

ISSN 1816-0301

ИНФОРМАТИКА

ТОМ 15



2

АПРЕЛЬ-ИЮНЬ

2018

ОТ РЕДАКЦИИ

В журнале «Информатика» публикуются оригинальные и обзорные статьи, описывающие результаты фундаментальных и прикладных исследований специалистов академического и вузовского профиля в области информатики и информационных технологий. Основной целью журнала является публикация наиболее значимых новых результатов в указанной области. Приветствуются статьи, описывающие заключительные результаты научных проектов и диссертационных исследований, открывающие новые направления исследований, которые находятся на стыке информатики и других наук.

К публикации приглашаются отечественные и зарубежные авторы, специалисты в области информационных технологий и молодые ученые.

В декабре 2017 г. журнал включен в базу данных Российского индекса научного цитирования (РИНЦ). С помощью инструментов и сервисов, доступных на платформе eLIBRARY (раздел «Личный кабинет»), можно самостоятельно корректировать список своих публикаций и цитирований в РИНЦ.

Журнал рассчитан на широкий круг специалистов в области информатики и информационных технологий.

Сайт журнала: <http://inf.grid.by>

Адрес редакции:

ул. Сурганова, 6, к. 305, г. Минск, Республика Беларусь

Тел. +375(017)284 26 22

E-mail: rio@newman.bas-net.by

ОБЪЕДИНЕННЫЙ ИНСТИТУТ ПРОБЛЕМ ИНФОРМАТИКИ
НАЦИОНАЛЬНОЙ АКАДЕМИИ НАУК БЕЛАРУСИ

ИНФОРМАТИКА

Том 15, № 2, апрель-июнь 2018

Ежеквартальный научный журнал

Издается с января 2004 г.

Учредитель – Объединенный институт проблем информатики
Национальной академии наук Беларуси

Г л а в н ы й р е д а к т о р

Александр Васильевич Тузиков – генеральный директор Объединенного института проблем информатики Национальной академии наук Беларуси, Минск, Республика Беларусь

Р е д а к ц и о н н а я к о л л е г и я

- М. Я. Ковалев** – Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Республика Беларусь (*заместитель главного редактора*)
- С. В. Абламейко** – Белорусский государственный университет, Минск, Республика Беларусь
- В. В. Анищенко** – Компания «СофтКлуб», Минск, Республика Беларусь
- П. Н. Бибило** – Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Республика Беларусь
- М. Н. Бобов** – ОАО «АГАТ – системы управления» – управляющая компания холдинга «Геоинформационные системы управления», Минск, Республика Беларусь
- А. Б. Долгий** – Высшая инженерная школа Бретани, Нант, Франция
- А. Н. Дудин** – Белорусский государственный университет, Минск, Республика Беларусь
- А. А. Карпов** – Институт информатики и автоматизации Российской академии наук (СПИИРАН), Санкт-Петербург, Российская Федерация
- С. Я. Килин** – Президиум Национальной академии наук Беларуси, Минск, Республика Беларусь
- В. В. Краснопрошин** – Белорусский государственный университет, Минск, Республика Беларусь
- А. М. Крот** – Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Республика Беларусь
- С. В. Кругликов** – Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Республика Беларусь
- С. П. Кундас** – Белорусский национальный технический университет, Минск, Республика Беларусь

- Н. А. Лиходед** – Белорусский государственный университет, Минск, Республика Беларусь
- П. П. Матус** – Институт математики Национальной академии наук Беларуси, Минск, Республика Беларусь
- А. А. Петровский** – Белорусский государственный университет информатики и радиоэлектроники, Минск, Республика Беларусь
- В. А. Скляр** – Университет Авейру, Португалия
- Ю. С. Сотсков** – Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Республика Беларусь
- А. Л. Стемпковский** – Институт проблем проектирования в микроэлектронике Российской академии наук, Москва, Российская Федерация
- Ю. С. Харин** – Научно-исследовательский институт прикладных проблем математики и информатики Белорусского государственного университета, Минск, Республика Беларусь
- А. Ф. Чернявский** – Институт прикладных физических проблем им. А. Н. Севченко Белорусского государственного университета, Минск, Республика Беларусь
- В. Н. Ярмолик** – Белорусский государственный университет информатики и радиоэлектроники, Минск, Республика Беларусь

ИНФОРМАТИКА

Том 15, № 2, апрель-июнь 2018

Ответственный за выпуск *С. С. Мойсейчик*
Редактор *Г. Б. Гончаренко*
Корректор *А. А. Михайлова*
Компьютерная верстка *О. Б. Бутевич*

Сдано в набор 04.05.2018. Подписано в печать 11.06.2018.
Формат 60×84 1/8. Бумага офсетная. Гарнитура Таймс. Ризография.
Усл. печ. л. 12,8. Уч.-изд. л. 12,5. Тираж 60 экз. Заказ 6.

Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя, распространителя печатных изданий № 1/274 от 04.04.2014.
ЛП № 02330/444 от 18.12.13.
Ул. Сурганова, 6, 220012, Минск.

THE UNITED INSTITUTE OF INFORMATICS PROBLEMS
OF THE NATIONAL ACADEMY OF SCIENCES OF BELARUS

INFORMATICS

Vol. 15, no. 2, April-June 2018

Published quarterly
Issued since January 2004

Founder – the United Institute of Informatics Problems
of the National Academy of Sciences of Belarus

Editor-in-Chief

Alexander V. Tuzikov – General Director of the United Institute of Informatics Problems
of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus

Editorial Board

- | | |
|--------------------------------|---|
| Mikhail Y. Kovalyov | – the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus (<i>Deputy Editor-in-Chief</i>) |
| Sergey V. Ablameyko | – Belarusian State University, Minsk, Republic of Belarus |
| Uladimir V. Anishchanka | – SoftClub Ltd., Minsk, Republic of Belarus |
| Petr N. Bibilo | – the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus |
| Mikhail N. Bobov | – Open Joint-Stock Company “AGAT – Control Systems – Managing Company of Geoinformation Control Systems Holding”, Minsk, Republic of Belarus |
| Alexandre B. Dolgui | – IMT Atlantique, Nantes, France |
| Alexander N. Dudin | – Belarusian State University, Minsk, Republic of Belarus |
| Alexey A. Karpov | – Institute of Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), St. Petersburg, Russian Federation |
| Sergey Ya. Kilin | – Presidium of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus |
| Viktor V. Krasnoproshin | – Belarusian State University, Minsk, Republic of Belarus |
| Alexander M. Krot | – the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus |
| Sergey V. Kruglikov | – the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus |
| Semen P. Kundas | – Belarusian National Technical University, Minsk, Republic of Belarus |

- Nikolai A. Likhoded** – Belarusian State University, Minsk, Republic of Belarus
Petr P. Matus – Institute of Mathematics of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus
Alexander A. Petrovsky – Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus
Valery A. Sklyarov – University of Aveiro, Portugal
Yuri N. Sotskov – the United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus
Alexander L. Stempkovsky – the Institute for Design Problems in Microelectronics of the Russian Federation Academy of Sciences, Moscow, Russian Federation
Yuriy S. Kharin – Research Institute for Applied Problems of Mathematics and Informatics of the Belarusian State University, Minsk, Republic of Belarus
Alexander F. Cherniavsky – A. N. Sevchenko Institute of Applied Physical Problems of the Belarusian State University, Minsk, Republic of Belarus
Vyacheslav N. Yarmolik – Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

INFORMATICS

Vol. 15, no. 2, April-June 2018

Issue Head *S. S. Maiseichyk*
Editor *H. B. Hancharenka*
Corrector *H. A. Mikhailava*
Computer Imposition *V. B. Butsevich*

Sent for press 04.05.2018. Output 11.06.2018.
Format 60×84 1/8. Offset paper. Headset Times. Riesography.
Printed sheets 12,8. Publisher's signatures 12,5. Circulation 60 copies. Order 6.

