6]. За счет того, что метод является оптическим, удается нивелировать недостатки, характерные для классических неоптических стационарных методов: длительное время проведения измерений; использование контактов; высокую чувствительность к потерям на конвекцию и лучистый теплообмен. СТО подходит как для измерения объемных материалов, так и для толстых и тонких пленок.

При СТО нагрев осуществляется лазером, модулированным прямоугольным сигналом, что позволяет достичь стационарного режима за время порядка десятков микросекунд. При измерении образцы, как правило, покрываются тонкой металлической пленкой с известным коэффициентом отражения, который изменяется при нагреве. Это изменение отслеживается сканирующим источником. Метод СТО предложен сравнительно недавно и уже использован для исследования теплопроводности образцов диоксида кремния, оксида алюминия, образцов алмаза и кварца, и показал высокую точность измерений [5]. СТО позволяет изме-

рять как продольную, так и поперечную составляющие теплопроводности.

Относительное изменение отражения $\Delta R/R$ при СТО пропорционально относительному изменению напряжения $\Delta V/V$, измеряемому детектором мощности излучения зондирующего лазера. Величина $\Delta V/V$, в свою очередь, пропорциональна абсолютному изменению температуры ΔT исследуемого образца, так что $\Delta V/V = \beta_{TO} \cdot \Delta T$, где β_{TO} — коэффициент термоотражения материала металлического покрытия либо образца. Принимая значение β_{TO} постоянным, можно рассчитать ΔT . При помощи образца с известной теплопроводностью, зная мощность поглощенного излучения, возможно проведение калибровки установки СТО — определение калибровочного коэффициента. После проведения калибровки возможно измерить теплопроводность других образцов.

Для проведения калибровки установки СТО целесообразно применять численное моделирование. Цель данной работы — создание и верификация численной модели нагрева различных образцов материалов за счет поглощения электромагнитного излучения при СТО, позволяющей проводить оценку поля температуры в образце. Модель учитывает оптические свойства исследуемых образцов и характер излучения (форму и диаметр пучка). Верификация модели включает в себя измерение температурного поля в образцах при помощи тепловизора и сравнение результатов моделирования с экспериментом.

Метод исследования

В экспериментальной части проведены измерения диаметра пучка падающего излучения. Определена функция зависимости мощности излучения от поперечной координаты. Выполнены измерения мощности падающего, прошедшего и отраженного от образцов излучений, а также температуры образцов с использованием тепловизора.

Обработка результатов эксперимента и моделирование включали в себя: аппроксимацию экспериментальной зависимости мощности излучения от поперечной координаты при помощи функции Гаусса; расчет стандартных отклонений; расчет коэффициента отражения и поглощения исследуемых образцов при помощи обработки данных о мощности падающего, прошедшего и отраженного излучения; численное моделирование процесса электромагнитного нагрева исследуемых образцов.

Исследуемые образцы. Исследование выполнено на четырех образцах: монокристаллический германий (Ge)¹, выращенный методом Чохральского (направление роста вдоль оси C_3); монокристаллический кремний (Si)²; монокристаллический арсенид галлия

¹ [Электронный ресурс]. Режим доступа: https://www.tydexoptics.com/ru/materials/for_transmission_optics/germanium/ (дата обращения: 28.09.2022).

² [Электронный ресурс]. Режим доступа: https://www.tydexoptics.com/ru/materials/for_transmission_optics/silicon/ (дата обращения: 28.09.2022).

(GaAs)¹, выращенный методом Чохральского (направление роста вдоль оси C_3) и ситалл СТ-50-12. При помощи микрометра (Micron 414780, класс точности — 1, погрешность 0,5 мкм) проведены измерения толщины l_s и диаметра d_s образцов. Геометрические параметры образцов представлены в табл. 1.

Характеристики излучения. Образцы прошли облучение непрерывным лазерным излучением (тип лазера — диодный, одномодовый, максимальная мощность 2 Вт) на длине волны 980 нм. Измерение значения мощности излучения в зависимости от поперечной координаты проведено при помощи метода ножа Фуко³. В результате получены зависимости после дифференцирования по x -координате (точки на рис. 1) и аппроксимации при помощи функции нормального распределения (сплошная линия на рис. 1). Значение стандартного отклонения σ рассчитано в программе OriginPro. Производная мощности излучения P по поперечной координате имеет вид:

$$\frac{dP_{in}(x)}{dx} = (P_{in}/(\omega_{st}\sqrt{\pi/2})) \cdot e^{-2(x/\omega_{st})^2},$$

где P_{in} — интегральная величина мощности; x — поперечная координата; ω_{st} — удвоенное стандартное отклонение.

Максимальное значение функции $\frac{dP_{in}(x)}{dx} P_{max}$ составляло 30,9 мВт/мм, значение P_{in} — $49,3 \pm 2,3$ мВт, значение стандартного отклонения $\sigma = \frac{\omega_{st}}{2} = 0,64$ мм.

Ширина пучка рассчитана как расстояние между точками (пунктирная линия на рис. 1), соответствующими значению мощности, равной:

$$\frac{\max\left(\frac{dP_{in}(x)}{dx}\right)}{e^2} = 4,2 \frac{\text{мВт}}{\text{мм}}.$$

Таким образом, ширина пучка составила 2,5 мм.

Измерение мощности падающего, прошедшего через образец и отраженного от образца излучения. Измерения выполнены фотодетектором (VLP-2000, предел измерений 2 Вт, погрешность измерений 0,06 Вт (3 %)). На рис. 2, *a* изображена схема эксперимента по измерению мощности прошедшего через образец излучения при нормальном падении излучения на образец. На рис. 2, *b* показана схема измерения мощности отраженного от образца излучения при падении пучка на образец под углом $\alpha = 45^\circ$. Угол падения регулировался при помощи механического ротатора производства ком-

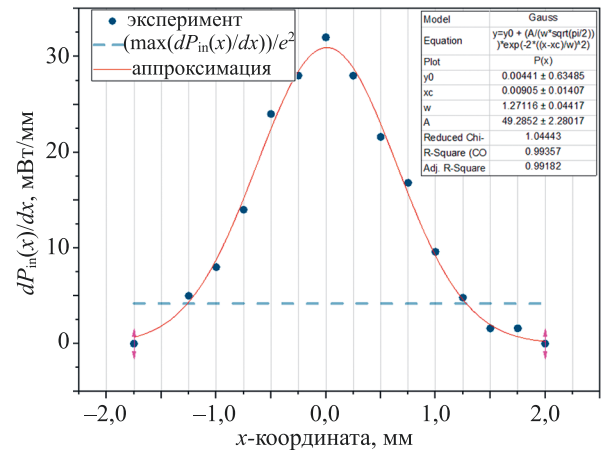


Рис. 1. Аппроксимация функции производной падающей мощности излучения по поперечной координате при помощи функции нормального распределения

Fig. 1. Approximation of the incident radiation power derivative with respect to the transverse coordinate using a normal distribution function

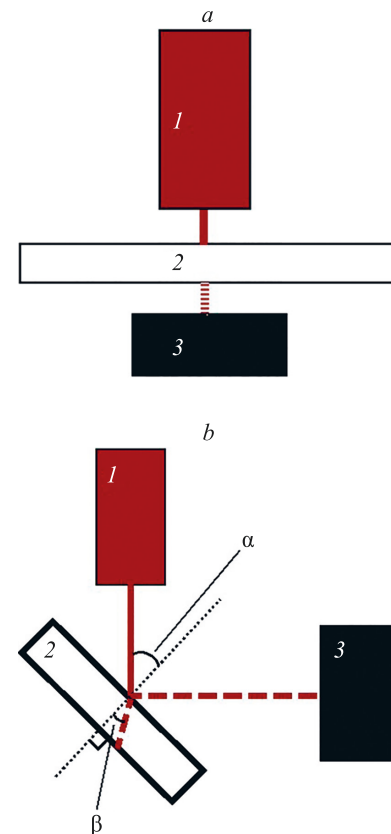


Рис. 2. Схема эксперимента. Измерение мощности прошедшего излучения нормально к плоскости образца (*a*) и отраженного излучения под углом $\alpha = 45^\circ$ (*b*).

β — угол преломления; 1 — источник излучения; 2 — образец; 3 — детектор, измеряющий мощность излучения

Fig. 2. The scheme of the experiment. Measurement of the transmitted radiation power normal to the sample plane (*a*); measurement of the reflected radiation power at an incident angle $\alpha = 45^\circ$ (*b*).

β — refraction angle, 1 — radiation source, 2 — sample, 3 — detector measuring radiation power

¹ [Электронный ресурс]. Режим доступа: https://www.tydexoptics.com/ru/materials/for_transmission_optics/gaas/ (дата обращения: 28.09.2022).

² [Электронный ресурс]. Режим доступа: http://speckip.ru/product/podlozhka_sitallovaya_st_50_1_1_0_6.html (дата обращения: 28.09.2022).

³ [Электронный ресурс]. Режим доступа: <https://cyberleninka.ru/article/n/sovremennyy-kolichestvennyy-tenevoy-metod-s-primeneniem-sredstv-kompyuternyh-tehnologiy/viewer> (дата обращения: 28.09.2022).

Таблица 1. Геометрические параметры образцов. Результаты измерения мощности падающего, прошедшего и отраженного излучения

Table 1. Geometric parameters of the samples. The results of the measurement of incident, transmitted, and reflected radiation power

Параметр		Материал			
		Ge	Si	Ситалл СТ-50-1	GaAs
Диаметр образца d_s , мм		17,0	15,5	55,0	25,4
Толщина образца i_s , мм		0,90	0,40	0,53	1,95
Мощность излучения, мВт	падающего P_{in}	54	250	250	665
	прошедшего P_{tr}	0,01	1,60	8,40	237
	отраженного P_{refl}	18,8	81,0	26,4	274,0

пании Standa LLC. Результаты измерений представлены в табл. 1.

Обработка результатов измерения и расчет коэффициентов отражения и поглощения исследуемых образцов

Воспользуемся математической моделью для взаимодействия электромагнитного излучения с ТЕ-поляризацией с материалом [7]. Зная мощности падающего P_{in} , прошедшего P_{tr} и отраженного P_{refl} излучений выполним расчет коэффициентов пропускания T и отражения R :

$$T = \frac{P_{tr}}{P_{in}}, \quad (1)$$

$$R = \frac{P_{refl}}{P_{in}}. \quad (2)$$

Рассчитаем амплитудный коэффициент отражения ρ на границе раздела «воздух–образец»:

$$\rho = \sqrt{R}. \quad (3)$$

Используя закон Френеля:

$$\rho = \frac{\cos \alpha - n_s \cos \beta}{\cos \alpha + n_s \cos \beta}, \quad (4)$$

и выражение для угла преломления, полученного из закона Снеллиуса

$$\beta = \arcsin\left(\frac{\sin \alpha}{n_s}\right), \quad (5)$$

можно выразить показатель преломления в материале n_s через ρ

$$n_s = \frac{1 + \rho}{1 - \rho}, \quad (6)$$

где α и β — углы падения и преломления излучения; показатель преломления воздуха равен 1.

Рассчитаем коэффициент экстинкции k_s по формуле:

$$k_s = \frac{c}{\omega l_s} \left\{ \ln \left[\frac{4n_s}{(n_s + 1)^2} \right] - \ln |\hat{H}_T| \right\}, \quad (7)$$

где $|\hat{H}_T| = \sqrt{T}$ — передаточная функция; ω — угловая частота; c — скорость света.

Коэффициент поглощения α_s :

$$\alpha_s = \frac{4\pi k_s}{\lambda}, \quad (8)$$

где λ — длина волны.

Расчет оптических свойств образцов с учетом эффекта Фабри–Перо. Для образца GaAs значение показателя преломления, рассчитанное по формуле (6) составило 4,2, в то время как в работе [8] — 3,52 (табл. 2). Разница значений может быть объяснена повторным отражением излучения на границе раздела сред «образец–воздух», обусловленным низким коэффициентом поглощением арсенида галлия. Для получения более точного результата необходимо использовать модель с учетом эффекта Фабри–Перо [7], описывающую отражение и пропускание на границе раздела сред «образец–воздух». В этом случае также используются формулы (1), (2) и (4)–(6), при условии, что формула (3) заменяется более сложной моделью.

Пусть τ — комплексный амплитудный коэффициент пропускания на границе раздела «воздух–образец», τ' и ρ' — комплексные амплитудные коэффициенты пропускания и отражения на границе раздела «образец–воздух». Эти коэффициенты можно выразить через комплексный показатель преломления материала образца $\hat{n}_s$ и углы α и β :

$$\rho' = \frac{\hat{n}_s \cos \beta - \cos \alpha}{\hat{n}_s \cos \beta + \cos \alpha},$$

$$\tau = \frac{2 \cos \alpha}{\cos \alpha + \hat{n}_s \cos \beta},$$

$$\tau' = \frac{2 \hat{n}_s \cos \beta}{\hat{n}_s \cos \beta + \cos \alpha},$$

$$\hat{n}_s = n_s - jk_s.$$

Выражение для модуля передаточной функции прошедшего излучения $|\hat{H}_T|$ будет иметь вид:

$$|\hat{H}_T| = \frac{|\hat{E}_{tr}|}{|\hat{E}_{in}|} = \sqrt{T} = \left| \frac{4 \hat{n}_s}{(\hat{n}_s + 1)^2} \exp \left\{ -\frac{\omega l_s}{c} (k_s + j[n_s - 1]) \right\} \widehat{FP} \right|, \quad (9)$$

где $|\hat{E}_{tr}|$ и $|\hat{E}_{in}|$ — модули амплитуды напряженности электрического поля прошедшего и падающего излучений.

Эффект Фабри–Перо:

$$\widehat{FP} = \left\{ 1 - \rho'^2 \exp \left[-2j\hat{n}_s \frac{\omega l_\beta}{c} \right] \right\}^{-1}, \quad (10)$$

где $l_\beta = l_s / \cos \beta$ — расстояние, пройденное волной в образце.

Подставляя (10) в (9) получим:

$$|\hat{H}_T| = \left| \frac{\frac{4\hat{n}_s}{(\hat{n}_s + 1)^2} \exp \left[-\frac{\omega l_s}{c} (k_s + j[n_s - 1]) \right]}{1 - \rho'^2 \exp \left[-2j\hat{n}_s \frac{\omega l_\beta}{c} \right]} \right|.$$

Модуль передаточной функции отраженного излучения $|\hat{H}_R|$ может быть представлен в виде отношения модуля амплитуды напряженности электрического поля отраженного излучения $|\hat{E}_{\text{refl}}|$ к $|\hat{E}_{\text{in}}|$:

$$|\hat{H}_R| = \frac{|\hat{E}_{\text{refl}}|}{|\hat{E}_{\text{in}}|} = \sqrt{R}.$$

При этом выражение для $|\hat{E}_{\text{refl}}|$ будет иметь вид:

$$|\hat{E}_{\text{refl}}| = |\rho| |\hat{E}_{\text{in}}| + \tau \tau' \rho' \exp \left[-2j\hat{n}_s \frac{\omega l_\beta}{c} \right] \widehat{FP} |\hat{E}_{\text{in}}|. \quad (11)$$

Разделим левую и правую часть уравнения (11) на $|\hat{E}_{\text{in}}|$:

$$\frac{|\hat{E}_{\text{refl}}|}{|\hat{E}_{\text{in}}|} = \left| \frac{\rho |\hat{E}_{\text{in}}|}{|\hat{E}_{\text{in}}|} + \frac{\tau \tau' \rho' \exp \left[-2j\hat{n}_s \frac{\omega l_\beta}{c} \right] \widehat{FP} |\hat{E}_{\text{in}}|}{|\hat{E}_{\text{in}}|} \right|. \quad (12)$$

Упрощая и подставляя в (12) выражение (10), получим:

$$|\hat{H}_R| = \left| \rho + \frac{\tau \tau' \rho' \exp \left[-2j\hat{n}_s \frac{\omega l_\beta}{c} \right]}{1 - \rho'^2 \exp \left[-2j\hat{n}_s \frac{\omega l_\beta}{c} \right]} \right|.$$

Расчет значений коэффициентов n_s и k_s проведен путем оптимизации функций $|\hat{H}_T|$ и $|\hat{H}_R|$ в среде MATLAB при помощи поиска минимума значения функции.

Сравним значения n_s и k_s для образцов германия, кремния и ситалла с данными, полученными в работах

[8–11]. Результаты расчетов совпали с данными других исследователей (табл. 2). Коэффициент поглощения в образцах рассчитывался по формуле (8).