State Scientific Institution «The United Institute of Informatics Problems
of the National Academy of Sciences of Belarus».
Certificate on the state registration of the publisher, manufacturer,
distributor of printing editions no. 1/274 dated 04.04.2014.
License for the press no. 02330/444 dated 18.12.13.
6, Surganov Str., 220012, Minsk, Republic of Belarus.

ISSN 1816-0301 (print)

СОДЕРЖАНИЕ

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ

Ковалев В. А., Козловский С. А., Калиновский А. А. Генерация искусственных рентгеновских изображений грудной клетки с использованием генеративно-состязательных нейронных сетей.....	7
Tuyet D. V., Ablameyko S. V. A model for medical diagnostic laboratory: multimedia database approach.....	17
Рыбенков Е. В., Петровский Н. А. Синтез FPGA-архитектур банков фильтров на основе блочной лестничной факторизации в алгебре кватернионов (часть 1).....	29
Лысы С. І., Станіславенка Г. Р., Гецэвіч Ю. С. Аўтаматызаваная генерацыя алфавітна-прадметнага паказальніка універсальнай дзесятковай класіфікацыі на беларускай мове.....	45
Залесский Б. А., Троцкий Ф. С. Параллельная версия детектора экстремальных особых точек изображений.....	55

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Ерофеенко В. Т. Модели базисных электромагнитных волн в биизотропной среде с пространственной дисперсией	64
Романчак В. М. Субъективное оценивание вероятности	74

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

Кругликов С. В., Бекиш А. Р. Обобщенная структура имитационной модели групповой деятельности операторов АСУ.....	83
Червяков Н. И., Коляда А. А., Коляда Н. А., Кучуков В. А., Протасеня С. Ю. Нейронные сети конечного кольца на основе редуцированной схемы позиционно-модулярного кодового преобразования	98

ISSN 1816-0301 (print)

CONTENTS

SIGNAL, IMAGE AND SPEECH PROCESSING

Kovalev V. A., Kozlovski S. A., Kalinovskiy A. A. Generation of artificial chest radiographs using generative adversarial neural networks.....	7
Tuyet D. V., Ablameyko S. V. A model for medical diagnostic laboratory: multimedia database approach.....	17
Rybenkov E. V., Petrovskiy N. A. Synthesis of FPGA architectures of block lifting-based filter banks in quaternion algebra (part 1).....	29
Lysy S. I., Stanislavenka H. R., Hetsevich Yu. S. Automated alphabetic subject index generation for universal decimal classification in Belarusian.....	45
Zalesky B. A., Trotski Ph. S. Parallel version of detector of extremal key points on images.....	55

MATHEMATICAL MODELING

Erofeenko V. T. Models of basic electromagnetic waves in the bi-isotropic medium with space dispersion.....	64
Romanchak V. M. The subjective measurement of probability.....	74

INTELLIGENT SYSTEMS

Kruglikov S. V., Bekish A. R. Generalized structure of the simulation model of the group activity of ACS operators.....	83
Chervyakov N. I., Kolyada A. A., Kolyada N. A., Kuchukov V. A., Protasenia S. U. Neural networks of the final ring based on the reduction scheme of the position-modular-code transformation.....	98

ISSN 1816-0301 (print)

ОБРАБОТКА СИГНАЛОВ, ИЗОБРАЖЕНИЙ И РЕЧИ*SIGNAL, IMAGE AND SPEECH PROCESSING*

УДК 004.9

Поступила в редакцию 29.11.2017

Received 29.11.2017

В. А. Ковалев, С. А. Козловский, А. А. Калиновский*Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск, Беларусь***ГЕНЕРАЦИЯ ИСКУССТВЕННЫХ РЕНТГЕНОВСКИХ ИЗОБРАЖЕНИЙ
ГРУДНОЙ КЛЕТКИ С ИСПОЛЬЗОВАНИЕМ
ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ НЕЙРОННЫХ СЕТЕЙ**

Аннотация. Рассматривается задача генерации правдоподобных (трудноотличимых от реальных) рентгеновских изображений грудной клетки человека в норме. Указанная задача решается с использованием генеративно-состязательных нейронных сетей (Generative Adversarial Nets). Степень правдоподобия получаемых результатов оценивается как визуально, так и количественно путем сравнения дескрипторов структуры изображений, основанных на локальных двоичных шаблонах.

Ключевые слова: нейронные сети, генеративно-состязательные сети, глубокое обучение, медицинские изображения, рентгеновские изображения грудной клетки

Для цитирования. Ковалев, В. А. Генерация искусственных рентгеновских изображений грудной клетки с использованием генеративно-состязательных нейронных сетей / В. А. Ковалев, С. А. Козловский, А. А. Калиновский // Информатика. – 2018. – Т. 15, № 2. – С. 7–16.

V. A. Kovalev, S. A. Kozlovski, A. A. Kalinovsky*The United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Belarus***GENERATION OF ARTIFICIAL CHEST RADIOGRAPHS USING
GENERATIVE ADVERSARIAL NEURAL NETWORKS**

Abstract. This paper deals with the problem of generating artificial chest x-ray images which expected to be almost undistinguishable from real ones. Generation was performed using Generative Adversarial Nets (GAN). Similarity of resultant artificial images to the real ones was evaluated both by visual examination and by quantitative assessment using commonly known Local Binary Patterns. It was concluded that GANs can be successfully employed for generating realistically appearing artificial chest radiographs. However, an automatic procedure of selecting “most realistic” results is necessary for excluding the final visual quality control stage and making the whole generation process fully automatic.

Keywords: neural networks, generative adversarial networks, deep learning, medical imaging, lungs x-ray images

For citation. Kovalev V. A., Kozlovski S. A., Kalinovsky A. A. Generation of artificial chest radiographs using generative adversarial neural networks. *Informatics*, 2018, vol. 15, no. 2, pp. 7–16 (in Russian).

Введение. В последние несколько лет наблюдается значительный прогресс в области разработки и использования сверточных нейронных сетей и методов так называемого глубокого, или глубинного, обучения (от англ. Deep Learning). В частности, данные методы показали высокую эффективность при решении широкого спектра задач анализа, классификации и распознавания биомедицинских изображений. Общее описание используемых подходов и примеры решения конкретных задач можно найти в обзорной публикации [1]. Результаты оценки сравнительной эффективности традиционных и нейросетевых методов решения задач класси-

фикационного и регрессионного типов на тестовой выборке рентгеновских изображений грудной клетки 10 000 здоровых людей в возрасте от 21 до 70 лет представлены в работах [2, 3]. Экспериментально показано, что использование сверточных нейронных сетей для выделения количественных признаков изображений является значительно более эффективным подходом по сравнению с традиционными методами. Так, например, при использовании нейросетевого подхода в тестовой задаче предсказания возраста человека по рентгеновскому снимку средне-квадратичная ошибка составляет 6,1 года, в то время как лучший результат, достигнутый при использовании традиционных методов, – 11,1 года, что почти в два раза хуже.