Численное моделирование электромагнитного нагрева образцов

Стационарная модель нагрева исследуемых образцов при взаимодействии с электромагнитным излучением создана при помощи метода конечных элементов [12] в программе COMSOL Multiphysics. Для моделирования применен модуль теплопередачи в твердом теле¹. Непосредственное моделирование электромагнитного излучения не проводилось, так как такая задача требует разбиения модели на конечные элементы с размером не менее 100–200 нм (10–20 % от длины волны), что приводит к крайне низкой производительности расчета и невозможности его проведения. Вместо этого применено соответствующее тепловое граничное условие в модуле теплопередачи в твердом теле — объемный источник теплоты Q_{in} . Функция Q_{in} включала как оптические свойства материала, так и параметры пучка. Это позволило провести достоверное 3D-моделирование со средним временем расчета порядка нескольких секунд. Модель была разбита на конечные элементы в форме тетраэдров, с более мелким разбиением в центре образца.

Теплофизические свойства материалов Ge, Si, GaAs взяты из встроенной библиотеки COMSOL, свойства ситалла СТ-50-1 были взяты из [13] и представлены в табл. 3. Коэффициент теплоотдачи α_{TO} , характеризующий процесс конвективного теплообмена, рассчитан с применением теории подобия [14].

Начальная температура всех элементов считалась равной комнатной и составляла 24°C. При этом уравнение теплового баланса имеет вид

$$\rho C_p \frac{dT}{dt} + \nabla \mathbf{q} = Q_{\text{in}},$$

где ρ — плотность; C_p — удельная теплоемкость при постоянном давлении; T — абсолютная температура;

¹ Heat Transfer Module User's Guide [Электронный ресурс]. Режим доступа: <https://doc.comsol.com/5.4/doc/com.comsol.help.heat/HeatTransferModuleUsersGuide.pdf> (дата обращения: 28.09.2022).

Таблица 2. Оптические свойства образцов
Table 2. Optical properties of the samples

Параметр		Материал			
		Ge	Si	Ситалл СТ-50-1	GaAs
Коэффициент отражения R		0,39	0,32	0,11	0,38
Коэффициент пропускания T		0,00	0,01	0,03	0,37
Показатель преломления n_s	расчет	4,31	3,64	1,96	3,50
	данные работ [8–10]	4,38 [9]	3,58 [10]	—	3,52 [8]
Коэффициент экстинкции $k_s \cdot 10^2$	расчет	20,30	0,08	0,05	1,50
	данные работ [9, 10, 11]	23,20 [9]	0,08 [10]	—	1,55 [11]
Коэффициент поглощения $\alpha_s, \text{м}^{-1}$		$264 \cdot 10^3$	108,3	59,8	21,0

Таблица 3. Теплофизические свойства материалов и параметры теплообмена
Table 3. Thermophysical properties of the samples and the parameters of heat and mass transfer

Параметр	Материал			
	Ge	Si	Ситалл СТ-50-1	GaAs
Плотность ρ , кг/м ³	5320	2330	2460	5320
Теплоемкость C_p , Дж/(кг·К)	320	703	920	327
Коэффициент теплопроводности $\lambda_{\text{тепл}}$, Вт/(м·К)	61,0	162,0	1,2	51,8
Коэффициент теплоотдачи $\alpha_{\text{ТО}}$, Вт/(м ² ·К)	5,2	5,3	3,9	4,7

Q_{in} — объемный источник тепла; $\nabla \mathbf{q}$ — дивергенция вектора плотности теплового потока.

Рассчитаем вектор плотности теплового потока по закону Фурье:

$$\mathbf{q} = -\lambda_{\text{тепл}} \nabla T,$$

где $\lambda_{\text{тепл}}$ — коэффициент теплопроводности; ∇T — градиент температуры.

Функция Q_{in} включает в себя функцию нормального распределения по координатам x и y , позволяющую учитывать форму пучка, и затухание электромагнитного излучения в материале за счет поглощения вдоль z -координаты [15]. Принимая значения стандартных отклонений σ_x и σ_y , равными друг другу и σ , и принимая за начало координат центр образца с координатами (0,0,0), выражение будет иметь вид:

$$Q_{\text{in}}(x, y, z) = Q_0(1 - R) \frac{\alpha_s}{\pi \sigma^2} e^{-0,5 \left[\frac{x^2 + y^2}{\sigma^2} \right]} e^{-\alpha_s z}.$$

При моделировании учтем конвективный и лучистый теплообмен¹:

$$-\mathbf{nq}_{\text{конв}} = \alpha_{\text{ТО}}(T_{\text{окр}} - T),$$

¹ Heat Transfer Module User's Guide [Электронный ресурс]. Режим доступа: <https://doc.comsol.com/5.4/doc/com.comsol.help.heat/HeatTransferModuleUsersGuide.pdf> (дата обращения: 28.09.2022).

$$-\mathbf{nq}_{\text{луч}} = \varepsilon \sigma_{\text{СБ}}(T_{\text{окр}}^4 - T^4),$$

где $\mathbf{n}$ — вектор нормали; ε — степень черноты тела; $\sigma_{\text{СБ}}$ — постоянная Стефана–Больцмана; $\mathbf{q}_{\text{конв}}$ и $\mathbf{q}_{\text{луч}}$ — вектора плотности теплового потока (конвективного и лучистого соответственно); $T_{\text{окр}}$ — температура окружающей среды, равная 24°C.

Результаты и обсуждения

Для всех исследуемых образцов выполним измерения температурного поля при поглощении падающего излучения. Измерения проведем при помощи тепловизора Testo 881 с температурной чувствительностью до 50 мК. Результаты расчета поля температуры в образцах и экспериментальные термограммы представлены на рис. 3. На экспериментальных термограммах (рис. 3, а, с, е, к) представлено поле температур в образце и локальные значения в точках М1, М2 и М3. Рассчитанное поле температуры (рис. 3, б, д, ф, л) в образцах представлено для стороны образца, обратной облучаемой стороне (с этой же стороны снимались термограммы).

Проведем анализ изменения температуры относительно ее начального значения (24°C), полученной при расчете $\Delta T_{\text{расч}}$, и в результате эксперимента $\Delta T_{\text{эксп}}$. Для метода СТО интерес представляет температура в области облучения пучком, т. е. в центре образца (точка М3

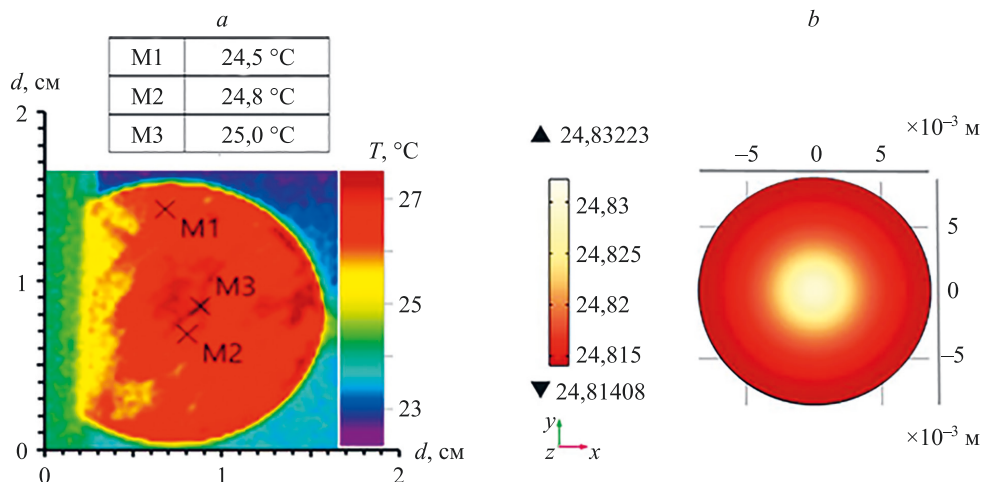


Рис. 3. Экспериментальные термограммы и расчеты полей температур образцов: Ge (а, б); Si (с, д); ситалла СТ-50-1 (е, ф) и GaAs (к, л)

Fig. 3. Experimental temperature patterns and calculation of the samples temperature fields: Ge (а, б); Si (с, д); sital CT-50-1 (е, ф) and GaAs (к, л)

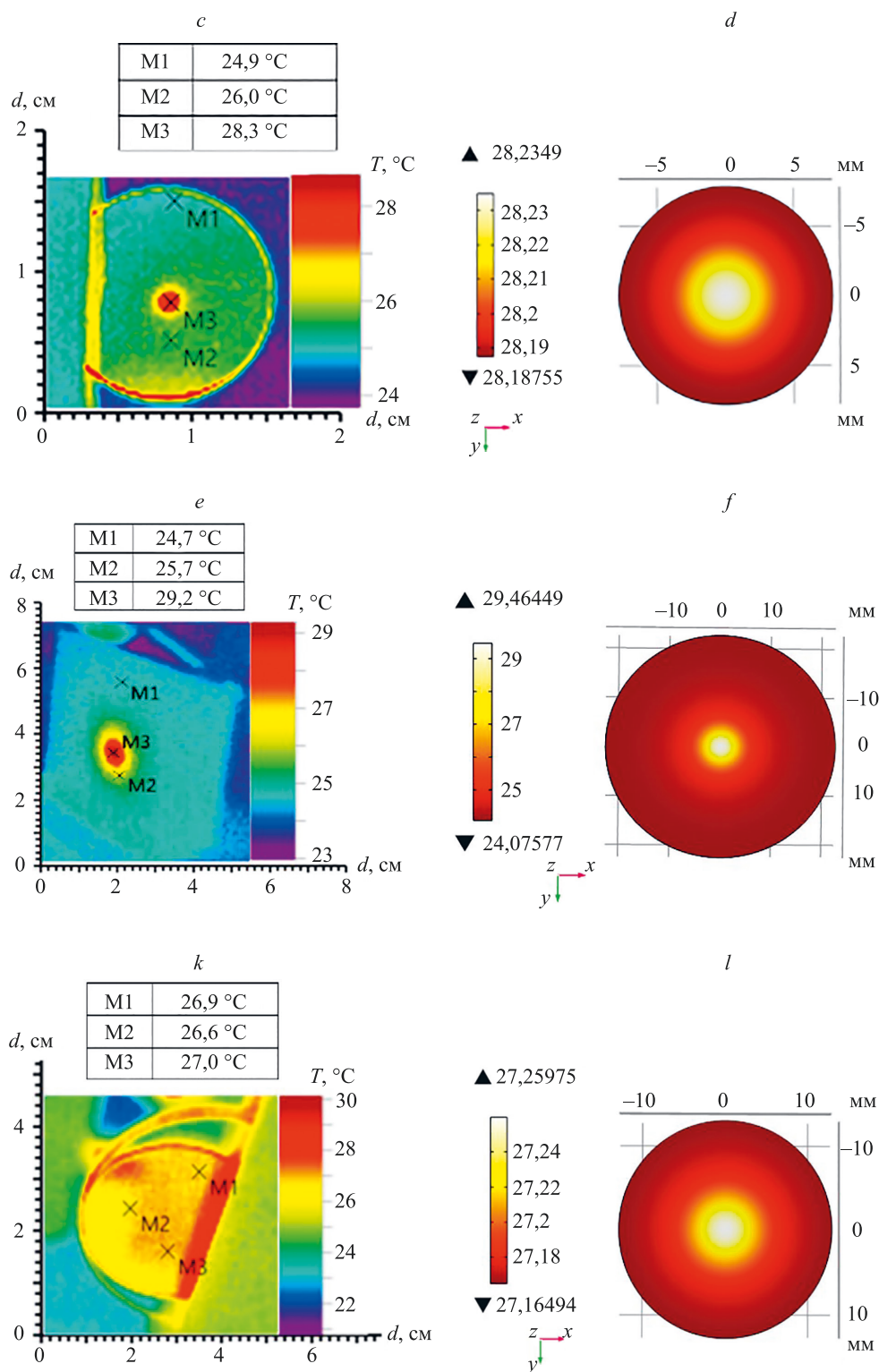


Рис. 3. Экспериментальные термограммы и расчеты полей температур образцов: Ge (a, b); Si (c, d); ситалл СТ-50-1 (e, f) и GaAs (k, l)

Fig. 3. Experimental temperature patterns and calculation of the samples temperature fields: Ge (a, b); Si (c, d); sitall CT-50-1 (e, f), and GaAs (k, l)

на экспериментальных термограммах). В таком случае отклонение данных расчета от экспериментальных данных составит:

$$\delta = (|\Delta T_{\text{расч}} - \Delta T_{\text{экс}}| / \Delta T_{\text{экс}}) \cdot 100\%.$$

Получим значения величины δ для образцов: 3,7 % (Ge), 1,6 % (Si), 5 % (ситалл СТ-50-1) и 8,3 % (GaAs).

Закключение

На основе предложенной методики создана численная модель нагрева материалов при поглощении электромагнитного излучения. Сравнение результатов расчета температуры в образцах и экспериментальных измерений показало, что расхождение данных не превышает 9 %. Полученный результат говорит об адекватности выбранной модели. Представленная модель может быть использована для калибровки установки стационарного термоотражения с применением эталонного образца, например сапфира, а также на изученных тонкопленочных образцах (например, на тонких плен-

ках кремния) и таких низкоразмерных объектах, как графен. В дальнейшем предложенный метод и установка могут применяться для исследования теплопроводности тонких пленок новых материалов.

Предложена методика расчета коэффициента поглощения, которая важна для точной оценки количества поглощенной образцом мощности излучения. Это необходимо для определения фактического теплового потока при расчете теплопроводности. Показано, что для образцов с высоким поглощением, таких как германий, достаточно простой модели. Для образцов с низким поглощением, таких как арсенид галлия, необходима модель, учитывающая эффект Фабри–Перо.

Литература

1. Маскаева Л.Н., Федорова Е.А., Марков В.Ф. Технология тонких пленок и покрытий: учебное пособие. Екатеринбург: Изд-во Урал. ун-та, 2019. 236 с.
2. Parker W.J., Jenkins R.J., Butler C.P., Abbott G.L. Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity // *Journal of Applied Physics*. 1961. V. 32. N 9. P. 1679–1684. <https://doi.org/10.1063/1.1728417>
3. Paddock C.A., Eesley G.L. Transient thermoreflectance from thin metal films // *Journal of Applied Physics*. 1986. V. 60. N 1. P. 285–290. <https://doi.org/10.1063/1.337642>
4. Schmidt J., Cheaito R., Chiesa M. A frequency-domain thermoreflectance method for the characterization of thermal properties // *Review of Scientific Instruments*. 2009. V. 80. N 9. P. 094901. <https://doi.org/10.1063/1.3212673>
5. Braun J.L., Olson D.H., Gaskins J.T., Hopkins P.E. A steady-state thermoreflectance method to measure thermal conductivity // *Review of Scientific Instruments*. 2019. V. 90. N 2. P. 024905. <https://doi.org/10.1063/1.5056182>
6. Zhao D., Qian X., Gu X., Jajja S.A., Yang R. Measurement techniques for thermal conductivity and interfacial thermal conductance of bulk and thin film materials // *Journal of Electronic Packaging*. 2016. V. 138. N 4. P. 040802. <https://doi.org/10.1115/1.4034605>
7. Naftaly M. *Terahertz Metrology*. Artech House, 2015. 378 p.
8. Skauli T., Kuo P.S., Vodopyanov K.L., Pinguet T.J., Levi O., Eyres L.A., Harris J.S., Fejer M.M., Gerard B., Becouarn L., Lallier E. Improved dispersion relations for GaAs and applications to nonlinear optics // *Journal of Applied Physics*. 2003. V. 94. N 10. P. 6447–6455. <https://doi.org/10.1063/1.1621740>
9. Amotchkina T., Trubetskov M., Hahner D., Pervak V. Characterization of e-beam evaporated Ge, YbF₃, ZnS, and LaF₃ thin films for laser-oriented coatings // *Applied Optics*. 2020. V. 59. N 5. P. A40–A47. <https://doi.org/10.1364/AO.59.000A40>
10. Green M.A. Self-consistent optical parameters of intrinsic silicon at 300 K including temperature coefficients // *Solar Energy Materials and Solar Cells*. 2008. V. 92. N 11. P. 1305–1310. <https://doi.org/10.1016/j.solmat.2008.06.009>
11. Adachi S. Optical dispersion relations for GaP, GaAs, GaSb, InP, InAs, InSb, Al_xGa_{1-x}As, and In_{1-x}Ga_xAsy P_{1-y} // *Journal of Applied Physics*. 1989. V. 66. N 12. P. 6030–6040. <https://doi.org/10.1063/1.343580>
12. Reddy J.N. *Introduction to the Finite Element Method*. 4th ed. New York: McGraw-Hill, 2019.
13. Егоров В.Н., Кондратенков В.И., Килессо В.С. Теплофизические свойства некоторых стекол и ситаллов // *Теплофизика высоких температур*. 1972. Т. 10. № 5. С. 1122–1123.
14. Москаленко Л.В., Стреляев Д.В. *Теплофизика: учебное пособие*. М.: МГТУ ГА, 2014. 100 с.
15. Yang S.T., Matthews M.J., Elhadj S., Cooke D., Guss G.M., Draggoo V.G., Wegner P.J. Comparing the use of mid-infrared versus far-infrared lasers for mitigating damage growth on fused silica // *Applied Optics*. 2010. V. 49. N 14. P. 2606–2616. <https://doi.org/10.1364/AO.49.002606>

References

1. Maskaeva L.N., Fedorova E.A., Markov V.F. *Technology of Thin Films and Coatings*. Yekaterinburg, Izatelstvo Uralskogo Universiteta, 2019, 236 p. (in Russian)
2. Parker W.J., Jenkins R.J., Butler C.P., Abbott G.L. Flash method of determining thermal diffusivity, heat capacity, and thermal conductivity. *Journal of Applied Physics*, 1961, vol. 32, no. 9, pp. 1679–1684. <https://doi.org/10.1063/1.1728417>
3. Paddock C.A., Eesley G.L. Transient thermoreflectance from thin metal films. *Journal of Applied Physics*, 1986, vol. 60, no. 1, pp. 285–290. <https://doi.org/10.1063/1.337642>
4. Schmidt J., Cheaito R., Chiesa M. A frequency-domain thermoreflectance method for the characterization of thermal properties. *Review of Scientific Instruments*, 2009, vol. 80, no. 9, pp. 094901. <https://doi.org/10.1063/1.3212673>
5. Braun J.L., Olson D.H., Gaskins J.T., Hopkins P.E. A steady-state thermoreflectance method to measure thermal conductivity. *Review of Scientific Instruments*, 2019, vol. 90, no. 2, pp. 024905. <https://doi.org/10.1063/1.5056182>
6. Zhao D., Qian X., Gu X., Jajja S.A., Yang R. Measurement techniques for thermal conductivity and interfacial thermal conductance of bulk and thin film materials. *Journal of Electronic Packaging*, 2016, vol. 138, no. 4, pp. 040802. <https://doi.org/10.1115/1.4034605>
7. Naftaly M. *Terahertz Metrology*. Artech House, 2015, 378 p.
8. Skauli T., Kuo P.S., Vodopyanov K.L., Pinguet T.J., Levi O., Eyres L.A., Harris J.S., Fejer M.M., Gerard B., Becouarn L., Lallier E. Improved dispersion relations for GaAs and applications to nonlinear optics. *Journal of Applied Physics*, 2003, vol. 94, no. 10, pp. 6447–6455. <https://doi.org/10.1063/1.1621740>
9. Amotchkina T., Trubetskov M., Hahner D., Pervak V. Characterization of e-beam evaporated Ge, YbF₃, ZnS, and LaF₃ thin films for laser-oriented coatings. *Applied Optics*, 2020, vol. 59, no. 5, pp. A40–A47. <https://doi.org/10.1364/AO.59.000A40>
10. Green M.A. Self-consistent optical parameters of intrinsic silicon at 300 K including temperature coefficients. *Solar Energy Materials and Solar Cells*, 2008, vol. 92, no. 11, pp. 1305–1310. <https://doi.org/10.1016/j.solmat.2008.06.009>
11. Adachi S. Optical dispersion relations for GaP, GaAs, GaSb, InP, InAs, InSb, Al_xGa_{1-x}As, and In_{1-x}Ga_xAsy P_{1-y}. *Journal of Applied Physics*, 1989, vol. 66, no. 12, pp. 6030–6040. <https://doi.org/10.1063/1.343580>
12. Reddy J.N. *Introduction to the Finite Element Method*. 4th ed. New York, McGraw-Hill, 2019.
13. Egorov V.N., Kondratenkov V.I., Kileso V.S. Thermophysical Properties of Some Glasses and Glass-Ceramics. *Teplofizika vysokikh temperatur*, 1972, vol. 10, no. 5, pp. 1122–1123.
14. Moskalenko L.V., Streljaev D.V. *Thermophysics*. Moscow, MSTUCA, 2014, 100 p. (in Russian)
15. Yang S.T., Matthews M.J., Elhadj S., Cooke D., Guss G.M., Draggoo V.G., Wegner P.J. Comparing the use of mid-infrared versus far-infrared lasers for mitigating damage growth on fused silica. *Applied Optics*, 2010, vol. 49, no. 14, pp. 2606–2616. <https://doi.org/10.1364/AO.49.002606>

Авторы

Тукмакова Анастасия Сергеевна — кандидат технических наук, старший научный сотрудник, ООО «Тидекс», Санкт-Петербург, 194292, Российская Федерация; доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57103805600](https://orcid.org/0000-0002-9123-8983), <https://orcid.org/0000-0002-9123-8983>, astukmakova@itmo.ru

Демченко Пётр Сергеевич — научный сотрудник, ООО «Тидекс», Санкт-Петербург, 194292, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57194214776](https://orcid.org/0000-0003-4948-2023), <https://orcid.org/0000-0003-4948-2023>, petr.s.demchenko@gmail.com

Тхоржевский Иван Леонидович — научный сотрудник, ООО «Тидекс», Санкт-Петербург, 194292, Российская Федерация; аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 57216155194](https://orcid.org/0000-0002-9781-0610), <https://orcid.org/0000-0002-9781-0610>, iltkhorzhevskiy@itmo.ru

Новотельнова Анна Владимировна — кандидат технических наук, доцент, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 56436689700](https://orcid.org/0000-0003-0073-2415), <https://orcid.org/0000-0003-0073-2415>, novotelnova@yandex.ru

Ходзицкий Михаил Константинович — кандидат физико-математических наук, начальник сектора, ООО «Тидекс», Санкт-Петербург, 194292, Российская Федерация, [sc 16444444600](https://orcid.org/0000-0001-7261-8350), <https://orcid.org/0000-0001-7261-8350>, khodzitskiy@yandex.ru

Authors

Anastasiia S. Tukmakova — PhD, Senior Researcher, OOO Tidex, Saint Petersburg, 194292, Russian Federation; Associate Professor, ITMO University, 197101, Russian Federation, [sc 57103805600](https://orcid.org/0000-0002-9123-8983), <https://orcid.org/0000-0002-9123-8983>, astukmakova@itmo.ru

Petr S. Demchenko — Scientific Researcher, OOO Tidex, Saint Petersburg, 194292, Russian Federation; PhD Student, ITMO University, 197101, Russian Federation, [sc 57194214776](https://orcid.org/0000-0003-4948-2023), <https://orcid.org/0000-0003-4948-2023>, petr.s.demchenko@gmail.com

Ivan L. Tkhorzhevskiy — Scientific Researcher, OOO Tidex, Saint Petersburg, 194292, Russian Federation; PhD Student, ITMO University, 197101, Russian Federation, [sc 57216155194](https://orcid.org/0000-0002-9781-0610), <https://orcid.org/0000-0002-9781-0610>, iltkhorzhevskiy@itmo.ru

Anna V. Novotelnova — PhD, Associate Professor, Associate Professor, ITMO University, 197101, Russian Federation, [sc 56436689700](https://orcid.org/0000-0003-0073-2415), <https://orcid.org/0000-0003-0073-2415>, novotelnova@yandex.ru

Mikhail K. Khodzitskiy — PhD (Physics & Mathematics), Head of Research Department, OOO Tidex, Saint Petersburg, 194292, Russian Federation, [sc 16444444600](https://orcid.org/0000-0001-7261-8350), <https://orcid.org/0000-0001-7261-8350>, khodzitskiy@yandex.ru

Статья поступила в редакцию 29.07.2022

Одобрена после рецензирования 06.10.2022

Принята к печати 27.11.2022

Received 29.07.2022

Approved after reviewing 06.10.2022

Accepted 27.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

doi: 10.17586/2226-1494-2022-22-6-1226-1236

УДК 533.6

Математическое и компьютерное моделирование однорядных и двухрядных шестилопастных винтокольцевых движителей

Сергей Юрьевич Дудников¹, Михаил Павлович Булат²,
Леонид Олегович Вокин³✉, Павел Николаевич Кузнецов⁴,
Павел Сергеевич Чернышов⁵

^{1,2,3,4} Севастопольский государственный университет, Севастополь, 299053, Российская Федерация

^{2,3,5} Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация

¹ sydudnikov@sevsu.ru, <https://orcid.org/0000-0003-1955-4313>

² bulat_mh@mail.ru, <https://orcid.org/0000-0001-9727-673X>

³ leonidvokin@mail.ru✉, <https://orcid.org/0000-0003-2944-1822>

⁴ PNKuznetsov@sevsu.ru, <https://orcid.org/0000-0002-1732-922X>

⁵ pashachp8@gmail.com, <https://orcid.org/0000-0003-1433-4194>

Аннотация

Предмет исследования. Рассмотрена проблема компьютерного моделирования эффективных движителей типа винт в кольце. По опыту применения в рулевых фенестронах вертолетов спаренных лопастей известно, что такая конфигурация создает меньше шума по сравнению с равномерным расположением лопастей по окружности. Однако поток за таким винтом менее равномерный, чем у обычного винта в кольце. Для беспилотных воздушных судов мультикоптерного типа и аэротакси ключевой проблемой является полет на режимах взлета и посадки, а также акустические и вихревые поля, создаваемые движителями на данных режимах. Снижение уровня шума у движителей со спаренными лопастями потенциально может сопровождаться усилением нестационарных вихревых воздействий на летательный аппарат, а также снижением удельной тяги. В данной работе предложена методика численного моделирования винтокольцевых движителей на режиме взлета и посадки, определения оптимального угла между лопастями, а также в сравнении винтокольцевого движителя со спаренными X-образными лопастями с обычными винтами. **Метод.** Расчет турбулентных течений выполнен с использованием нестационарных осредненных по Рейнольдсу уравнений Навье–Стокса с привлечением модели Shear Stress Transport (SST) модели турбулентности и моделирования крупных вихрей с применением модели подсеточной вязкости WALE (Wall-Adapting Local Eddy-viscosity). В расчетах использована модификация γ - Re_θ Transition SST модели турбулентности Лэнгтри–Ментера, в которой имеются соотношения для критерия перемежаемости. Это позволило учесть ламинарно-турбулентный переход и появление тонких ламинарных отрывных пузырей, влияющих как на тягу винта, так и на неравномерность течения за ним. Тестирование проведено на четырехлопастных винтокольцевых движителях по известным результатам эталонных экспериментов Центрального аэрогидродинамического института имени профессора Н.Е. Жуковского.

Основные результаты. Тестирование γ - Re_θ Transition Shear Stress Transport (tSST) модели турбулентности Лэнгтри–Ментера показало, что она лучше стандартной SST-модели воспроизводит зависимость коэффициентов тяги и мощности от угла установки лопастей. Расчеты показали наличие четко выраженного оптимума по углу между спаренными лопастями. Выполненное сравнение трехлопастного, шестилопастного одинарного и шестилопастного со сдвоенными лопастями винтокольцевых движителей подтвердило, что последний вариант имеет несколько лучшие характеристики тяги и создает существенно меньший уровень шума на местности.

Практическая значимость. Исследованные характеристики винтокольцевого движителя продемонстрировали перспективы применения винтов со сдвоенными лопастями в воздушных судах с вертикальным взлетом и посадкой. Оработанная численная методика может быть непосредственно использована для промышленных расчетов винтов и вентиляторов.

Ключевые слова

беспилотный летательный аппарат, беспилотное воздушное судно, винтокольцевой движитель, математическое и компьютерное моделирование, оптимизация, фенестрон

Благодарности

Работа выполнена при финансовой поддержке Министерства науки и высшего образования Российской Федерации в ходе реализации проекта «Фундаментальные основы механики, систем контроля и управления беспилотных авиационных систем с формообразующими конструкциями, глубоко интегрированными с силовыми установками, и уникальными свойствами, не применяемыми сегодня в пилотируемой авиации», № FEFM-2020-0001.

Ссылка для цитирования: Дудников С.Ю., Булат М.П., Вокин Л.О., Кузнецов П.Н., Чернышов П.С. Математическое и компьютерное моделирование однорядных и двухрядных шестилопастных винтокольцевых движителей // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1226–1236. doi: 10.17586/2226-1494-2022-22-6-1226-1236

Modeling and simulation of one- and two-row six-bladed ducted fans

Sergey Yu. Dudnikov¹, Mikhail P. Bulat², Leonid O. Vokin³, Pavel N. Kuznetsov⁴,
Pavel S. Chernyshov⁵

^{1,2,3,4} Sevastopol State University, Sevastopol, 299053, Russian Federation

^{2,3,5} Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation

¹ sydudnikov@sevsu.ru, <https://orcid.org/0000-0003-1955-4313>

² bulat_mh@mail.ru, <https://orcid.org/0000-0001-9727-673X>

³ leonidvokin@mail.ru, <https://orcid.org/0000-0003-2944-1822>

⁴ PNKuznetsov@sevsu.ru, <https://orcid.org/0000-0002-1732-922X>

⁵ pashachp8@gmail.com, <https://orcid.org/0000-0003-1433-4194>

Abstract

The problem of simulation of efficient ducted fan type propulsors is considered. From experience of operation of twin blades in fantails of helicopters, it is known that this configuration creates less noise compared to a uniform arrangement of the blades around the circumference. However, the flow behind such fan is less uniform than that of a conventional ducted fan. For multicopter-type unmanned aircraft and air taxis, the key problem is flight in take-off and landing modes as well as acoustic and vortex fields created by propulsors in these modes. The decrease in the noise level in propellers with twin blades can potentially be accompanied by an increase in non-stationary vortex effects on the aircraft as well as a decrease in specific thrust. The objectives were to develop a method for simulation of ducted fan propellers in the takeoff and landing mode, to determine the optimal angle between the blades, and to compare a ducted fan with twin X-shaped blades to conventional blade position. Turbulent flows were calculated using transient Reynold-averaged Navier-Stokes equations, complemented by SST turbulence model, and large eddy simulation with WALE subgrid viscosity model. The calculations used the modification γ - Re_θ Transition SST of the Langtry-Menter turbulence model, where there are relations for the intermittency criterion, which made it possible to consider the laminar-turbulent transition and the appearance of thin laminar separation bubbles that affect both the thrust of the propeller and the non-uniformity of the flow behind it. Testing was carried out on four-bladed propellers according to the known results of the TsAGI reference experiments. Testing of the γ - Re_θ Transition SST Langtry-Menter turbulence model showed that it reproduces the dependence of the thrust coefficient and power factor on the blade angle better than the standard SST model. Calculations have shown that there is a clearly defined optimum angle between the paired blades. A comparison of three-bladed, six-bladed single and six-bladed propellers with twin blades showed that the latter option has slightly better thrust characteristics and creates a significantly lower noise level on the ground. The studied characteristics of ducted fans demonstrate the prospects for the use of propellers with twin blades in aircraft with vertical takeoff and landing. The developed numerical method can be directly used for industrial calculations of propellers and fans.

Keywords

unmanned aerial vehicle, unmanned aircraft, ducted fan, modeling and simulation, optimization, fenestron

Acknowledgements

This work was financially supported by the Ministry of Science and Higher Education of the Russian Federation in the course of the project “Fundamental bases of mechanics, control and management systems for unmanned aircraft systems with shaping structures deeply integrated with propulsion systems and unique properties not used today in manned aviation”, No. FEFM-2020-0001.

For citation: Dudnikov S.Yu., Bulat M.P., Vokin L.O., Kuznetsov P.N., Chernyshov P.S. Modeling and simulation of one- and two-row six-bladed ducted fans. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1226–1236 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1226-1236

Введение

Цель работы — разработка метода численного моделирования винтокольцевых движителей (ВКД) и исследование винтов со спаренными лопастями (рис. 1). Винт в кольце является более сложным движителем, чем винт. Впервые он широко был применен в качестве

рулевого винта на вертолетах. Такие винты получили название фенестрон (зарегистрированное название, принадлежащее компании Eurocopter, в настоящее время Airbus Helicopters). Спаренные лопасти также получили впервые распространение на фенестронах вертолетов как более «тихие» движители, чем обычные винты.