Наряду с указанными выше достоинствами современные методы глубокого обучения имеют один существенный недостаток: для эффективного обучения сверточной нейронной сети требуется очень большое количество (десятки и сотни тысяч) аннотированных изображений, что особенно трудно обеспечить в случае биомедицинских изображений. Действительно, в задачах машинного обучения, которые зачастую оперируют с обычными цифровыми фотоснимками, проблема получения большого их количества является хотя и трудоемкой, но вполне решаемой. Кроме того, ручная аннотация снимков – выделение таких целевых объектов, как люди, автомобили, здания, животные и т. п., – не требует никакой профессиональной подготовки и может осуществляться практически кем угодно.

В случае биомедицинских изображений ситуация существенно отличается от описанной выше: количество доступных изображений естественным образом ограничивается поступающим потоком пациентов, существующими возможностями и технологией накопления и цифрового архивирования данных, доступных для лечебного учреждения, уровнем компьютерной грамотности, квалификацией медицинского персонала и др. Кроме того, имеется еще целый ряд дополнительных факторов, специфичных для данной области. Например, известно, что существует большой перечень достаточно редких заболеваний и редких форм распространенных заболеваний, по которым большое количество снимков в рамках одного или нескольких родственных медицинских учреждений не может быть получено в принципе. Ситуация с ручной аннотацией медицинских изображений также гораздо более тяжелая, так как такие изображения относятся к профессиональным данным, аннотация которых может осуществляться только людьми с соответствующим специальным образованием и практическим опытом.

Целью данной работы является экспериментальное исследование возможности генерации правдоподобных рентгеновских изображений грудной клетки для их последующего использования вместо реальных изображений при решении различных практических задач. Поскольку искусственные медицинские изображения, за редким исключением, не имеют никаких законодательных и этических ограничений для распространения, они могут свободно использоваться также в учебных целях, в качестве иллюстраций в современных цифровых мультимедийных продуктах и презентациях, в виде пиктограмм в компьютерных программах и т. п.

Следует подчеркнуть, что в настоящей статье приводятся результаты лишь первого этапа исследований, направленных на разработку методов и программного обеспечения генерации правдоподобных медицинских изображений. Поэтому проблема использования получаемых искусственных изображений для обучения сверточных нейронных сетей с целью диагностики заболеваний, которая является одной из конечных целей данного направления, здесь не рассматривается.

Архитектура генеративно-сопоставительных нейронных сетей. Генеративно-сопоставительные нейронные сети (Generative Adversarial Nets, GAN) были предложены в работе [4]. Наряду с вариационным автокодировщиком [5] данный тип нейросетевых архитектур используется в качестве одного из современных методов обучения без учителя (Unsupervised Learning [4]). Как уже упоминалось выше, разработка эффективных методов обучения без учителя является одной из ключевых проблем в области глубокого обучения, так как подготовка размеченных данных – достаточно трудоемкая и дорогостоящая процедура. Классической архитектурой GAN (рис. 1) является модель с двумя модулями: генератором и дискриминатором. Задача генератора заключается в преобразовании (генерации) случайного вектора z (зачастую из равномерного распределения) в вектор более высокой размерности, соответствующий размерности реальных данных. При этом в процессе обучения модель генератора стремится подобрать внут-

ренные параметры так, чтобы выход был «похож» на примеры из имеющегося набора данных. Критерий похожести определяется другим модулем – дискриминатором. Дискриминатор в процессе обучения старается подобрать внутренние параметры так, чтобы он мог отличить данные из реального набора от данных, созданных генератором.

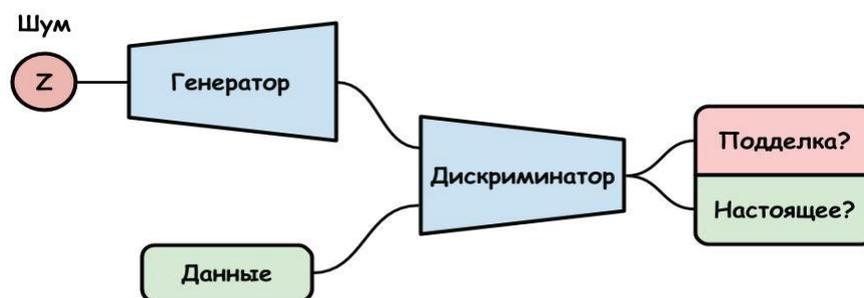


Рис. 1. Общая архитектура GAN

При обучении GAN каждый из модулей конкурирует с другим в процессе минимизации совместной функции ошибки $V(D, G)$, которую можно формализовать как минимакс следующим образом [4]:

$$\min_G \max_D V(D, G) = E_{p \sim p_{data}(x)} \log(D(x)) + E_{z \sim p_z(z)} \log(1 - D(G(z))),$$

где D – модель дискриминатора; G – модель генератора; x – реальные данные, набираемые из априорного распределения $p_{data}(x)$ (в рассматриваемых экспериментах это равномерное распределение); z – случайный низкоразмерный вектор, формируемый из случайного равномерного распределения $p_z(z)$.

Из-за локального характера структуры изображения (близкие пиксели коррелируют, далекие – нет) многослойная сверточная сеть является эталонным выбором в качестве архитектуры для классификации изображений [6]. Поэтому для работы с изображениями в роли дискриминатора выступает многослойная сверточная сеть. В работе [7] была представлена архитектура dcGAN (глубокая сверточная генеративно-состязательная сеть), где в качестве генератора авторы предложили применять многослойную деконволюционную («разверточную») сеть (рис. 2).

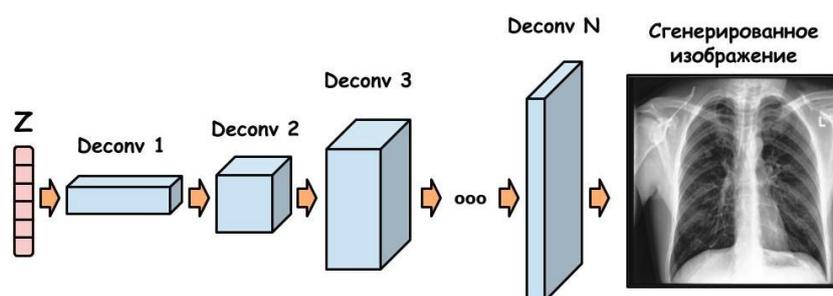


Рис. 2. Деконволюционная архитектура для модели генератора изображений

В настоящем исследовании для генерации изображений использовалась модель dcGAN в виде ее открытой реализации [8]. Как было показано в работе [7], для более стабильной сходимости процесса обучения сети необходимо использовать:

- конволюционные (сверточные) и деконволюционные (разверточные) слои с пропусками вместо операций пулинга;
- батч-нормализацию как в дискриминаторе, так и в генераторе;
- в качестве нелинейной функции активации после сверточного слоя в модели генератора функцию ReLU на всех слоях, кроме выходного, а также гиперболический тангенс на выходе;
- в качестве нелинейной функции в модели дискриминатора на всех слоях LeakyReLU.

Примеры применения генеративно-сопоставительных сетей в области биомедицинских исследований и разработок. Следует отметить, что в настоящее время количество работ, посвященных использованию GAN для генерации правдоподобных медицинских изображений, невелико. Вероятнее всего, это объясняется новизной данного подхода. Среди исследований данного направления прежде всего стоит упомянуть работу [9], которая посвящена генерации синтетических компьютерно-томографических (КТ) изображений пациентов на основе их изображений, полученных с помощью современной ядерной магнитно-резонансной томографии (МРТ). Практическая мотивация данного исследования заключается в том, что КТ-изображения необходимы для планирования локальной радиотерапии пациентов, но их съемка сама по себе предполагает получение пациентом определенной дозы рентгеновского облучения. В то же время выполнение сканирования тела при МРТ предполагает размещение пациента в магнитном поле, что практически безвредно. По результатам работы авторы делают вывод, что получаемые синтетические КТ-изображения представляют анатомию пациента достаточно точно, а алгоритм генерации КТ-изображений по МРТ-томограммам работает устойчиво. К особенностям данного исследования следует отнести тот факт, что для повышения качества результирующих КТ-изображений на этапе обучения нейронной сети авторы предлагают включать в целевую функцию градиентное изображение. В работе [10] авторы решают аналогичную задачу генерации искусственного КТ-изображения на основе МРТ-изображений, однако существенным отличием здесь является то, что на этапе тренировки сети не предполагается использование так называемых парных изображений, т. е. изображений одного и того же пациента. Кроме того, генерация трехмерных КТ-изображений осуществляется послойно, т. е. по двумерным изображениям.

Другое важное исследование данного направления [11] посвящено генерации правдоподобных изображений пациентов, получаемых с помощью позитронной эмиссионной томографии (ПЭТ), широко используемой при диагностике различных форм рака. Решение данной задачи мотивируется тем, что использование методов глубокого обучения для анализа и распознавания ПЭТ-изображений предполагает наличие значительного количества томограмм, представляющих большую вариабельность результатов ПЭТ-сканирования. Указанная вариабельность обусловлена такими факторами, как различия в типах и размерах опухолей и их локализации; естественная вариабельность анатомического строения и размеров тела разных людей; особенности, обусловленные полом и возрастом обследуемого пациента; вариабельность изображений, вызванная техническими различиями томографов от различных производителей, и др. В работе [11] эксперименты по генерации ПЭТ-изображений проводились с использованием совместных трехмерных КТ-, ПЭТ-сканов 50 пациентов, больных раком легкого. На основании этих экспериментов авторы делают вывод, что полученные искусственные изображения достаточно близки к реальным. В частности, результаты распознавания опухолей с использованием нейронной сети, натренированной на синтетических изображениях, лишь на 2,8 % хуже, чем в случае тренировки на изображениях реальных пациентов. Такие результаты выглядят многообещающе, однако в контексте настоящей работы следует подчеркнуть, что ПЭТ-изображения опухолей имеют типичный «туманный», сглаженный, слабоконтрастный вид с характерным «облачным» строением. Это коренным образом отличает их от рентгеновских снимков, представляющих анатомию тела в мельчайших деталях. В работе [12] предлагается другой подход к решению проблемы недостатка входных данных для задачи семантической сегментации МРТ-изображений агрессивного рака простаты. Суть подхода заключается в интеграции GAN в общую схему сегментации с целью своеобразного «отсеивания» неправдоподобных промежуточных вариантов на основе оценки глобальной структуры объекта. Утверждается, что получены многообещающие результаты при использовании ограниченного набора МРТ-изображений (всего лишь 152 пациентов).

Известно, что базовая задача компьютерной томографии как таковой – это задача восстановления объемного изображения по проекциям, которая относится к классу некорректно поставленных. В современных условиях естественно ожидать, что генеративные сети могут быть использованы при решении задачи восстановления в качестве своеобразного «нейросетевого регуляризационного выражения», помогающего уменьшать неопределенность, вызванную не-

хваткой данных из-за стремления уменьшить время сканирования и (или) сократить дозу облучения. В работе [13] предлагается использовать GAN при восстановлении (генерации) МРТ-изображений, которые зачастую требуют слишком много времени и, как следствие, предполагают временное обездвиживание пациента. В условиях сканирования детей показано, что время съемки изображений может быть сокращено на два порядка по сравнению с существующими методами. Известной проблемой в области анализа медицинских изображений является также автоматическая сегментация исследуемого органа, которую необходимо выполнить непосредственно перед процедурой анализа. В работе [14] рассматривается достаточно сложная задача сегментации печени на трехмерных КТ-изображениях. Следуя общей методологии связательных сетей, авторы комбинируют обычную энкодер-декодер (т. е. кодировщик-декодировщик)-сеть сегментации с дискриминационной сетью, которая «следит» за качеством результата. Предложенный метод протестирован на 1000 КТ-изображений. Показано, что новый подход обеспечивает лучшие результаты по сравнению с существующими методами сегментации печени как «в норме», так и при наличии патологии.

Исходные данные. В качестве исходных данных использовалась тестовая база цифровых рентгеновских изображений грудной клетки 10 000 здоровых людей в возрасте от 21 до 70 лет (рис. 3).

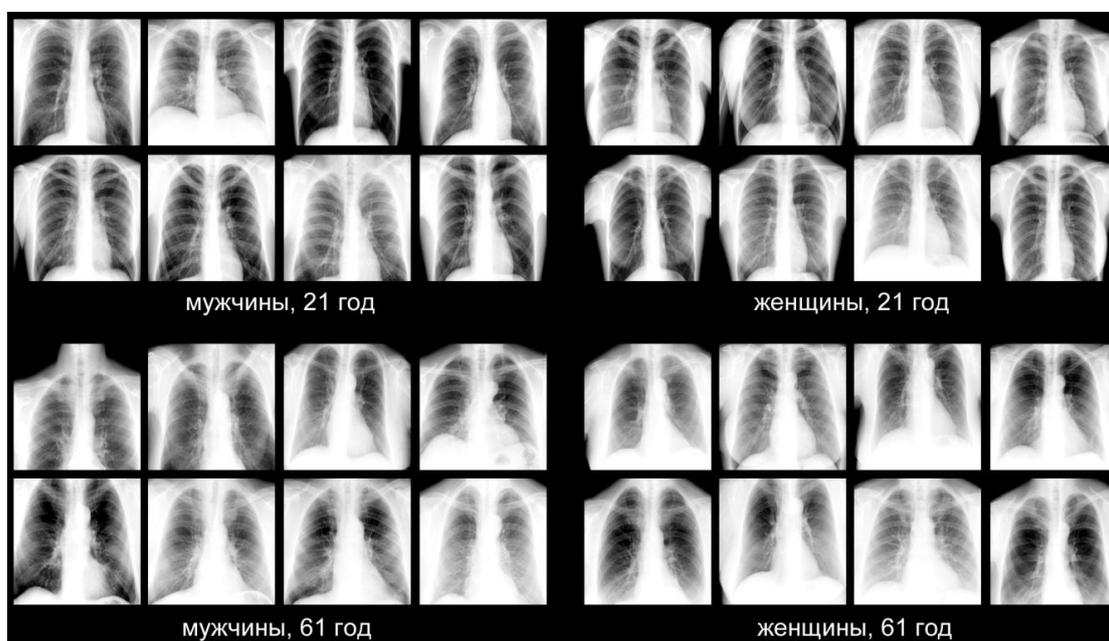


Рис. 3. Примеры исходных изображений

Возраст измерялся в количестве полных лет жизни на дату съемки. База изображений представляла собой тестовую группу, полностью сбалансированную по полу и по возрасту. В частности, на каждый конкретный год жизни из указанного выше диапазона в 50 лет были представлены снимки 200 человек (100 женщин и 100 мужчин).

Процедура генерации изображений и оценки их качества. Следует отметить, что из-за новизны исследований, использующих сверточные нейронные сети и глубокое обучение, большинство работ данного направления носит преимущественно экспериментальный характер. При отсутствии фундаментальных теоретических результатов и обоснований проблема выбора рациональных архитектур и управляющих параметров решается чаще всего экспериментально, методом проб и ошибок.