Рис. 1. Фенестрон со спаренными лопастями вертолета Ka-62

Fig. 1. Fenestron with twin blades of the Ka-62 helicopter

Среди потенциальных транспортных малогабаритных беспилотных летательных аппаратов (БЛА) получил развитие новый класс сравнительно небольших мультироторных летательных аппаратов. Квадрокоптеры (летательные аппараты с четырьмя несущими винтами, две пары из которых вращаются в противоположных направлениях) обладают рядом преимуществ (компактность, маневренность, малая взлетная масса при значительной массе полезной нагрузки) перед другими типами БЛА [1]. Для обеспечения эффективного функционирования на различных режимах полета на этапе проектирования необходимо комплексное решение ряда проблем, связанных с улучшением аэродинамических характеристик летательного аппарата и его различных систем. В частности, важным является вопрос, связанный с определением характеристик винтового движителя и его влияния на другие части БЛА [2]. Имеющиеся результаты численных и экспериментальных исследований показали, что работа винтового движителя оказывает существенное влияние на аэродинамические силы и моментные характеристики БЛА [3, 4].

Улучшение аэродинамических и аэроакустических характеристик квадрокоптера тесно связано с определением оптимальных параметров ВКД. Эта задача является сложной и многокритерийной, наличие кольца еще более усложняет анализ, так как добавляются дополнительные переменные, описывающие взаимодействие винта и профилированного канала [5–9].

Один из инструментов оптимизации существующих и поиска новых схем и компоновок БЛА — численное моделирование. При этом имеется возможность не только рассчитывать тяговые характеристики ВКД, но и исследовать течение, формирующееся за винтом. В работе [10] выполнен обзор экспериментальных и численных исследований. Исследования движителей подобного типа осуществляются методом натурного эксперимента [11–14] и численными методами, в том числе RANS/URANS [15–17] и вихреразрешающими подходами к моделированию турбулентных течений [18]. Сравнение результатов расчета течения в ВКД, перемещающемся со скольжением, с эксперименталь-

ными данными [19] продемонстрировали хорошее совпадение. Разработанные программные средства в настоящее время позволяют моделировать особенности обтекания как изолированного несущего винта с упругими лопастями, совершающими маховое движение, как у вертолета, так и совместно с фюзеляжем и рулевым винтом [20–22]. В частности, вихреразрешающие методы моделирования турбулентных течений Detached Eddy Simulation (DES) применяются в работе [23] для решения сопряженной задачи расчета аэроупругой деформации лопастей вертолетного винта.

Важное значение также приобретает анализ акустического шума, генерируемого ВКД и БЛА в целом. Регистрация и обработка акустического сигнала представляет интерес для идентификации БЛА и выполняемого им маневра, а также для извлечения информации, полезной для контроля за местоположением и параметрами движения. Жесткие требования по уровню шума на местности предъявляются к БЛА транспортного назначения и будущему перспективному виду транспорта — аэротакси.

Возможности по моделированию аэродинамики винтов, которые открываются с появлением нелинейной нестационарной вихревой теории винта на основе тонкой несущей поверхности, позволили приступить к исследованию аэроакустических характеристик винтов в дальнем поле. Метод определения нелинейных аэродинамических характеристик в нестационарной постановке на основе тонкой несущей поверхности представлен в работе [24]. В [25] смоделировано обтекание шестилопастного винта с использованием TVD-схем (Total Variation Diminishing) на тетраэдральной сетке, а расчет акустического шума в дальнем поле произведен с помощью методики Фокса Уилльямса–Хокингса. Данная методика использована в настоящей работе. Моделирование обтекания несущего винта в общем случае движения, определение его аэродинамических и акустических характеристик проведены в работах [26, 27].

Для изучения закономерностей движения квадрокоптера часто используется математическая модель, описывающая движение твердого тела с шестью степенями свободы [28, 29], которая может применяться и с учетом увеличения подъемной силы за счет экранного эффекта, возникающего вблизи поверхности земли [30].

В данной работе предложена методика численного эксперимента и исследован винт в кольце ВКЗХ-К6 со двоянными лопастями (рис. 2), где К6 — тип лопасти.

Геометрическая модель

Рассмотрим геометрическую модель со следующими параметрами: угол смещения лопастей винта в двухрядной конфигурации $\Delta\psi$, продольное смещение лопастей δ , угол отклонения ВКД от вертикали γ . Величины $\Delta\psi$ и δ являются варьируемыми и подлежат оптимизации, угол наклона — внешняя переменная или параметр решаемой задачи γ (примем $\gamma = 0$).

Для сравнения выбраны ВКД с трех-, четырех- и шестилопастными винтами (рис. 3, а–с), а также двухрядным шестилопастным винтом (рис. 3, d). Отработка

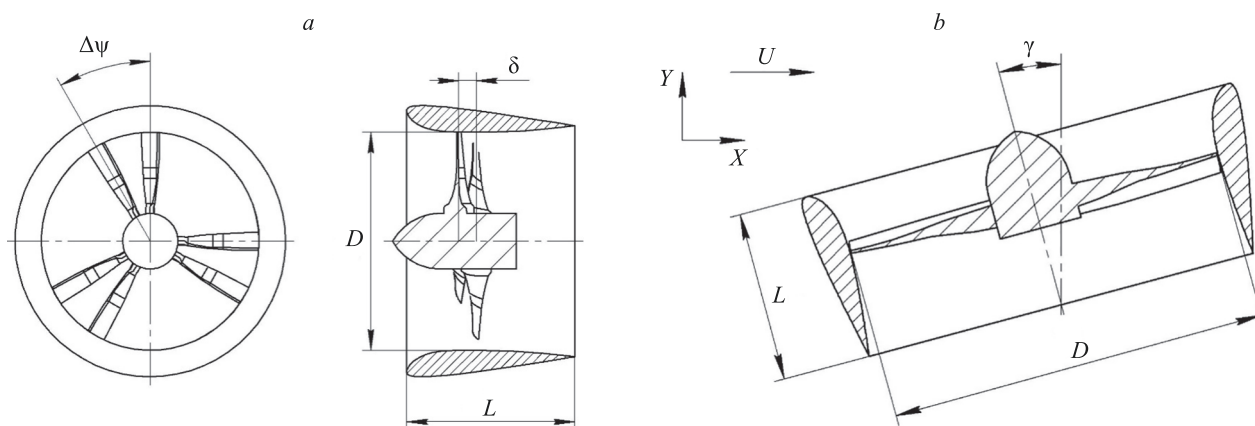


Рис. 2. Схема исследуемого винтокольцевого движителя: $\Delta\psi$ — угол смещения лопастей винта; D — диаметр винта; δ — продольное смещение лопастей; L — осевая протяженность винтокольцевого движителя; U — скорость набегающего потока; γ — угол отклонения винтокольцевого движителя от вертикали.

a — вид при $\gamma = 0$; b — вид при $\gamma \neq 0$

Fig. 2. Scheme of the researched ducted fan: $\Delta\psi$ — displacement angle of the fan blades; D — fan diameter; δ — longitudinal displacement of the blades; L — axial extension of the ducted fan; U — incoming flow velocity; γ — angle of deviation of the ducted fan from the vertical

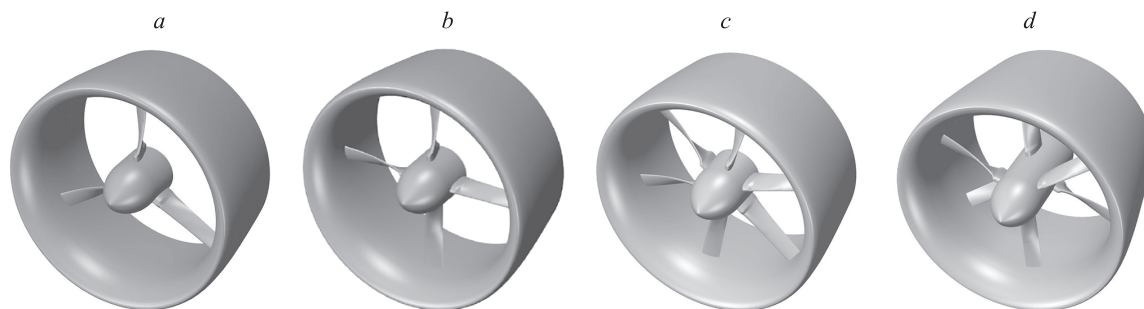


Рис. 3. Винтокольцевые движители, выбранные для сравнительных исследований: BK3-K6 (a), BK4-K6 (b), BK6-K6 (c), BK3 × 3-K6 (d)

Fig. 3. Ducted fans selected for comparative studies

численной методики выполнена на ВКД с четырехлопастным винтом BK4-K184B.

Моделирование выполнено для диаметров: винта D_B , равном 760 мм, и втулки винта — $0,258 D_B$, при характеристиках: лопастей винтов K6 (рис. 4) и винта K184B (рис. 5). Относительная ширина лопасти принята $\bar{b} = 0,0875$.

Во внутренних сечениях лопасти винта K6 до $r < 0,572R$ использованы профили П-117, во внешних — профили П-105а. У профиля П-105а кривизна постоянная $\bar{f} = 7\%$ (горизонтальный участок $\bar{f} = 7\%$, рис. 4, a), у профиля П-117 меняется по закону $\bar{f} = 11 - \frac{c}{3}$ (криволинейный участок при $\bar{r} \lesssim 0,6$, рис. 4, a).

Угол установки лопастей $\varphi_{0,75}$ на относительном радиусе $\bar{r} = r/R = 0,75$, где R — радиус винта задан на уровне $\varphi_{0,75} = 28^\circ$. Данный размер угла выбран исходя из известных эффективных углов установки существующих тяжело нагруженных винтов и угла поворота профиля на комеле лопатки, который не должен превышать 90° .

Расчетная область и разностная сетка

Расчетная область состоит из вращающейся (ротор) и стационарной (статор) подобластей, для учета взаи-

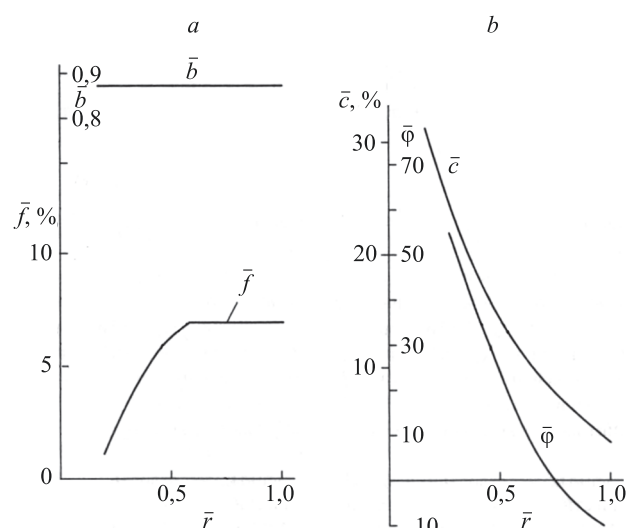


Рис. 4. Геометрические характеристики лопасти винта K6: $\bar{f}$ и $\bar{b}$ — относительные кривизна и ширина профиля (a); $\bar{c}$ — относительная толщина профиля и $\bar{\varphi}$ — крутка лопасти (b); $\bar{r}$ — относительный радиус сечения лопасти

Fig. 4. Geometric characteristics of the K6 fan blade $\bar{f}$ — relative profile curvature; $\bar{b}$ — relative fan blade width (a); $\bar{c}$ — relative profile thickness; $\bar{\varphi}$ — blade twist (b); $\bar{r}$ — relative sectional radius of the fan blade

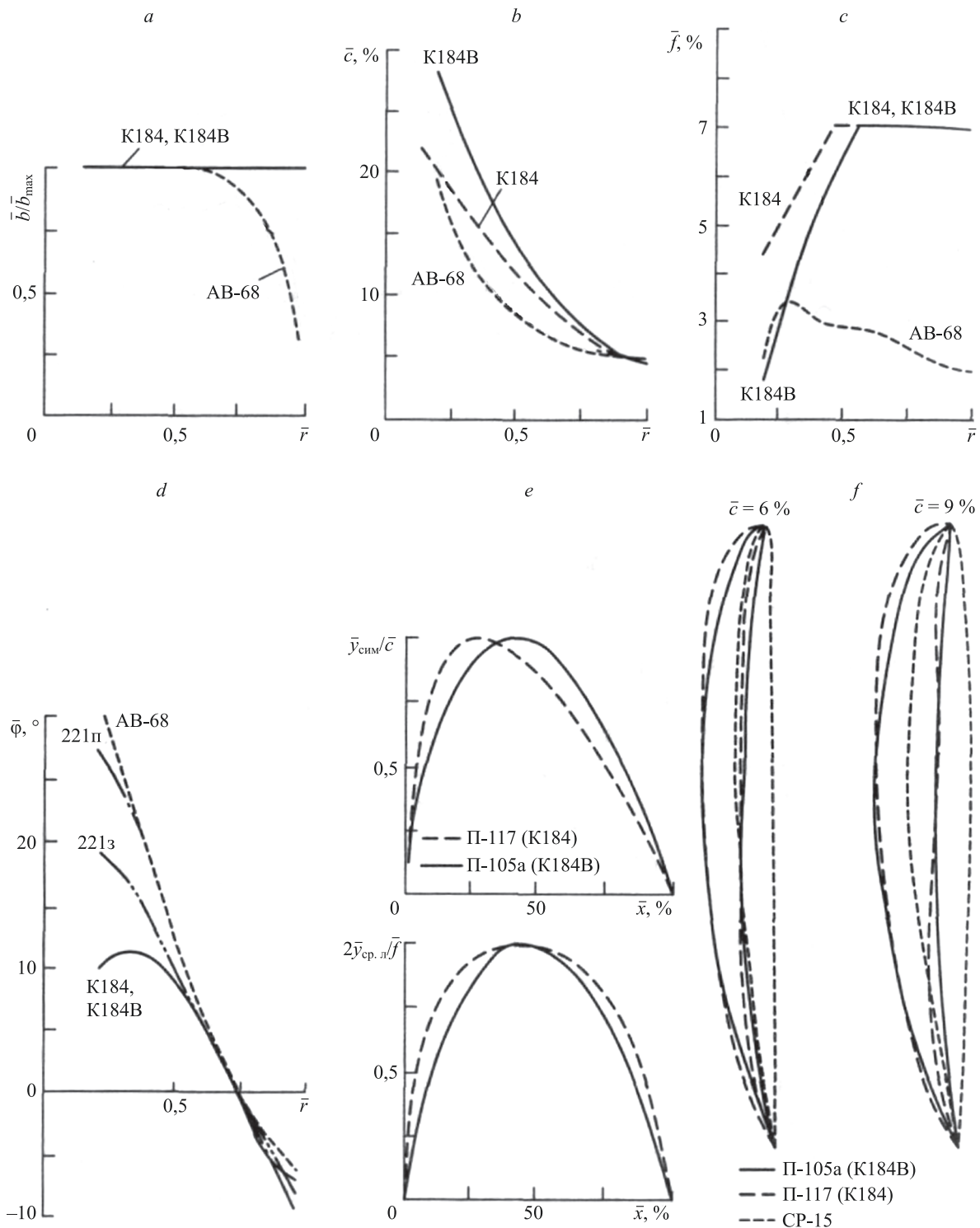


Рис. 5. Сравнение характеристик лопасти винта K184/K184B и самолетных винтов (a-d) и серии профилей (e, f).

221п (передний) и 221з (задний) — винты самолета Ан-22; АВ-68 — винт самолета Ан-24; П-117, П-105а, СР-15 — профили $\bar{y}_{сим}$ и $\bar{y}_{ср.л}$ — координаты симметричной части и средней линии профиля; $\bar{x}$ — относительная длина хорды

Fig. 5. Characteristics of the K184 / K184V fan blade and the series of profiles used. For comparison, the characteristics of aircraft fans 221п (front), 221з (rear) of the An-22 aircraft and AV-68 of the An-24 aircraft are given: $\bar{f}$ — relative profile curvature, $\bar{c}$ — relative profile thickness, $\bar{\varphi}$ — blade twist, $\bar{y}_{сим}$ — coordinate of the symmetrical part of the profile, $\bar{y}_{ср.л}$ — profile midline coordinate, $\bar{x}$ — relative chord length

модействия которых и сшивки результатов расчетов в каждой из подобластей применена технология скользящих сеток. Граница разбиения не является конструктивным элементом устройства, а представляет собой

вспомогательную поверхность вращения, разделяющую расчетную сетку на несколько частей. Ротор рассчитан во вращающейся системе координат, а статор — в стационарной. При таком подходе к решению задачи

изменение расчетной сетки при изменении положения вращающегося тела не происходит, поскольку сетка в роторе вращается за подвижными деталями устройства.