В настоящем исследовании процедура генерации правдоподобных изображений состояла из пяти основных этапов:

- формирование обучающих выборок, включающих рентгеновские снимки реальных людей;
- обучение генеративно-связательной сети с использованием различного количества эпох, т. е. с различной длительностью обучения;

– генерация искусственных рентгеновских изображений с использованием обученной генеративно-состязательной сети и различных векторов случайных чисел z в качестве входных данных;

– визуальная оценка получаемых результатов и выбор окончательной конфигурации генератора изображений;

– количественная оценка отличий реальных и искусственных изображений.

Поскольку изображения мужчин и женщин естественным образом различаются, было принято решение обучать генеративные сети для мужчин и женщин отдельно. С этой целью для обучения сетей из имеющейся базы рентгеновских снимков были выделены два сбалансированных по возрастам набора – снимки мужчин (набор 1) и снимки женщин (набор 2), по 5000 изображений в каждом наборе.

Сеть с архитектурой dcGAN [8] была обучена отдельно на наборах 1 и 2 с числом эпох от 10 до 100 с шагом 10. В результате визуальной оценки получаемых искусственных изображений было установлено, что использование 50 эпох в качестве параметра обеспечивает разумный компромисс между детальностью представления структуры искусственных изображений и необходимым разнообразием (вариабельностью) их общего вида. При малом количестве эпох наблюдался эффект «недообучения», проявляющийся в размытости, слабой контрастности генерируемых снимков. При большом количестве эпох сеть проявляла тенденцию к переобучению. Указанная тенденция проявлялась в том, что получаемые изображения становились все более и более похожи друг на друга и, как следствие, терялось естественное разнообразие, характерное для набора снимков реальных людей.

Для объективной оценки степени похожести реальных и искусственных изображений использовался метод оценки их близости в пространстве признаков на основе количественных дескрипторов. В качестве дескрипторов, описывающих структуру изображений, были выбраны так называемые локальные бинарные шаблоны (Local Binary Patterns, LBP), которые широко используются при решении самых различных задач анализа и распознавания изображений. В частности, в работе [2] было показано, что на рассматриваемой здесь базе рентгеновских снимков использование LBP-дескрипторов обеспечивает достаточно высокое качество классификации возрастных групп людей (площадь под ROC-кривой варьирует от 0,80 до 0,91 в зависимости от сравниваемых возрастных групп и типа используемого классификатора).

LBP-дескрипторы были подсчитаны для 2000 реальных и 2000 синтезированных снимков. Группа реальных снимков выбиралась равномерно, по 40 изображений для каждого года жизни в интервале от 21 года до 70 лет. Группа из 2000 синтезированных снимков формировалась путем случайного выбора из 5000 сгенерированных изображений. Для оценки степени близости дескрипторов был применен алгоритм t-SNE [15], который оценивал близость векторов признаков в терминах евклидова расстояния.

Результаты исследования. Примеры реальных и искусственных рентгеновских изображений мужчин и женщин представлены на рис. 4 и 5 соответственно.

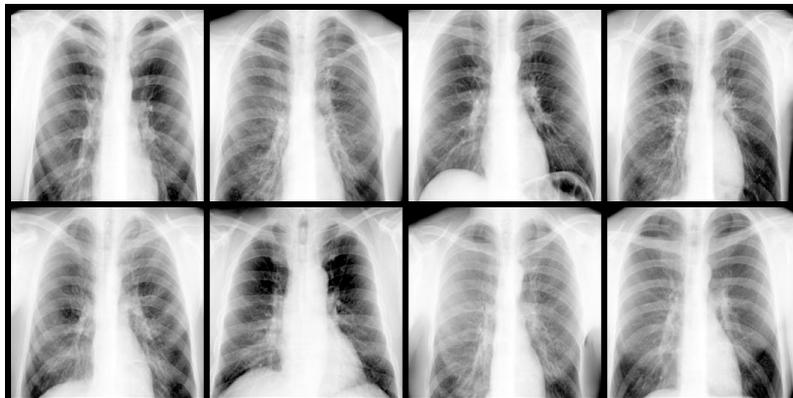


Рис. 4. Реальные снимки грудной клетки мужчин (первый ряд – 1, 3; второй ряд – 2, 4) и сгенерированные снимки (первый ряд – 2, 4; второй ряд – 1, 3)



Рис. 5. Реальные снимки грудной клетки женщин (первый ряд – 1, 3; второй ряд – 2, 4) и сгенерированные снимки (первый ряд – 2, 4; второй ряд – 1, 3)

Результаты визуального сравнения реальных и сгенерированных изображений, представленных на рис. 4 и 5, а также сравнения большого количества изображений, проведенного авторами и их коллегами, позволяют утверждать, что синтетические изображения выглядят достаточно реалистично, разнообразно и многие из них визуально неотличимы от реальных.

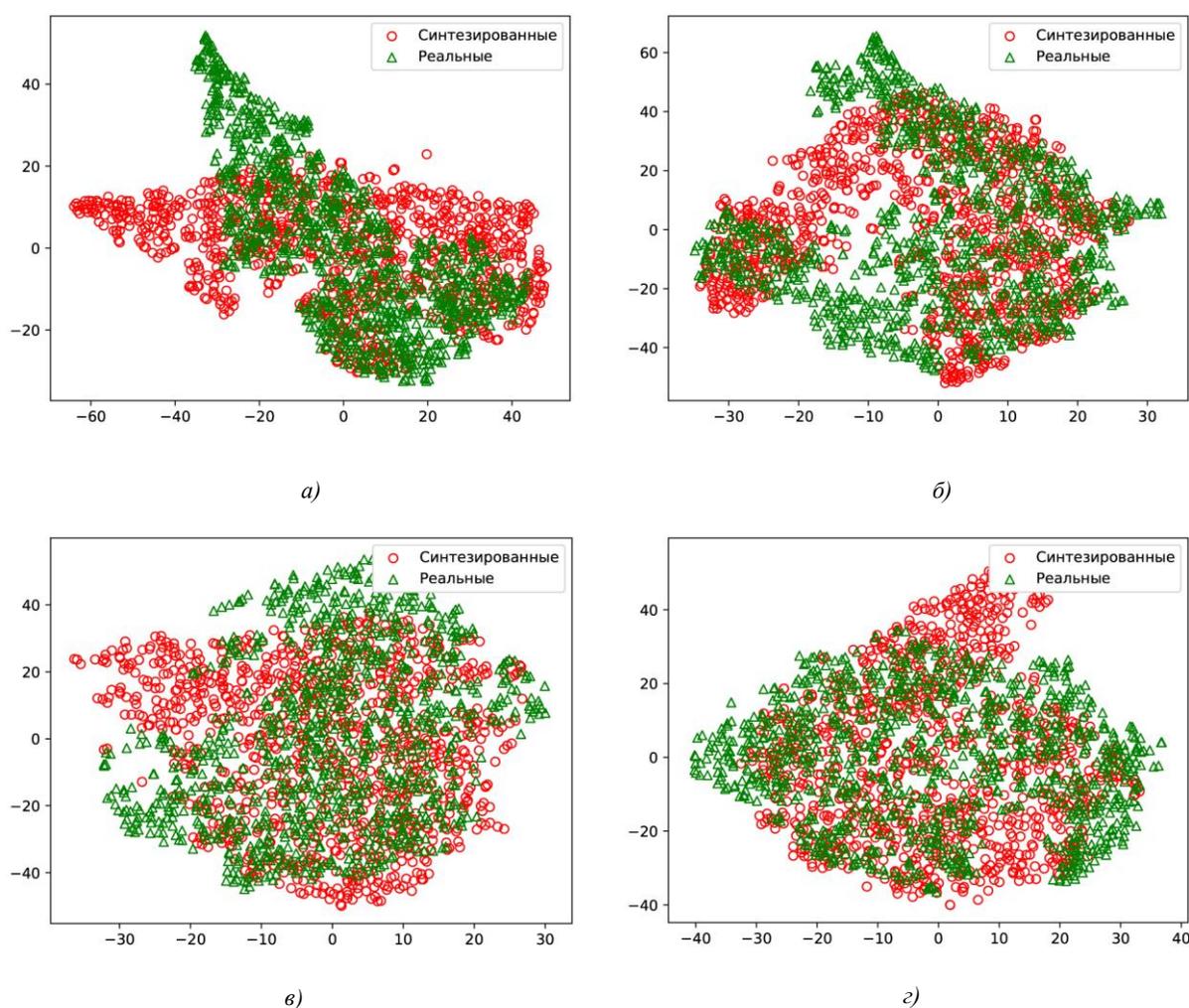


Рис. 6. Распределение реальных и синтетических изображений грудной клетки мужчин (а, б) и женщин (в, г) для LBP-признаков с радиусами $R = 1$ пиксел (а, б) и $R = 4$ пиксела (в, г), редуцированное в двумерное пространство методом t-SNE

Скаттерграммы, представленные на рис. 6, визуализируют расположение реальных изображений мужчин и женщин в пространстве признаков (зеленые треугольники) и сгенерированных (красные кольца). Для двухмерной визуализации исходное 256-мерное пространство LBP-признаков было редуцировано в эквивалентное двухмерное представление с помощью метода t-SNE. Видно, что в пространстве LBP-признаков синтетические и реальные изображения расположены достаточно близко и рассеяны стохастично, что подтверждает субъективное предположение о довольно высокой степени правдоподобия и разнообразии синтезированных изображений. Тем не менее следует отметить присутствие областей пространства, в которые попадают изображения только одной категории. Это можно объяснить тем, что, с одной стороны, синтетические изображения не подвергались никакому отбору и содержат явно неудачные экземпляры (рис. 7), а с другой – ограниченностью возможностей обученной сети, которая обусловлена сложностью обучения генеративно-состязательных сетей как таковых. Как видно из рис. 7, в некоторых случаях сгенерированные изображения демонстрируют необычную для нормы анатомию, могут иметь нетипичные затемнения в одном из легких, а также обладают некоторыми другими нехарактерными элементами.

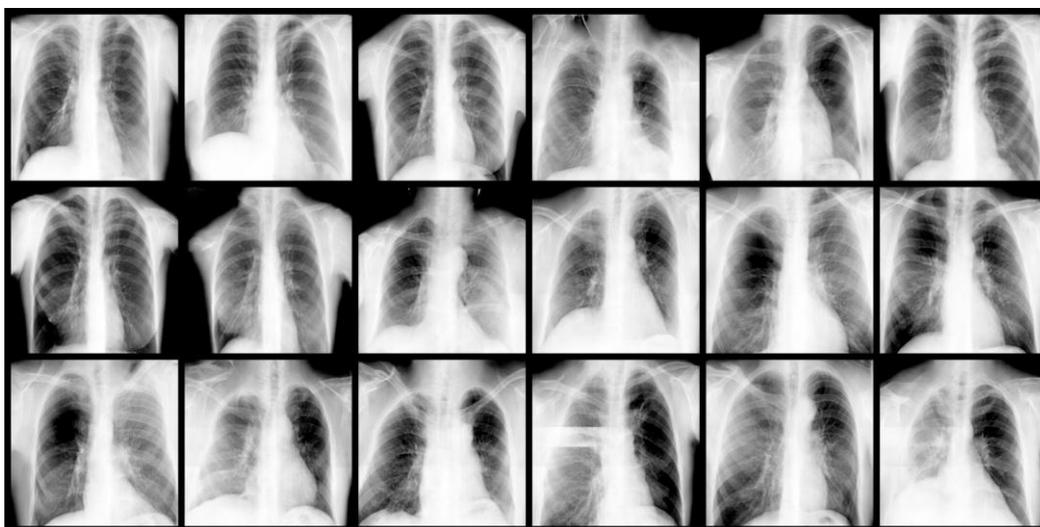


Рис. 7. Примеры неудачных сгенерированных изображений

На рис. 6 можно также заметить, что при использовании LBP-признаков с радиусом $R = 4$, т. е. при рассмотрении более «глобальной» структуры сравниваемых изображений, схожесть сгенерированных и реальных изображений выше, чем при рассмотрении локальной структуры (межпиксельное расстояние $R = 1$). Для получения интегральных количественных оценок степени различия сгенерированных и реальных изображений решалась модельная задача классификации 4000 изображений испытываемых по полу, т. е. задача предсказания пола по рентгеновскому снимку грудной клетки, представленному его LBP-признаками (таблица).

Результаты классификации при обучении на реальных и сгенерированных изображениях

Номер строки	Классификатор	Радиус LBP	Тип изображений обучающей выборки	Точность, %
1	kNN, $k = 20$	$R = 2$	реальные	84,0
2	kNN, $k = 20$	$R = 2$	сгенерированные	64,5
3	kNN, $k = 20$	$R = 4$	реальные	84,3
4	kNN, $k = 20$	$R = 4$	сгенерированные	74,8
5	SVM	$R = 2$	реальные	80,7
6	SVM	$R = 2$	сгенерированные	60,0
7	SVM	$R = 4$	реальные	81,0
8	SVM	$R = 4$	сгенерированные	72,4

Классификация выполнялась с помощью двух известных классификаторов kNN и SVM (rbf). В первой серии экспериментов (см. нечетные строки таблицы) классификаторы обучались на выборке реальных изображений и тестировались также на независимой выборке реальных изображений. Вторая серия экспериментов (см. четные строки таблицы) была аналогична первой, за исключением того, что обучение классификаторов проводилось на искусственных изображениях. Из таблицы видно, что, несмотря на наблюдаемые флуктуации точности классификации в зависимости от параметров дескрипторов изображений и параметров классификаторов, точность классификации при обучении классификаторов на искусственных изображениях пока отстает от точности, получаемой на реальных изображениях. Данный вывод хорошо согласуется с рассмотренными выше результатами анализа распределения изображений в пространстве признаков, а также с фактом наличия среди результатов генерации неудачных экземпляров (см. рис. 7).

Заключение. Полученные результаты позволяют утверждать, что аппарат генеративно-состязательных сетей может быть использован при решении задачи синтеза правдоподобных изображений грудной клетки человека. Применение данного подхода позволяет получать изображения, которые выглядят правдоподобно как при их визуальном сравнении с реальными, так и при количественном сравнении с использованием LBP-дескрипторов.

В то же время в ходе выполнения исследования возник ряд вопросов, получить удовлетворительные ответы на которые авторам пока не удалось:

Как с помощью вектора латентных параметров z влиять на параметры конечного изображения? Например, как задавать возраст человека?

Как интерпретировать оценку близости к реальным изображениям, выдаваемую дискриминатором, которая в общем случае не соответствует визуальной оценке близости, определяемой человеком?

Каким образом реализовать процедуру автоматического отбора лучших изображений при условии, что количество генерируемых изображений практически ничем не ограничено?

Дальнейшим развитием данного направления будет проведение комплекса экспериментальных исследований для оценки возможности применения синтезированных изображений в качестве входных данных для обучения классификационных и иных моделей.

Работа была выполнена при финансовой поддержке гранта Президента Республики Беларусь (распоряжение Президента Республики Беларусь № 32рп от 19 января 2018 г.).