Расчет выполним на неструктурированной сетке, состоящей из тетраэдральных ячеек. Общее количество ячеек составляет примерно три миллиона для основных расчетов. К твердым стенкам примыкает часть сетки с призматическими ячейками. Толщина призматического слоя предположительно равна 1 мм. Сетка сгенерирована со сгущением на всех твердых поверхностях, с целью получения параметра $y^+ \approx 1$. Для этого задана толщина первого слоя в 0,002 мм, отношение толщин соседних слоев — 2, общее количество слоев — 15. Дополнительное условие на допустимый угол кривизны в 4° задано на входной кромке с целью неискажения ее поверхности при генерации сетки. В связи с симметрией винта и осевым режимом обтекания смоделирован не весь винт, а сектор с одной сдвоенной лопастью с постановкой периодических граничных условий на меридиональных плоскостях сектора.

Математическая модель и метод

Выполним численные расчеты при фиксированной скорости вращения $f = 3000$ об/мин. Число Рейнольдса рассчитано по хорде лопасти и ее концевой скорости и равно $Re = 5,8 \cdot 10^5$, при этом скорость набегающего потока равна нулю. Расчеты проведены для углов смещения лопастей винта $\Delta\psi$ в интервале от 20 до 40° и продольных смещений лопастей винта δ — от 63 до 91 мм.

Для моделирования течения, индуцированного вращением лопастей, использованы полные уравнения Навье–Стокса, описывающие течение вязкого сжимаемого газа. Моделирование турбулентности осуществлено при помощи нестационарных осредненных по Рейнольдсу уравнений Навье–Стокса с привлечением γ - Re_θ Transition Shear Stress Transport (tSST) модели турбулентности Лэнгтри–Ментера. Проведено моделирование крупных вихрей с использованием модели подсеточной вязкости WALE (Wall-Adapting Local Eddy-viscosity). Для вычисления плотности выбрана модель идеального газа. Вязкость среды считается постоянной.

Для дискретизации основных уравнений применен метод конечных объемов на неструктурированных сетках и среднемедианный контрольный объем [31]. Интегрирование по времени выполнено методом Рунге–Кутты 3-го порядка. Дискретизация невязких потоков осуществлена при помощи монотонной противопоточной схемы для законов сохранения Monotonic Upstream Schemes for Conservation Laws (MUSCL), а вязких потоков — с помощью централизованной схемы 2-го порядка точности. Использование схемы MUSCL позволило повысить порядок аппроксимации по пространственным переменным без потери монотонности решения и соблюсти условие TVD. Схема представляет собой комбинацию централизованных конечных разностей 2-го порядка и диссипативного члена, для переключения между которыми служит ограничитель потока, построенный на основе характеристических переменных. Определение градиента и псевдолапласиана

в срединной точке грани контрольного объема произведено на основе соотношений, приспособленных для расчетов на сильно растянутых сетках, используемых в пограничном слое. Для решения системы разностных уравнений применен геометрический многосеточный метод [32]. Система сеток различной разрешающей способности построена при помощи метода схлопывающихся граней.

На входе в расчетную область задана скорость и полная температура, а на остальных границах зафиксировано статическое давление. Степень турбулентности на входной границе установлена равной 5 %, а отношение турбулентной вязкости к динамической вязкости воздуха отличается в 10 раз. На стенках для динамических переменных применены условия прилипания и непротекания, а для температуры — равенства нулю теплового потока.

Для ускорения сходимости задано ненулевое начальное распределение скоростей и давления. Для их нахождения использована модель замороженного ротора, которая позволила получить близкое к реальному начальное распределение искомых функций на достаточно грубой сетке. В рамках приближения замороженного ротора геометрия замирает в определенном положении, что позволило исследовать поле потока для выбранного положения ротора (лопасти неподвижны относительно канала, а к окружающей области приложены центробежные силы). При использовании модели замороженного ротора также отключается обновление положения вращающейся сетки относительно неподвижной, что уменьшило время расчета одной итерации.

В результате численного моделирования ВКД получены значения: поля давления; скорости в расчетной зоне и непосредственно на поверхности лопасти, что позволяет определить аэродинамические характеристики движителя на различных режимах.

По результатам численного расчета определены: сила тяги, создаваемая лопастями винта P_d ; сила тяги кольца ВКД P_k ; сила сопротивления обтекателя втулки воздушного винта P_0 ; момент сопротивления вращению M . Полученные силовые характеристики позволили определить суммарную силу тяги винта при условии $P = P_d + P_k - P_0$.

Определим необходимую мощность на валу:

$$N = \pi n M / 30,$$

где N — потребляемая мощность винта; n — скорость вращения.

Вычислим коэффициент тяги по формуле

$$\alpha = P / \rho n_c^2 D^4,$$

где P — сила тяги; ρ — плотность воздуха; n_c — число оборотов в секунду; D — диаметр винта.

Коэффициент мощности вычислим по формуле

$$\beta = N / \rho n_c^3 D^5.$$

Результаты аэродинамических расчетов используем для моделирования распространения акустических возмущений в ближнем и дальнем полях и для опреде-

ления аэроакустических характеристик ВКД. Расчеты проведены с шагом по времени $5 \cdot 10^{-6}$, что соответствует разрешению волн с частотой до 50 кГц. Расчет акустических характеристик ВКД основан на интегральном методе Фокса Уилльямса–Хокинга. Решение акустической задачи найдено в виде суммы поверхностного интеграла от функции давления и пульсации скорости и пространственного интеграла от функции распределения квадрупольных источников. При малых числах Маха квадрупольными источниками можно пренебречь, а контрольную поверхность интегрирования необходимо сравнить с поверхностью обтекаемого тела [33]. В этом случае акустическое поле обтекаемого тела определим величиной нестационарных динамических нагрузок, действующих на него со стороны потока. Учет поверхностных интегралов существенно усложняет метод при исследовании шумоизлучения тел сложной пространственной конфигурации и снижает производительность расчетов.

Результаты и их анализ

Произведем расчеты в следующей последовательности. Протестируем модели турбулентности: Спаларта–Аллараса (SA) [34], $k-\omega$ SST Ментора [35] и четырехпараметрическая Transition SST $\gamma-Re_\theta$ модель Ленгтри (tSST) [36]. При помощи выбранной модели исследуем ВКД с X-образными сдвоенными лопастями (рис. 2), определим оптимальные угол $\Delta\psi$ и зазор δ . Оптимальный ВКД ВКЗХ-К6 сравним с остальными ВКД по тяге и потребляемой мощности на валу. Сравнение по уровню акустического шума на местности выполним с ВКД ВК6-К6, который наиболее близок изучаемому ВКЗХ-К6 по тяговым характеристикам.

Тестирование моделей турбулентности осуществим на постепенно сгущающейся сетке до тех пор, пока необходимые измерения не выполнены для всех моделей. Окончательная разностная сетка состоит из 980 475 ячеек. Для модели SST сделаем сравнение двух решателей CFX и Fluent, включенных в пакет ANSYS, которые отличаются тем, что в CFX значения переменных вычисляются в вершинах ячеек, а во Fluent в центрах

ячеек. Тестирование моделей показало, что tSST-модель имеет наилучшее совпадение с результатами эксперимента (рис. 6, а). Видно, что до угла установки лопасти $\varphi_{0,75} = 35^\circ$, когда винт переходит на закритические отрывные режимы, результаты, полученные при помощи tSST, совпадают с экспериментальными. У стандартной SST-модели падение тяги начинается раньше. Причина состоит в том, что при угле $\varphi_{0,75} = 35^\circ$ на лопастях образуется тонкий ламинарный пузырь, который меняет обтекание лопасти таким образом, как будто увеличивается кривизна профиля. Модель tSST благодаря наличию механизма учета ламинарно-турбулентного перехода адекватно воспроизводит это явление. Модель SST считает поток с развитой турбулентностью, поэтому воспроизводить ламинарные отрывные пузыри она не может. Отметим, что для моделирования течения в ВКД на закритических режимах необходимо привлечь вихреразрешающие методы.

На рис. 6, б приведено сравнение результатов, полученных во Fluent и CFX для модели SST. На безотрывных режимах до $\varphi_{0,75} = 35^\circ$ они идентичны. В момент появления ламинарных отрывных пузырей CFX предсказывает более развитое отрывное течение, сопровождающееся сильным падением тяги.

Дальнейшие расчеты выполнены для решателя во Fluent с применением tSST-модели турбулентности. Результаты моделирования показали, что до угла между сдвоенными лопастями $\Delta\psi = 30^\circ$ характер течения меняется слабо, тяга и потребляемая мощность постепенно растут, а параметр «мощность на единицу тяги» (это замена понятия коэффициент полезного действия при работе винта на месте) N/P остается примерно постоянным. Линия стекания эжектируемого потока на внешней поверхности кольца почти вертикальная. За линией стекания вектор скорости направлен по потоку, т. е. эжектируемый поток вносит свой вклад в создание тяги (рис. 7, а). При дальнейшем увеличении $\Delta\psi$ линии тока за линией растекания отклоняются против потока. Образуется отрывной пузырь, в котором поток движется навстречу вектору тяги (рис. 7, б), и только за счет вязкого взаимодействия с выхлопной струей ВКД он разворачивается по потоку, но при этом уменьшается

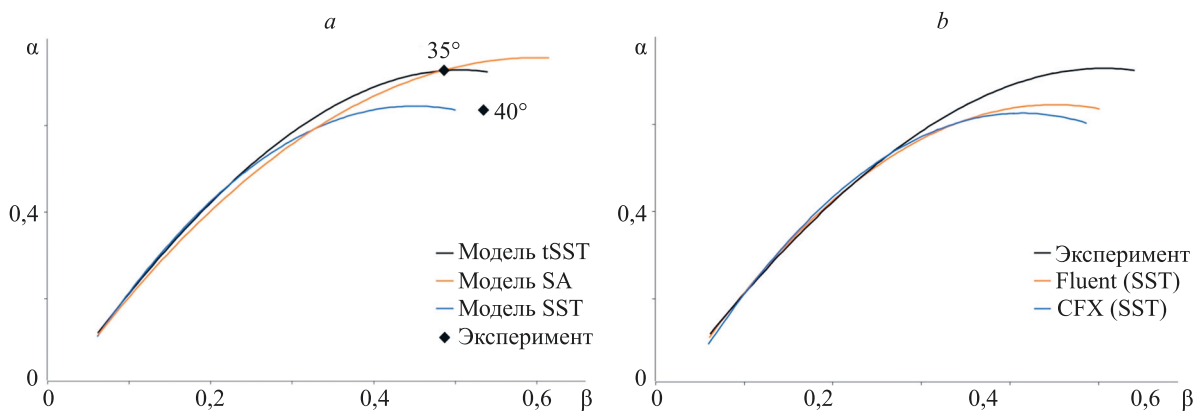


Рис. 6. Сравнения моделей турбулентности с результатами эксперимента (а) и решателей Fluent и CFX (б).

Символом $\blacklozenge$ отмечены углы установки лопасти $\varphi_{0,75}$ равные 35° и 40° , при которых винт работает на закритических режимах

Fig. 6. Comparison of various turbulence models with experimental results (a) and Fluent and CFX solvers (b).

The symbol $\blacklozenge$ marks the blade installation angles $\varphi_{0,75} = 35^\circ$ and 40° at which the fan operates in supercritical modes

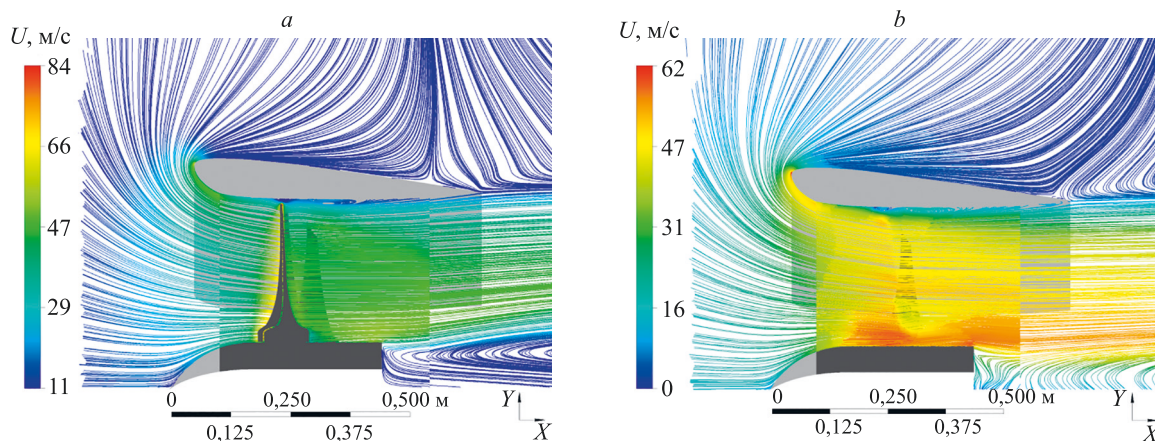


Рис. 7. Сравнение картины течения при угле между лопастями $\Delta\psi = 30^\circ$ (a) и $\Delta\psi = 40^\circ$ (b)
 Fig. 7. Comparison of the flow pattern at an angle between the blades $\Delta\psi = 30^\circ$ (a) and $\Delta\psi = 40^\circ$ (b)

Таблица. Сравнение характеристик ВКД при различных $\Delta\psi$ и δ
 Table. Comparison of ducted fans characteristics of at different $\Delta\psi$ and δ

δ , мм	$\Delta\psi$	Мощность N , кВт	Тяга P , кгс	N/P , кВт/кгс	Мощность N , кВт	Тяга P , кгс	N/P , кВт/кгс
63	20°	21,32	78,68	0,27	15,75	68,26	0,23
	30°	22,51	81,01	0,28		63,91	0,25
	40°	23,07	79,85	0,29		61,89	0,25
91	30°	22,20	79,62	0,28		63,32	0,25

импульс потока за ВКД, следовательно, уменьшается и сила тяги.

После определения оптимального угла $\Delta\psi$ выполнено исследование характеристик ВКД при разных его значениях и граничных расстояниях δ от 63 до 91 мм. Значение $\delta_{\text{опт}} = 91$ мм установлено оптимальным и является минимальным значением N/P .

Картина течения при этом практически не отличается от рис. 7, а. Обобщенные результаты расчетов характеристик для граничных расстояний и при фиксированной мощности двигателя в 15 кВт приведены в таблице. Зависимость тяги от $\Delta\psi$ приведена на рис. 8.

ВКД с трехлопастным винтом уступает при максимальном использовании подведенной мощности всем

другим типам винтов примерно на 20 %. Если же нагрузка мощностью умеренная, то данный ВКД создает несколько большую тягу.