Список использованных источников

1. A survey on deep learning in medical image analysis / G. Litjens [et al.] // *Medical Image Analysis*. – 2017. – Vol. 42. – P. 60–88.
2. A comparison of conventional and Deep Learning methods of image classification on a database of chest radiographs / V. A. Kovalev [et al.] // *Intern. J. of Computer Assisted Radiology and Surgery*. – 2017. – Vol. 12, suppl. 1. – P. S139–S140.
3. Benchmarking the efficiency of Deep Learning methods on the problem of predicting subjects' age by chest radiographs / V. A. Kovalev [et al.] // *Proc. 3rd Intern. Conf. Big Data and Advanced Analytics*. – Minsk : BSUIR, 2017. – P. 75–82.
4. Generative Adversarial Networks / I. J. Goodfellow [et al.] // *arXiv.org* [Electronic resource]. – 2014. – Mode of access: <https://arxiv.org/pdf/1406.2661>. – Date of access: 25.09.2017.
5. Kingma, D. P. Auto-encoding variational bayes / D. P. Kingma, M. Welling // *arXiv.org* [Electronic resource]. – 2014. – Mode of access: <https://arxiv.org/pdf/1312.6114>. – Date of access: 25.09.2017.
6. LeCun, Y. Gradient-based learning applied to document recognition / Y. LeCun // *Proc. of the IEEE*. – 1998. – Vol. 86(11). – P. 2278–2324.
7. Unsupervised representation learning with deep convolutional generative adversarial networks / A. Radford, L. Metz, S. Chintala // *arXiv.org* [Electronic resource]. – 2016. – Mode of access: <https://arxiv.org/pdf/1511.06434>. – Date of access: 25.09.2017.
8. DCGAN in tensorflow [Electronic resource]. – Mode of access: <https://github.com/carpedm20/DCGAN-tensorflow>. – Date of access: 25.09.2017.
9. Medical image synthesis with context-aware generative adversarial networks / D. Nie [et al.] // *MICCAI, LNCS*. – Springer, Cham, 2017. – Vol. 10435. – P. 417–425.
10. Deep MR to CT Synthesis using Unpaired Data / J. M. Wolterink [et al.] // *arXiv.org* [Electronic resource]. – 2017. – Mode of access: <https://arxiv.org/pdf/1708.01155>. – Date of access: 22.09.2017.
11. Synthesis of Positron Emission Tomography (PET) images via multi-channel Generative Adversarial Networks (GANs) / L. Bi [et al.] // *Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment*. – Springer, Cham, 2017. – Vol. 10555. – P. 43–51.

12. Adversarial Networks for the Detection of Aggressive Prostate Cancer / S. Kohl [et al.] // arXiv.org [Electronic resource]. – 2017. – Mode of access: <https://arxiv.org/pdf/1702.08014>. – Date of access: 28.09.2017.
13. Deep Generative Adversarial Networks for Compressed Sensing Automates MRI / M. Mardani [et al.] // arXiv.org [Electronic resource]. – 2017. – Mode of access: <https://arxiv.org/pdf/1706.00051>. – Date of access: 25.09.2017.
14. Automatic Liver Segmentation Using an Adversarial Image-to-Image Network / D. Yang [et al.] // arXiv.org [Electronic resource]. – 2017. – Mode of access: <https://arxiv.org/pdf/1707.08037>. – Date of access: 24.10.2017.
15. Van der Maaten, L. Visualizing data using t-SNE / L. Van der Maaten, G. E. Hinton // *J. of Machine Learning Research*. – 2008. – Vol. 9. – P. 2579–2605.

References

1. Litjens G., Kooi T., Bejnordi B., Setio A., Ciompi F., Ghafoorian M. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017, vol. 42, pp. 60–88. doi: 10.1016/j.media.2017.07.005
2. Kovalev V., Kalinovsky A., Liauchuk V., Shukelovich A. A comparison of conventional and Deep Learning methods of image classification on a database of chest radiographs. *Intern. J. of Computer Assisted Radiology and Surgery*, 2017, vol. 12, suppl. 1, pp. S139–S140.
3. Kovalev V., Kalinovsky A., Liauchuk V., Shukelovich A. Benchmarking the efficiency of Deep Learning methods on the problem of predicting subjects' age by chest radiographs. *Proc. 3rd Intern. Conf. Minsk, BSUIR*, 2017, pp. 75–82.
4. Goodfellow I. J., Pouget-Adabie J., Mirza M., Xu B., Warde-Farley D., ... Bengio Y. *Generative Adversarial Networks*. ArXiv.org, 2014. Available at: <https://arxiv.org/pdf/1406.2661> (accessed 25.09.2017).
5. Kingma D. P., Welling M. *Auto-encoding variational bayes*. ArXiv.org, 2014. Available at: <https://arxiv.org/pdf/1312.6114> (accessed 25.09.2017).
6. LeCun Y. Gradient-based learning applied to document recognition. *Proc. of the IEEE*, 1998, vol. 86(11), pp. 2278–2324.
7. Radford A., Metz L., Chintala S. *Unsupervised representation learning with deep convolutional generative adversarial networks*. ArXiv.org, 2016. Available at: <https://arxiv.org/pdf/1511.06434> (accessed 25.09.2017).
8. *DCGAN in tensorflow*. Available at: <https://github.com/carpedm20/DCGAN-tensorflow> (accessed 25.09.2017).
9. Nie D., Trullo R., Lian J., Petitjean C., Ruan S. Medical image synthesis with context-aware generative adversarial networks. *MICCAI, LNCS*, 2017, vol. 10435, pp. 417–425.
10. Wolterink I. M., Dinkla A. M., Savenije M. H. F., Seevinck P.R., van den Berg C. A. T., Isgum I. *Deep MR to CT Synthesis using Unpaired Data*. ArXiv.org, 2017. Available at: <https://arxiv.org/pdf/1708.01155> (accessed 22.09.2017).
11. Bi L., Kim J., Kumar A., Feng D., Fulham M. Synthesis of Positron Emission Tomography (PET) images via multi-channel Generative Adversarial Networks (GANs). *Springer, Cham*, 2017, vol. 10555, pp. 43–51.
12. Kohl S., D. Bonekamp D., Schlemmer H. P., Yaqubi K., Hohenfellner M., Hadaschik B., Radtke J. P. *Adversarial Networks for the Detection of Aggressive Prostate Cancer*. ArXiv.org, 2017. Available at: <https://arxiv.org/pdf/1702.08014> (accessed 28.09.2017).
13. Mardani M., Gong E., Cheng J. Y., Vasawala S., Zaharchuk G., ... Xing L. *Deep Generative Adversarial Networks for Compressed Sensing Automates MRI*. ArXiv.org, 2017. Available at: <https://arxiv.org/pdf/1706.00051> (accessed 25.09.2017).
14. Yang D., Xu D., Zhou S. K., Georgescu B., Chen M., ... Comaniciu D. *Automatic Liver Segmentation Using an Adversarial Image-to-Image Network*. ArXiv.org, 2017. Available at: <https://arxiv.org/pdf/1707.08037> (accessed 24.10.2017).
15. Van der Maaten L., Hinton G. E. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 2008, vol. 9, pp. 2579–2605.