Выполним полученных анализ результатов. Винт с $\Delta\psi = 30^\circ$ соответствует ВКЗ $\times$ 3-К6 (рис. 3, с). На максимальной мощности такой винт продемонстрировал максимальную из всех вариантов тягу. Однако при заданной мощности около 15 кВт, когда потенциал винта использован не полностью и частота вращения меньше 3000 об/мин наибольшую тягу дает винт с $\Delta\psi = 20^\circ$. Наивысшую эффективность N/P имеет также Х-образный винт с $\Delta\psi = 20^\circ$. При максимальной частоте

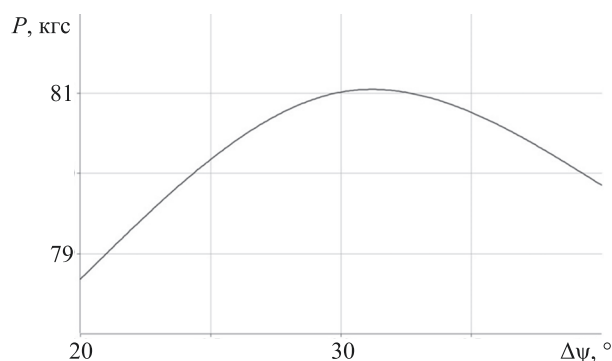


Рис. 8. Зависимость тяги P от угла смещения лопастей винта $\Delta\psi$ при расстоянии между дисками винта $\delta = 63$ мм
 Fig. 8. Dependence of thrust P on the angle of displacement of the fan blades $\Delta\psi$ at a distance between the fan disks $\delta = 63$ mm

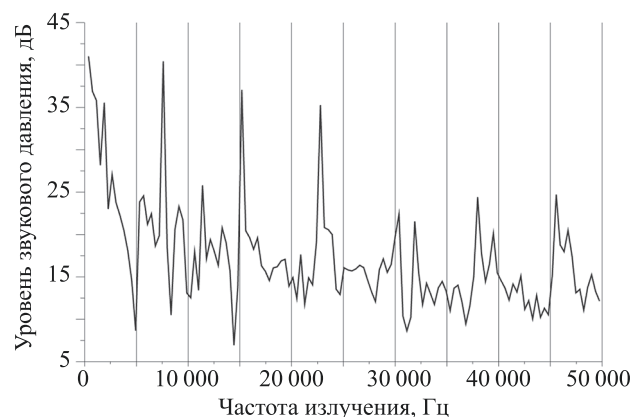


Рис. 9. Спектр шума, создаваемого ВКД ВК6-К6 на оси, на расстоянии 3 м за винтом
 Fig. 9. Noise spectrum generated by ducted fan ВК6-К6 on the axis, at a distance of 3 m behind the fan

те вращения 3000 об/мин получен парадоксальный на первый взгляд результат $P_{\Delta\psi = 40^\circ} > P_{\Delta\psi = 20^\circ}$, который характерен для существенно большей подведенной мощности.

Таким образом, наилучшими характеристиками обладает ВКД ВКЗХ-К6 с $\Delta\psi = 20^\circ$. Оптимальное значение расстояния между дисками $\delta_{\text{опт}} = 91$ мм.

Полученные в процессе нестационарного газодинамического расчета пульсации давления во времени позволили определить акустические характеристики ВКД как в ближнем, так и в дальнем полях. Наиболее важными характеристиками являются: зависимость общего уровня звукового давления от направления на точку наблюдения и спектральный состав акустического сигнала в точках наблюдения.

Акустические характеристики изучены для шестилопастного ВКД ВК6-К6 и ВКЗХ-К6 с $\Delta\psi = 20^\circ$. Интуитивно понятно, что повышение эффективности ВКЗХ-К6 с $\Delta\psi = 20^\circ$ достигается за счет уменьшения индуктивного сопротивления, соответственно, снижается интенсивность концевых вихрей, что должно сопровождаться снижением уровня шума. Измерения проведены на расстоянии 4 м в плоскости вращения винта и на расстоянии 3 м за винтом на оси его вращения. Для ВКЗХ-К6 с $\Delta\psi = 20^\circ$ интегральный уровень шума составил 82 дБ в плоскости вращения и 78 дБ на оси вращения. Для ВК6-К6 — 100,1 дБ и 89,2 дБ соответственно. Отметим, что у обычного винта присутствуют выраженные частоты дискретного тона, причем в высокочастотной области (рис. 9). У ВКЗХ-К6 с $\Delta\psi = 20^\circ$ спектр излучения равномерный, пик смещен в область низкочастотного излучения с частотой $f = 50\text{--}200$ Гц.

Заключение

Разработан метод численного моделирования винтокольцевых движителей. Исследованы модели турбулентности Спаларта-Аллмараса (SA), $k\text{--}\omega$ SST (Shear Stress Transport) модели турбулентности Ментора и четырехпараметрическая Transition SST $\gamma\text{--Re}_\theta$ модель Ленгтри (tsST). Показано, что наилучшие результаты в моделировании потока внутри кольца продемонстрировала модель Transition SST, что связано со способностью данной модели воспроизводить тонкие ламинарные отрывные пузыри. Выполнено численное исследование винтокольцевых движителей с винтами нескольких типов: трехлопастным, шестилопастным и двумя трехлопастными винтами, вращающимися в одну сторону, лопасти которых смещены на различные углы. В результате получено, что двухрядный винт с расстоянием между дисками вращения 91 мм и углом смещения лопастей 20° показал наивысшую эффективность работы за счет наименьшего уровня индуктивного сопротивления. Этот же винт является самым маломощным, кроме того, в спектре создаваемого им акустического излучения отсутствуют высокочастотные гармоники дискретного тона, а максимум спектра смещен в низкочастотную область, что субъективно ощущается более благоприятно. Преимущество в тяге и эффективности проявилось при максимальной нагрузке винта мощностью. При средней нагрузке наилучшие характеристики имеет обычный трехлопастный винт. При этом снижение шума при Х-образной компоновке лопастей наблюдается при любой подведенной мощности.

Таким образом, винты и вентиляторы с неравномерным размещением лопастей по азимуту должны исследоваться активнее.

Литература

1. Chovancová A., Fico T., Chovanec L., Hubinský P. Mathematical modelling and parameter identification of quadrotor (a survey) // *Procedia Engineering*. 2014. V. 96. P. 172–181. <https://doi.org/10.1016/j.proeng.2014.12.139>
2. Остроухов С.П. Аэродинамика воздушных винтов и винтокольцевых движителей. М.: Физматлит, 2014. 328 с.
3. Бузыкин О.Г., Казаков А.В., Шустов А.В. Численное моделирование аэродинамических характеристик малоразмерного летательного аппарата // *Ученые записки ЦАГИ*. 2010. Т. 41. № 5. С. 21–31.
4. Назаров Д.В., Кондрякова А.В. Исследование обтекания винта с применением численных и экспериментальных методов // *Известия Самарского научного центра Российской академии наук*. 2018. Т. 20. № 4-1. С. 70–75.
5. Rumsey C.L., Biedron R., Farassat F., Spence P. Ducted-fan engine acoustic predictions using a Navier-Stokes code // *Journal of Sound and Vibration*. 1998. V. 213. N 4. P. 643–664. <https://doi.org/10.1006/jsvi.1998.1519>
6. Reboul G., Polacsek C., Lewy S., Heib S. Aeroacoustic computation of ducted-fan broadband noise using LES data // *Journal of the Acoustical Society of America*. 2008. V. 123. N 5. P. 3539. <https://doi.org/10.1121/1.2934519>
7. Myers L., Rhee W., McLaughlin D. Aeroacoustics of vertical lift ducted rotors // *Proc. of the 15th AIAA/CEAS Aeroacoustics Conference (30th AIAA Aeroacoustics Conference)*. 2009. P. 2009-3333. <https://doi.org/10.2514/6.2009-3333>
8. Astley R., Sugimoto R., Achunche I., Kewin M., Mustafi P., Deane E. A review of CAA for fan duct propagation and radiation, with application to liner optimisation // *Procedia Engineering*. 2010. V. 6. P. 143–152. <https://doi.org/10.1016/j.proeng.2010.09.016>

References

1. Chovancová A., Fico T., Chovanec L., Hubinský P. Mathematical modelling and parameter identification of quadrotor (a survey). *Procedia Engineering*, 2014, vol. 96, pp. 172–181. <https://doi.org/10.1016/j.proeng.2014.12.139>
2. Ostroukhov S.P. *Aerodynamics of Aircraft Propellers and Annular Circular Propellers*. Moscow, Fizmatlit, 2014, 328 p. (in Russian)
3. Buzikin O.G., Kazakov A.V., Shustov A.V. Numerical simulation of the aerodynamic performance of a micro air vehicle. *TSAGI Science Journal*. 2010, vol. 41, no. 5, pp. 535–550. <https://doi.org/10.1615/TsAGISciJ.v41.i5.40>
4. Nazarov D., Kondryakova A. Analysis of the flow past the screw with the use of numerical and experimental simulation. *Izvestia of Samara Scientific Center of the Russian Academy of Sciences*, 2018, vol. 20, no. 4-1, pp. 70–75. (in Russian)
5. Rumsey C.L., Biedron R., Farassat F., Spence P. Ducted-fan engine acoustic predictions using a Navier-Stokes code. *Journal of Sound and Vibration*, 1998, vol. 213, no. 4, pp. 643–664. <https://doi.org/10.1006/jsvi.1998.1519>
6. Reboul G., Polacsek C., Lewy S., Heib S. Aeroacoustic computation of ducted-fan broadband noise using LES data. *Journal of the Acoustical Society of America*, 2008, vol. 123, no. 5, pp. 3539. <https://doi.org/10.1121/1.2934519>
7. Myers L., Rhee W., McLaughlin D. Aeroacoustics of vertical lift ducted rotors. *Proc. of the 15th AIAA/CEAS Aeroacoustics Conference (30th AIAA Aeroacoustics Conference)*, 2009, pp. 2009-3333. <https://doi.org/10.2514/6.2009-3333>
8. Astley R., Sugimoto R., Achunche I., Kewin M., Mustafi P., Deane E. A review of CAA for fan duct propagation and radiation, with application to liner optimization. *Procedia Engineering*, 2010, vol. 6, pp. 143–152. <https://doi.org/10.1016/j.proeng.2010.09.016>

9. Malgoezar A.M., Vieira A., Snellen M., Simons D.G., Veldhuis L.L. Experimental characterization of noise radiation from a ducted propeller of an unmanned aerial vehicle // *International Journal of Aeroacoustics*. 2019. V. 18. N 4-5. P. 372–391 <https://doi.org/10.1177/1475472X19852952>
10. Zhang T., Barakos G.N. Review on ducted fans for compound rotorcraft // *The Aeronautical Journal*. 2020. V. 124. N 1277. P. 941–974. <https://doi.org/10.1017/aer.2019.164>
11. Akturk A., Shavalikul A., Camci C. PIV measurements and computational study of a 5-inch ducted fan for V/STOL UAV applications // *Proc. of the 47th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*. 2009. P. 2009-332. <https://doi.org/10.2514/6.2009-332>
12. Akturk A., Camci C. Experimental and computational assessment of a ducted-fan rotor flow model // *Journal of Aircraft*. 2012. V. 49. N 3. P. 885–897. <https://doi.org/10.2514/1.C031562>
13. Yilmaz S., Erdem D., Kavsaoglu M. Effects of duct shape on a ducted propeller performance // *Proc. of the 51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*. 2013. P. 2013–0803. <https://doi.org/10.2514/6.2013-803>
14. Akturk A., Camci C. Tip clearance investigation of a ducted fan used in VTOL unmanned aerial vehicles. Part I: Baseline experiments and computational validation // *Journal of Turbomachinery*. 2014. V. 136. N 2. P. 021004. <https://doi.org/10.1115/1.4023468>
15. Xu H.-Y., Xing S.-L., Ye Z.-Y. Numerical study of ducted-fan lip stall suppression based on inflatable leading lip cell // *Procedia Engineering*. 2015. V. 126. P. 158–162. <https://doi.org/10.1016/j.proeng.2015.11.202>
16. Biava M., Barakos G.N. Optimisation of ducted propellers for hybrid air vehicles using high-fidelity CFD // *The Aeronautical Journal*. 2016. V. 120. N 1232. P. 1632–1657. <https://doi.org/10.1017/aer.2016.78>
17. Chen J., Li L., Huang G., Xiang X. Numerical investigations of ducted fan aerodynamic performance with tip-jet // *Aerospace Science and Technology*. 2018. V. 78. P. 510–521. <https://doi.org/10.1016/j.ast.2018.05.016>
18. Ohanian O.J., Karni E.D., Londenberg W.K., Gelhausen P.A., Inman D.J. Ducted-fan force and moment control via steady and synthetic jets // *Journal of Aircraft*. 2011. V. 48. N 2. P. 514–526. <https://doi.org/10.2514/1.C031110>
19. Мойзын Е.И., Завалов О.А., Кузнецов А.В. Экспериментальные исследования аэродинамических характеристик дистанционно-пилотируемого летательного аппарата с несущей системой «винт в кольце» // *Труды МАИ*. 2012. № 50. С. 11.
20. Dehaeze F., Barakos G.N. Hovering rotor computations using an aeroelastic blade model // *The Aeronautical Journal*. 2012. V. 116. N 1180. P. 621–649. <https://doi.org/10.1017/S0001924000007107>
21. Абалакин И.В., Бахвалов П.А., Бобков В.Г., Козубская Т.К., Аникин В.А. Численное моделирование аэродинамических и акустических характеристик винта в кольце // *Математическое моделирование*. 2015. Т. 27. № 10. С. 125–144.
22. Абалакин И.В., Аникин В.А., Бахвалов П.А., Бобков В.Г., Козубская Т.К. Численное исследование аэродинамических и акустических свойств винта в кольце // *Известия Российской академии наук. Механика жидкости и газа*. 2016. № 3. С. 130–145. <https://doi.org/10.7868/S0568528116030026>
23. Диизи Ф., Баракос Дж., Кусюмов А.Н., Кусюмов С.А., Михайлов С.А. DES-моделирование обтекания несущего винта вертолета // *Известия вузов. Авиационная техника*. 2018. № 1. С. 40–46.
24. Головкин М.А., Кочиш С.И., Крицкий Б.С. Методика расчета аэродинамических характеристик комбинированной несущей системы летательного аппарата // *Труды МАИ*. 2012. № 55. С. 5.
25. Копьев В.Ф., Титарев В.А., Беляев И.В. Разработка методологии расчета шума винтов с использованием суперкомпьютеров // *Ученые записки ЦАГИ*. 2014. Т. 45. № 2. С. 78–106.
26. Costes M., Renaud T., Rodriguez B. Rotorcraft simulations: a challenge for CFD // *International Journal of Computational Fluid Dynamics*. 2012. V. 26. N 6-8. P. 383–405. <https://doi.org/10.1080/10618562.2012.726710>
27. Costes M., Renaud T., Rodriguez B., Reboul G. Application of vorticity confinement to rotor wake simulations // *International Journal of Engineering Systems, Modelling and Simulation*. 2012. V. 4. N 1-2. P. 102–112. <https://doi.org/10.1504/ijesms.2012.044848>
28. Попов Н.И., Емельянова О.В., Яцун С.Ф. Моделирование динамики полета квадрокоптера // *Вестник Воронежского института ГПС МЧС России*. 2014. № 4(13). С. 69–75.
9. Malgoezar A.M., Vieira A., Snellen M., Simons D.G., Veldhuis L.L. Experimental characterization of noise radiation from a ducted propeller of an unmanned aerial vehicle. *International Journal of Aeroacoustics*, 2019, vol. 18, no. 4-5, pp. 372–391 <https://doi.org/10.1177/1475472X19852952>
10. Zhang T., Barakos G.N. Review on ducted fans for compound rotorcraft. *The Aeronautical Journal*, 2020, vol. 124, no. 1277, pp. 941–974. <https://doi.org/10.1017/aer.2019.164>
11. Akturk A., Shavalikul A., Camci C. PIV measurements and computational study of a 5-inch ducted fan for V/STOL UAV applications. *Proc. of the 47th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2009, pp. 2009-332. <https://doi.org/10.2514/6.2009-332>
12. Akturk A., Camci C. Experimental and computational assessment of a ducted-fan rotor flow model. *Journal of Aircraft*, 2012, vol. 49, no. 3, pp. 885–897. <https://doi.org/10.2514/1.C031562>
13. Yilmaz S., Erdem D., Kavsaoglu M. Effects of duct shape on a ducted propeller performance. *Proc. of the 51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition*, 2013, pp. 2013–0803. <https://doi.org/10.2514/6.2013-803>
14. Akturk A., Camci C. Tip clearance investigation of a ducted fan used in VTOL unmanned aerial vehicles. Part I: Baseline experiments and computational validation. *Journal of Turbomachinery*, 2014, vol. 136, no. 2, pp. 021004. <https://doi.org/10.1115/1.4023468>
15. Xu H.-Y., Xing S.-L., Ye Z.-Y. Numerical study of ducted-fan lip stall suppression based on inflatable leading lip cell. *Procedia Engineering*, 2015, vol. 126, pp. 158–162. <https://doi.org/10.1016/j.proeng.2015.11.202>
16. Biava M., Barakos G.N. Optimisation of ducted propellers for hybrid air vehicles using high-fidelity CFD. *The Aeronautical Journal*, 2016, vol. 120, no. 1232, pp. 1632–1657. <https://doi.org/10.1017/aer.2016.78>
17. Chen J., Li L., Huang G., Xiang X. Numerical investigations of ducted fan aerodynamic performance with tip-jet. *Aerospace Science and Technology*, 2018, vol. 78, pp. 510–521. <https://doi.org/10.1016/j.ast.2018.05.016>
18. Ohanian O.J., Karni E.D., Londenberg W.K., Gelhausen P.A., Inman D.J. Ducted-fan force and moment control via steady and synthetic jets. *Journal of Aircraft*, 2011, vol. 48, no. 2, pp. 514–526. <https://doi.org/10.2514/1.C031110>
19. Mozyzh E.I., Zavalov O.A., Kuznetsov A.V. Experimental investigation of aerodynamic characteristics of unmanned aerial vehicle with lifted system of shrouded rotor. *Trudy MAI*, 2012, no. 50, pp. 11. (in Russian)
20. Dehaeze F., Barakos G.N. Hovering rotor computations using an aeroelastic blade model. *The Aeronautical Journal*, 2012, vol. 116, no. 1180, pp. 621–649. <https://doi.org/10.1017/S0001924000007107>
21. Abalakin I.V., Bakhvalov P.A., Bobkov V.G., Kozubskaya T.K., Anikin V.A. Numerical simulation of aerodynamic and acoustic characteristics of rotor in ring. *Matematicheskoe modelirovanie*, 2015, vol. 27, no. 10, pp. 125–144. (in Russian)
22. Abalakin I.V., Anikin V.A., Bakhvalov P.A., Bobkov V.G., Kozubskaya T.K. Numerical investigation of the aerodynamic and acoustical properties of a shrouded rotor. *Fluid Dynamics*, 2016, vol. 51, no. 3, pp. 419–433. <https://doi.org/10.1134/S0015462816030145>
23. Dehaeze F., Barakos G.N., Kusunyov A.N., Kusunyov S.A., Mikhailov S.A. Exploring the detached-eddy simulation for main rotor flows. *Russian Aeronautics*, 2018, vol. 61, no. 1, pp. 37–44. <https://doi.org/10.3103/S1068799818010063>
24. Golovkin M.A., Kochish S.I., Kritsky B.S. Calculation procedure of aerodynamic characteristics of the combined carrying system of the aircraft. *Trudy MAI*, 2012, no. 55, pp. 5. (in Russian)
25. Kopiev V.F., Titarev V.A., Belyaev I.V. Development of a methodology for propeller noise calculation on high-performance computer. *TSAGI Science Journal*, 2014, vol. 45, no. 3-4, pp. 293–327. <https://doi.org/10.1615/TSAGISciJ.2014011857>
26. Costes M., Renaud T., Rodriguez B. Rotorcraft simulations: a challenge for CFD. *International Journal of Computational Fluid Dynamics*, 2012, vol. 26, no. 6-8, pp. 383–405. <https://doi.org/10.1080/10618562.2012.726710>
27. Costes M., Renaud T., Rodriguez B., Reboul G. Application of vorticity confinement to rotor wake simulations. *International Journal of Engineering Systems, Modelling and Simulation*, 2012, vol. 4, no. 1-2, pp. 102–112. <https://doi.org/10.1504/ijesms.2012.044848>

29. Канатников А.Н., Акопян К.Р. Управление плоским движением квадрокоптера // Математика и математическое моделирование. 2015. № 2. С. 23–36. <https://doi.org/10.7463/mathm.0215.0789477>
30. Шайдаков В.И. Влияние близости земли на аэродинамические характеристики летательного аппарата с несущей системой «винт в кольце» // Труды МАИ. 2011. № 49. С. 24.
31. 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.
32. Volkov K. Multigrid and preconditioning techniques in CFD applications // CFD Techniques and Thermo-Mechanics Applications. Springer International Publishing, 2018. P. 83–149. https://doi.org/10.1007/978-3-319-70945-1_6
33. Суворов А.С., Коротин П.И., Соков Е.М. Метод конечно-элементного моделирования шумоизлучения, генерируемого неоднородностями тел, движущихся в турбулентном потоке жидкости // Акустический журнал. 2018. Т. 64. № 6. С. 756–767. <https://doi.org/10.1134/S0320791918060114>
34. Spalart P.R., Allmaras S.R. A one-equation turbulence model for aerodynamic flows // Proc. of the 30th Aerospace Sciences Meeting and Exhibit. 1992. P. 1992-0439. <https://doi.org/10.2514/6.1992-439>
35. Menter F.R. Zonal two-equation k- ω turbulence models for aerodynamic flows // Proc. of the 23rd Fluid Dynamics, Plasmadynamics, and Lasers Conference. 1993. P. 1993-2906. <https://doi.org/10.2514/6.1993-2906>
36. Menter F.R., Langtry R., Volker S. Transition modelling for general purpose CFD codes // Flow, Turbulence and Combustion. 2006. V. 77. N 1-4. P. 277–303. <https://doi.org/10.1007/s10494-006-9047-1>
28. Popov N.I., Emelianova O.V., Jatsun S.F. Modelling of dynamics of flight of a quadrotor helicopter. *Vestnik Voronezhskogo instituta GPS MChS Rossii*, 2014, no. 4(13), pp. 69–75.
29. Kanatnikov A. N., Akopyan K. R. The plane motion control of the quadcopter. *Mathematics and Mathematical Modelling*, 2015, no. 2, pp. 23–36. (in Russian). <https://doi.org/10.7463/mathm.0215.0789477>
30. Shydakov V.I. Ground effect on aerodynamic characteristics of aerial vehicle with lifted system of shrouded rotor. *Trudy MAI*, 2011, no. 49, pp. 24. (in Russian)
31. 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.
32. Volkov K. Multigrid and preconditioning techniques in CFD applications. *CFD Techniques and Thermo-Mechanics Applications*. Springer International Publishing, 2018, pp. 83–149. https://doi.org/10.1007/978-3-319-70945-1_6
33. Suvorov A.S., Korotin P.I., Sokov E.M. Finite element method for simulating noise emission generated by inhomogeneities of bodies moving in a turbulent fluid flow. *Acoustical Physics*, 2018, vol. 64, no. 6, pp. 778–788. <https://doi.org/10.1134/S1063771018060106>
34. Spalart P.R., Allmaras S.R. A one-equation turbulence model for aerodynamic flows. *Proc. of the 30th Aerospace Sciences Meeting and Exhibit*, 1992, pp. 1992-0439. <https://doi.org/10.2514/6.1992-439>
35. Menter F.R. Zonal two-equation k- ω turbulence models for aerodynamic flows. *Proc. of the 23rd Fluid Dynamics, Plasmadynamics, and Lasers Conference*, 1993, pp. 1993-2906. <https://doi.org/10.2514/6.1993-2906>
36. Menter F.R., Langtry R., Volker S. Transition modelling for general purpose CFD codes. *Flow, Turbulence and Combustion*, 2006, vol. 77, no. 1-4, pp. 277–303. <https://doi.org/10.1007/s10494-006-9047-1>

Авторы

Дудников Сергей Юрьевич — кандидат физико-математических наук, директор Института национальной технологической инициативы, Севастопольский государственный университет, Севастополь, 299053, Российская Федерация, [sc 23992474600](https://orcid.org/0000-0003-1955-4313), <https://orcid.org/0000-0003-1955-4313>, syududnikov@sevsu.ru

Булат Михаил Павлович — кандидат технических наук, старший научный сотрудник, Севастопольский государственный университет, Севастополь, 299053, Российская Федерация; старший научный сотрудник, Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация, [sc 55973496000](https://orcid.org/0000-0001-9727-673X), <https://orcid.org/0000-0001-9727-673X>, bulat_mh@mail.ru

Вокни Леонид Олегович — младший научный сотрудник, Севастопольский государственный университет, Севастополь, 299053, Российская Федерация; младший научный сотрудник, Балтийский государственный технический университет «ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация, [sc 57222385381](https://orcid.org/0000-0003-2944-1822), <https://orcid.org/0000-0003-2944-1822>, leonidvokin@mail.ru

Кузнецов Павел Николаевич — кандидат технических наук, старший научный сотрудник, Севастопольский государственный университет, Севастополь, 299053, Российская Федерация, [sc 57206301213](https://orcid.org/0000-0002-1732-922X), <https://orcid.org/0000-0002-1732-922X>, PNKuznetsov@sevsu.ru

Чернышов Павел Сергеевич — младший научный сотрудник, «Балтийский государственный технический университет ВОЕНМЕХ» им. Д.Ф. Устинова, Санкт-Петербург, 190005, Российская Федерация, [sc 57215086432](https://orcid.org/0000-0003-1433-4194), <https://orcid.org/0000-0003-1433-4194>, pashachp8@gmail.com

Authors

Sergey Yu. Dudnikov — PhD (Physics & Mathematics), Director of the National Technology Initiative Institute, Sevastopol State University, Sevastopol, 299053, Russian Federation, [sc 23992474600](https://orcid.org/0000-0003-1955-4313), <https://orcid.org/0000-0003-1955-4313>, syududnikov@sevsu.ru

Mikhail P. Bulat — PhD, Senior Researcher, Sevastopol State University, Sevastopol, 299053, Russian Federation; Senior Researcher, Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation, [sc 55973496000](https://orcid.org/0000-0001-9727-673X), <https://orcid.org/0000-0001-9727-673X>, bulat_mh@mail.ru

Leonid O. Vokin — Junior Researcher, Sevastopol State University, Sevastopol, 299053, Russian Federation; Junior Researcher, Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation, [sc 57222385381](https://orcid.org/0000-0003-2944-1822), <https://orcid.org/0000-0003-2944-1822>, leonidvokin@mail.ru

Pavel N. Kuznetsov — PhD, Senior Researcher, Sevastopol State University, Sevastopol, 299053, Russian Federation, [sc 57206301213](https://orcid.org/0000-0002-1732-922X), <https://orcid.org/0000-0002-1732-922X>, PNKuznetsov@sevsu.ru

Pavel S. Chernyshov — Junior Researcher, Baltic State Technical University “VOENMEH” named after D.F. Ustinov, Saint Petersburg, 190005, Russian Federation, [sc 57215086432](https://orcid.org/0000-0003-1433-4194), <https://orcid.org/0000-0003-1433-4194>, pashachp8@gmail.com

Статья поступила в редакцию 07.09.2022

Одобрена после рецензирования 27.10.2022

Принята к печати 29.11.2022

Received 07.09.2022

Approved after reviewing 27.10.2022

Accepted 29.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

КРАТКИЕ СООБЩЕНИЯ

BRIEF PAPERS

doi: 10.17586/2226-1494-2022-22-6-1237-1240

УДК 536.6

**Дифференциально-разностная модель теплопереноса в твердых телах
с использованием метода параметрической идентификации**

**Николай Васильевич Пилипенко¹, Павел Андреевич Колодийчук²✉,
Юрий Петрович Заричняк³**

^{1,2,3} Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

¹ pilipenko38@mail.ru, <https://orcid.org/0000-0001-9328-3166>

² pkolodiychuk@itmo.ru✉, <https://orcid.org/0000-0002-9984-1613>

³ zarich4@gmail.com, <https://orcid.org/0000-0001-8713-3583>

Аннотация

Рассмотрена задача параметрической идентификации дифференциально-разностной модели процесса теплопереноса в сферическом теле. При разработке модели использован оригинальный расширенный фильтр Калмана, который позволяет учесть зависимость теплофизических свойств исследуемого объекта от температуры. Постановка задачи и полученное решение дают возможность учитывать различный характер внешнего теплового воздействия и процессы, происходящие внутри тел, в особенности при фазовых переходах в системах тел. Продemonстрировано хорошее совпадение результатов параметрической идентификации и численного моделирования с использованием программного пакета «Ansys». Показано, что предложенный в работе метод позволяет определять не только температуру в разных точках объекта, но и восстанавливать нестационарный тепловой поток на границе объекта, а также уточнять его теплофизические свойства. Рассмотренный метод параметрической идентификации дифференциально-разностной модели теплопереноса может найти применение при определении эффективности аккумуляторов тепловой энергии.

Ключевые слова

дифференциально-разностная модель, сферическое тело, сферическая симметрия, фильтр Калмана, параметрическая идентификация, теплотметрия

Ссылка для цитирования: Пилипенко Н.В., Колодийчук П.А., Заричняк Ю.П. Дифференциально-разностная модель теплопереноса в твердых телах с использованием метода параметрической идентификации // Научно-технический вестник информационных технологий, механики и оптики. 2022. Т. 22, № 6. С. 1237–1240. doi: 10.17586/2226-1494-2022-22-6-1237-1240

**Differential-difference model of heat transfer in solids using the method
of parametric identification**

Nikolai V. Pilipenko¹, Pavel A. Kolodiychuk²✉, Yuri P. Zarichnyak³

^{1,2,3} ITMO University, Saint Petersburg, 197101, Russian Federation

¹ pilipenko38@mail.ru, <https://orcid.org/0000-0001-9328-3166>

² pkolodiychuk@itmo.ru✉, <https://orcid.org/0000-0002-9984-1613>

³ zarich4@gmail.com, <https://orcid.org/0000-0001-8713-3583>

Abstract

The paper considers the problem of parametric identification of a differential-difference model of the heat transfer process in a spherical body. When developing the model, the original extended Kalman filter is used which allows taking into account the dependence of the thermophysical properties of the object under study on temperature. This formulation and the obtained solution of the problem make it possible to take into account the different nature of the external thermal effect and the processes occurring inside the bodies, in particular, during phase transitions in systems of bodies. The research results obtained using parametric identification and Ansys software are in good agreement. However, the method

we have considered, in contrast to the Ansys software, allows not only to determine the temperature at different points of the object, but also to restore the non-stationary heat flow at the object boundary as well as to refine its thermophysical properties. The considered method of parametric identification of the differential-difference model of heat transfer can be successfully used in determining the efficiency of heat energy storage devices.

Keywords

differential-difference model, spherical body, spherical symmetry, Kalman filter, parametric identification, heat measurement

For citation: Pilipenko N.V., Kolodiyuk P.A., Zarichnyak Yu.P. Differential-difference model of heat transfer in solids using the method of parametric identification. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2022, vol. 22, no. 6, pp. 1237–1240 (in Russian). doi: 10.17586/2226-1494-2022-22-6-1237-1240

В настоящее время актуальным является решение прямых и обратных задач теплопроводности при известных граничных условиях, в особенности если внутри тела или системы тел происходят фазовые переходы. Существуют различные методы восстановления нестационарной температуры при решении прямых задач теплопроводности и нестационарного теплового потока обратных задач теплопроводности. В работе [1] приведено описание метода параметрической идентификации дифференциально-разностных моделей теплопереноса на основе рекуррентных алгоритмов фильтра Калмана, который позволил решить ряд сложных нестационарных задач [2–8]. Алгоритмы программ для решения ряда прямых и обратных задач теплопроводности при использовании дифференциально-разностных моделей и неопределенности, полученные при решении различных задач, приведены в работах [2–8].

Представляет интерес использование подхода [2–8] в аккумуляторах тепловой энергии [9, 10] с целью увеличения эффективности действующего объекта. В отличие от рассмотренных в [2–8] задач, где выполнен анализ тела или системы тел в виде плоских стенок с расположенными на их поверхностях преобразователями теплового потока, в работах [9, 10] аккумулятор имеет сферическую форму со специфическими процессами теплопереноса.

Отметим, что в настоящей работе использован оригинальный расширенный фильтр Калмана, который позволяет учесть нелинейность задачи, а именно зависимость теплофизических свойств от температуры.

Цель работы — получение дифференциально-разностной модели теплопереноса сферического тела и определение адекватности разработанной модели реальному процессу теплообмена в процессе решения прямой задачи теплопроводности с использованием расширенного фильтра Калмана и с помощью программного обеспечения (ПО) «Ansys».

В качестве физической модели выбран стальной шарик диаметром 30 мм, находящийся в текучей среде, температура которой меняется по гармоническому закону $t_{cp} = 100\sin(5\tau)$. Гидродинамические параметры среды таковы, что коэффициент теплоотдачи к поверхности шарика постоянен и равен $\alpha = 100 \text{ Вт}/(\text{м}^2 \cdot \text{К})$. Требуется установить динамику температуры центральной части исследуемого тела.

Решение задач выполнено на основании алгоритма, рассмотренного в работах [1, 3]: объект условно разбит на n блоков (например $n = 20$), и для каждого составлено уравнение теплового баланса.

После проведения ряда преобразований составлена дифференциально-разностная модель процесса теплопереноса от первого до последнего блока с указанием граничного условия на первом блоке.

В общем случае дифференциально-разностная модель имеет вид [1]

$$\frac{d}{d\tau} \mathbf{T}(\tau) = \mathbf{F}(\tau) \mathbf{T}(\tau) + \mathbf{G}(\tau) \mathbf{U}(\tau),$$

где $\mathbf{F}(\tau)$ и $\mathbf{G}(\tau)$ — матрицы обратных связей и управления; $\mathbf{U}(\tau)$ и $\mathbf{T}(\tau)$ — вектора управления и состояния; τ — время.

Приведем выражения для расчета скоростей изменения температуры на первом (внешнем), втором, девятнадцатом и двадцатом (центральном) блоках исследуемого тела в связи с наличием ряда особенностей при их выводе. Выражения для скоростей изменения температуры блоков 3–19 идентичны второму, поэтому выражения 3–18 не приводятся.

$$\begin{aligned} \frac{dt_1}{d\tau} &= \frac{3q(\tau)S}{4c\rho\pi(R_2^3 - R_1^3)} - \frac{3aS_1t_2}{(R_1^{-1} - R_2^{-1})(R_2^3 - R_1^3)} + \\ &+ \frac{3aS_1t_1}{(R_1^{-1} - R_2^{-1})(R_2^3 - R_1^3)}, \\ \frac{dt_2}{d\tau} &= \frac{3aS_2t_3}{(R_2^{-1} - R_3^{-1})(R_3^3 - R_2^3)} - \\ &- \left(\frac{S_2}{(R_2^{-1} - R_3^{-1})} - \frac{S_1}{(R_1^{-1} - R_2^{-1})} \right) 3at_2 + \frac{3aS_1t_1}{(R_1^{-1} - R_2^{-1})(R_3^3 - R_2^3)}, \\ \frac{dt_{19}}{d\tau} &= \frac{3aS_{19}t_{20}}{(R_{19}^{-1} - R_{20}^{-1})(R_{20}^3 - R_{19}^3)} - \\ &- \left(\frac{S_{19}}{(R_{19}^{-1} - R_{20}^{-1})} - \frac{S_{18}}{(R_{18}^{-1} - R_{19}^{-1})} \right) 3at_{19} + \\ &+ \frac{3aS_{18}t_{18}}{(R_{18}^{-1} - R_{19}^{-1})(R_{20}^3 - R_{19}^3)}, \\ \frac{dt_{20}}{d\tau} &= - \frac{3aS_{19}t_{20}}{(R_{19}^{-1} - R_{20}^{-1})(R_{20}^3 - R_{19}^3)} + \\ &+ \frac{3aS_{19}t_{19}}{(R_{19}^{-1} - R_{20}^{-1})(R_{20}^3 - R_{19}^3)}, \end{aligned}$$

где $q(\tau)$ — поступающий на поверхность тела нестационарный тепловой поток; τ — время; $\alpha = \frac{\lambda}{c\rho}$ — темпе-

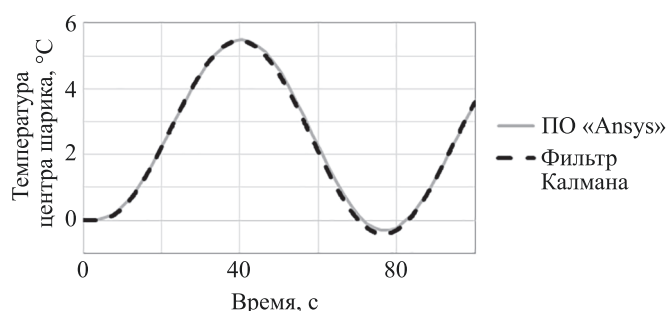


Рисунок. Результаты решения прямой задачи теплопроводности
Figure. The results of solving the direct problem of heat conduction

ратуропроводность; i — номер блока; t_i — температура i -го блока; R_i — внешний радиус i -го блока; c — удельная теплоемкость; ρ — плотность; λ — теплопроводность; S_i — площадь внешней поверхности i -го блока.

На рисунке представлены результаты расчета температуры центра модельного шара при выбранном законе изменения температуры окружающей среды. Расчеты выполнены с использованием метода параметрической идентификации дифференциально-разностных моделей процесса теплопереноса и численного метода ПО «Ansys».

Результаты, полученные с помощью метода параметрической идентификации дифференциально-разностных моделей процесса теплопереноса и ПО «Ansys», совпадают при решении прямой задачи, что говорит об адекватности выбранной модели теплопереноса. Отметим, что рассмотренный метод параметрической идентификации, в отличие от численного ПО, позволил не только определить температуру в разных точках объекта, но и восстановить значение нестационарного теплового потока на границе объекта и уточнить его теплофизические свойства.

Литература

1. Пилипенко Н.В., Заричняк Ю.П., Иванов В.А., Халявин А.М. Параметрическая идентификация дифференциально-разностных моделей теплопереноса в одномерных телах на основе алгоритмов фильтра Калмана // Научно-технический вестник информационных технологий, механики и оптики. 2020. Т. 20. № 4. С. 584–588. <https://doi.org/10.17586/2226-1494-2020-20-4-584-588>
2. Кириллов К.В., Пилипенко Н.В. Алгоритмы программ для решения прямых и обратных задач теплопроводности при использовании дифференциально-разностных моделей // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2010. № 5(69). С. 106–110.
3. Пилипенко Н.В. Неопределенность восстановления нестационарного теплового потока путем параметрической идентификации дифференциально-разностных моделей теплопереноса // Известия высших учебных заведений. Приборостроение. 2017. Т. 60. № 7. С. 664–671. <https://doi.org/10.17586/0021-3454-2017-60-7-664-671>
4. Pilipenko N.V., Gladskih D.A. Determination of the heat losses of buildings and structures by solving inverse heat conduction problems // Measurement Techniques. 2014. V. 57. N 2. P. 181–186. <https://doi.org/10.1007/s11018-014-0427-y>
5. Сиваков И.А., Пилипенко Н.В. Применение фильтра Калмана при восстановлении плотности теплового потока на поверхности объекта исследования в импульсной аэродинамической трубе // Сборник докладов IV Всероссийской научно-практической конференции с международным участием «Информационные технологии в науке, образовании и экономике». 2012. С. 55–58.
6. Пилипенко Н.В. Динамические характеристики различных типов приемников тепловых потоков на основе дифференциально-разностных моделей теплопереноса // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2009. № 3(61). С. 52–58.
7. Пилипенко Н.В. Неопределенность измерения нестационарной температуры поверхности массивных тел // Известия высших учебных заведений. Приборостроение. 2016. Т. 59. № 9. С. 767–772. <https://doi.org/10.17586/0021-3454-2016-59-9-767-772>
8. Пилипенко Н.В. Использование расширенного фильтра Калмана в нестационарной теплотометрии при решении обратных задач

References

1. Pilipenko N.V., Zarichnyak Yu.P., Ivanov V.A., Khalyavin A.M. Parametric identification of differential-difference models of heat transfer in one-dimensional bodies based on Kalman filter algorithms. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2020, vol. 20, no. 4, pp. 584–588. (in Russian). <https://doi.org/10.17586/2226-1494-2020-20-4-584-588>
2. Kirillov K., Pilipenko N. Solution algorithms for direct and backward heat conductivity problems by means of differential-difference models. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2010, no. 5(69), pp. 106–110. (in Russian)
3. Pilipenko N.V. Uncertainty of non-stationary heat flux recovery by parametric identification of differential-difference model of heat transmission. *Journal of Instrument Engineering*, 2017, vol. 60, no. 7, pp. 664–671. (in Russian). <https://doi.org/10.17586/0021-3454-2017-60-7-664-671>
4. Pilipenko N.V., Gladskih D.A. Determination of the heat losses of buildings and structures by solving inverse heat conduction problems. *Measurement Techniques*, 2014, vol. 57, no. 2, pp. 181–186. <https://doi.org/10.1007/s11018-014-0427-y>
5. Sivakov I.A., Pilipenko N.V. Kalman filter application when restoring heat flux density on study object surface in pulsed wind tunnel. *Proc. IV All-Russian Scientific-Practical Conference with International Participation "Information Technologies in Science, Education and Economy"*, 2012, pp. 55–58. (in Russian)
6. Pilipenko N. Dynamic characteristics for different types of heat flow receivers based on differential-difference models of heat transfer. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2009, no. 3(61), pp. 52–58. (in Russian)
7. Pilipenko N.V. The uncertainty of measuring unsteady surface temperature of massive bodies. *Journal of Instrument Engineering*, 2016, vol. 59, no. 9, pp. 767–772. (in Russian). <https://doi.org/10.17586/0021-3454-2016-59-9-767-772>
8. Pilipenko N.V. Using the extended Kalman filter in nonstationary thermal measurement when solving inverse heat transfer problems. *Journal of Instrument Engineering*, 2019, vol. 62, no. 3, pp. 212–217. (in Russian). <https://doi.org/10.17586/0021-3454-2019-62-3-212-217>
9. Bosholm F., López-Navarro A., Gamarra M., Corberán J.M., Payá J. Reproducibility of solidification and melting processes in a latent heat thermal storage tank. *International Journal of Refrigeration*, 2016, vol. 62, pp. 85–96. <https://doi.org/10.1016/j.ijrefrig.2015.10.016>

- теплопроводности // Известия высших учебных заведений. Приборостроение. 2019. Т. 62. № 3. С. 212–217. <https://doi.org/10.17586/0021-3454-2019-62-3-212-217>
9. Bosholm F., López-Navarro A., Gamarra M., Corberán J.M., Payá J. Reproducibility of solidification and melting processes in a latent heat thermal storage tank // International Journal of Refrigeration. 2016. V. 62. P. 85–96. <https://doi.org/10.1016/j.ijrefrig.2015.10.016>
 10. Sharma A., Tyagi V.V., Chen C.R., Buddhi D. Review on thermal energy storage with phase change materials and applications // Renewable and Sustainable Energy Reviews. 2009. V. 13. N 2. P. 318–345. <https://doi.org/10.1016/j.rser.2007.10.005>

Авторы

Пилипенко Николай Васильевич — доктор технических наук, профессор, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 7006938207](https://orcid.org/0000-0001-9328-3166), <https://orcid.org/0000-0001-9328-3166>, pilipenko38@mail.ru

Колодийчук Павел Андреевич — аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, <https://orcid.org/0000-0002-9984-1613>, pkolodiychuk@itmo.ru

Заричняк Юрий Петрович — доктор физико-математических наук, профессор, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, [sc 6701513411](https://orcid.org/0000-0001-8713-3583), <https://orcid.org/0000-0001-8713-3583>, zarich4@gmail.com

Authors

Nikolai V. Pilipenko — D. Sc., Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 7006938207](https://orcid.org/0000-0001-9328-3166), <https://orcid.org/0000-0001-9328-3166>, pilipenko38@mail.ru

Pavel A. Kolodiychuk — PhD Student, ITMO University, Saint Petersburg, 197101, Russian Federation, <https://orcid.org/0000-0002-9984-1613>, pkolodiychuk@itmo.ru

Yuri P. Zarichnyak — D. Sc. (Physics & Mathematics), Professor, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, [sc 6701513411](https://orcid.org/0000-0001-8713-3583), <https://orcid.org/0000-0001-8713-3583>, zarich4@gmail.com

Статья поступила в редакцию 16.07.2022

Одобрена после рецензирования 25.10.2022

Принята к печати 16.11.2022

Received 16.07.2022

Approved after reviewing 25.10.2022

Accepted 16.11.2022



Работа доступна по лицензии
Creative Commons
«Attribution-NonCommercial»

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.

PROCEEDINGS OF THE XXXII SCHOOL ON HOLOGRAPHY

Part II

Pulse recording of dynamic holograms in bismuth silicate crystal in a broad wavelength range

Hybrid endoscope with television and multispectral image processing for the internal organs cancer early diagnostics

Modelling of a composite waveguide holographic display

Dadenkov I.G., Tolstik A.L., Miksiuk Yu.I., Saechnikov K.A. 1025

Zaichenko K.V., Gurevich B.S., Belyaev A.V., Svyatkina V.I. 1031

Kharitonov D.Yu., Akhmetov D.M., Muslimov E.R., Gilfanov A.R., Pavlycheva N.K. 1037

OPTICAL ENGINEERING

Application of infrared spectroscopy methods in studying compositions for paper sizing

Distribution optimization method of pixel density by surveillance area

Babakhanova Kh.A., Sadridinov A.A., Abdunazarov M.M., Babakhanova M.A., Gromyko I.G. 1048

Volkhonskiy V.V., Kovalevskiy V.A. 1055

AUTOMATIC CONTROL AND ROBOTICS

Evaluation and development of method for compensating the positioning error of computer numeric control equipment (in English)

Compensation of output external disturbances for a class of linear systems with control delay

Rizki M.A., Fedosov Yu.V. 1063

Bui V.H., Margun A.A. 1072

MATERIAL SCIENCE AND NANOTECHNOLOGIES

Luminescence technique for studying the growth of AgInS₂ quantum dots (in English)

Peculiarities of pulsed laser deposition of thin InGaAsN films in an active background gas atmosphere

Determination of the electron distribution in thin barrier AlGaAs/GaAs superlattices by capacitance-voltage profiling

Spectral and kinetic properties of silver sulfide quantum dots in an external electric field

Ahmad A.K., Mohammed A.H., Skaptsov A.A., Devitsky O.V. 1078

Vasilkova E.I., Pirogov E.V., Sobolev M.S., Baranov A.I., Gudovskikh A.S., Bouravleuv A.D. 1092

Daibagya D.S., Ambrozevich S.A., Perepelitsa A.S., Zakharchuk I.A., Osadchenko A.V. Bezverkhnyaya D.M., Avramenko A.I., Selyukov A.S. 1098

Lukyantsev D.S., Lubchenko A.V., Ivanov D.A., Lubchenko O.I., Fedotov A.S. 1104

Osadchenko A.V., Vashchenko A.A., Zakharchuk I.A., Daibagya D.S., Ambrozevich S.A., Volodin N.Yu., Cheptsov D.A., Dolotov S.M., Traven V.F., Avramenko A.I., Semenova S.L., Selyukov A.S. 1112

Kumar N., Gupta P., Singh R.K. 1119

Influence of nano-sized horizontal inhomogeneities on surface profiling by means of XPS

Organic light-emitting diodes with new dyes based on coumarin

Fabrication and characterization of hybrid composite of Al₆₀82/SiC/rice husk powder using friction stir processing (in English)

COMPUTER SCIENCE

A multi-path secure routing for the detection of node capturing attack in wireless sensor network (in English)

A method for documenting architectural solutions of computing platforms

Improving out of vocabulary words recognition accuracy for an end-to-end Russian speech recognition system (in English)

Method for monitoring the state of elements of cyber-physical systems based on time series analysis

Application of the text wave model to the sentiment analysis problem

Automated evaluation of ECG parameters during the COVID-19 pandemic

Kolangiappan J., Senthil Kumar A. 1127

Gorbachev Ya.G. 1136

Andrusenko A.Yu., Romanenko A.N. 1143

Semenov V.V. 1150

Gruzdeva A.S., Iurev R.N., Bessmertny I.A. 1159

Vatian A.S., Gusarova N.F., Dobrenko N.V., Zmievsky D.A., Kabyshev M.V., Polevaya T.A., Tatarinova A.A., Tomilov I.V. 1166

Gribanov T.A., Filchenkov A.A., Azarov A.A., Shalyto A.A. 1178

Rybkin K.E., Filchenkov A.A., Azarov A.A., Zabashta A.S., Shalyto A.A. 1187

Multi-agent adaptive routing by multi-head-attention-based twin agents using reinforcement learning

Joint learning of agents and graph embeddings in a conveyor belt control problem

MODELING AND SIMULATION

Simulation of radiative transfer in gas-liquid foams

The effect of signal-to-noise ratio value on the error in measuring acoustic emission parameters: statistical assessment

Simulating the process of steady-state thermorefectance for measuring the thermal conductivity of materials

Modeling and simulation of one- and two-row six-bladed ducted fans

Isaeva A.A., Isaeva E.A., Pantyukov A.V. 1197

Fedorov A.V., Altay Y., Stepanova K.A., Kuzivanov D.O. 1205

Tukmakova A.S., Demchenko P.S., Tkhorzhevskiy I.L., Novotelnova A.V., Khodzitskiy M.K. 1216

Dudnikov S.Yu., Bulat M.P., Vokin L.O., Kuznetsov P.N., Chernyshov P.S. 1226

BRIEF PAPERS

Differential-difference model of heat transfer in solids using the method of parametric identification

Pilipenko N.V., Kolodiychuk P.A., Zarichnyak Yu.P. 1237