Информация об авторах

Ковалев Василий Алексеевич – кандидат технических наук, заведующий лабораторией анализа биомедицинских изображений, Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: vassili.kovalev@gmail.com

Козловский Сергей Александрович – инженер-программист, Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: kozlovski.serge@gmail.com

Калиновский Александр Александрович – научный сотрудник, Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: gakarak@gmail.com

Information about the authors

Vassiliy A. Kovalev – Ph. D. (Technical Sciences), Head of Laboratory of Biomedical Images Analysis, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganova Str., 220012, Minsk, Republic of Belarus). E-mail: vassili.kovalev@gmail.com

Sergei A. Kozlovski – Software engineer, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganova Str., 220012, Minsk, Republic of Belarus). E-mail: kozlovski.serge@gmail.com

Alexander A. Kalinowski – Researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganova Str., 220012, Minsk, Republic of Belarus). E-mail: gakarak@gmail.com

ISSN 1816-0301 (print)
UDC 681.3

Received 15.01.2018
Поступила в редакцию 15.01.2018

D. V. Tuyet^{3,4}, S. V. Ablameyko^{1,2}

¹*Belarusian State University, Minsk, Belarus*

²*United Institute of Informatics Problems of the National Academy of Sciences
of Belarus, Minsk, Belarus*

³*Binh Duong University, Ho Chi Minh City, Vietnam*

⁴*Vietnam National Space Center of Vietnam Academy of Science
and Technology, Hanoi, Vietnam*

A MODEL FOR MEDICAL DIAGNOSTIC LABORATORY: MULTIMEDIA DATABASE APPROACH

Abstract. There are two main tendencies in Medical Informatics. One is the development of core theory in the information technology itself and the other one is the use of technologies in applications for health improvement. In recent years the deployment of Health Information System (HIS) applications and Laboratory Information System (LIS) in some big hospitals have enhanced the quality of administrative management and actively supported the diagnosis and treatment tasks of the physician. In this paper we will present an architectural model of the Medical Diagnostic Laboratory (MDL) and show its applications.

Keywords: medical diagnostic laboratory, laboratory information system, radiology information systems, picture archiving and communication system, digital image and communication in medicine, health level

For citation. Tuyet D. V., Ablameyko S. V. A model for medical diagnostic laboratory: multimedia database approach. *Informatics*, 2018, vol. 15, no. 2, pp. 17–28.

Д. В. Туйет^{3,4}, С. В. Абламейко^{1,2}

¹*Белорусский государственный университет, Минск, Беларусь*

²*Объединенный институт проблем информатики Национальной академии
наук Беларуси, Минск, Беларусь*

³*Университет Бинь Зьонг, Хошимин, Вьетнам*

⁴*Вьетнамский национальный космический центр
Вьетнамской академии наук и технологий, Ханой, Вьетнам*

МОДЕЛЬ МЕДИЦИНСКОЙ ДИАГНОСТИЧЕСКОЙ ЛАБОРАТОРИИ: МУЛЬТИМЕДИЙНЫЙ ПОДХОД К БАЗЕ ДАННЫХ

Аннотация. В медицинской информатике существуют две основные тенденции. Одной из них является разработка теории в области информационных технологий, а другой – использование информационных технологий в приложениях для медицины. В последние годы развертывание медицинских информационных систем и лабораторных информационных систем в больницах повысило качество административного управления и улучшило процессы диагностики заболеваний и лечения пациентов. В статье представлена архитектурная модель медицинской диагностической лаборатории и показаны ее приложения.

Ключевые слова: медицинская диагностическая лаборатория, лабораторная информационная система, радиологические информационные системы, система архивации и передачи изображений, цифровые изображения и их передача в медицине, уровень здравоохранения

Для цитирования. Туйет, Д. В. Модель медицинской диагностической лаборатории: мультимедийный подход к базе данных / Д. В. Туйет, С. В. Абламейко // Информатика. – 2018. – Т. 15, № 2. – С. 17–28.

Introduction. Radiology Information Systems (RIS) use digital images (in DICOM, JPEG formats) instead of the traditional method of using films. A doctor can easily interact, process images during the analysis using computer software. DICOM standard continues to evolve to meet the growing and changing the role of imaging in clinical and research environments, as emphasized by the many technical working groups [1]. It will help to reduce the need for using films in diagnosis, save the cost of films and the storage issues. RIS is a core of the HIS with the added requirement of

institution's Picture Archiving and Communication System (PACS) and workflow seen in a radiology department. A distinguishing aspect of RIS vs. PACS is that the former is typically constrained to non-imaging data, whereas PACS is primarily concerned with the storage and manipulating of the imaging data [1]. RIS also help to reduce the number of medical images when patient data are stored in a computer system.

All data related to patient diagnosis are those such as medical images, diagnosis, management reporting, a unified storage, synchronization and security. Therefore, doctors and even patients can view details about the medical history over time easily. This will be a great assistance in the examination and treatment of disease.

The system allows multiple access and fast response. Doctors can access data and images of the patient quickly and accurately every time, anywhere based on computing technology in a network system.

It is extremely important in medicine, more than in any other field, that the accuracy is comparable to experts. Diagnoses that are incorrect, or other diagnoses which are missed, may result in serious consequences for the patients. HIS provides support to medical personnel improving the reliability and quality of treatment [2].

MDL occupies a large part of the structure of diagnostic research, both in the quantity of research and the clinical importance of test results – which are an important source of diagnostic information for modern medical diagnostic processes. According to world statistics, in previous decades the quantity of performed clinical laboratory tests and their diagnostic importance exponentially increased – and it continues to increase [3]. Analogous to RIS for radiology, LIS deal with the requirement of laboratory and pathology departments, providing interfaces to the various instruments used to assess chemistry, hematology, immunology, microbiology, genetic and other histopathologic markers. The use of LIS has now become the standard of MDL activity.

The integration of LIS-RIS-PACS-Teleradiology forms a perfect combination to meet the needs of professional work of physicians. This combination forms a closed system and can replace systems based on traditional medical records, which require significant efforts on consultations during analysis of medical images and video sequences in the case of an emergency or difficult cases. PACS and teleradiology combined model in which the PACS also supports imaging centers for radiology reading is shown in fig. 1.

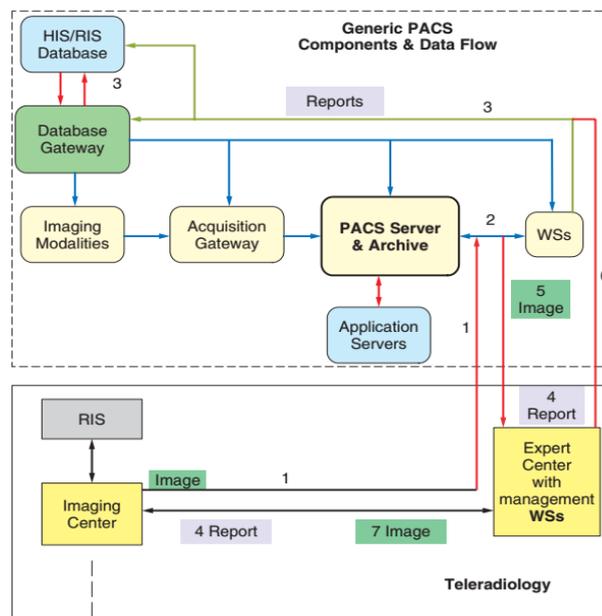


Fig. 1. PACS and teleradiology combined model

Teleradiology can be compiled together with PACS as a healthcare enterprise operation and the workflow of this combined model consists of seven steps:

- Radiologists at PACS WSs read exams from outside imaging centers (step 1).

– After reading by PACS radiologists from its own WSs (step 2), the reports are sent to the HIS database via the database gateway for its own record (step 3) and to the expert center (step 4, 5) from where the report (step 4) is also sent back to the imaging center.

– PACS can also send its own exams to outside expert center for reading (step 5). The expert center returns the report to the PACS database gateway (step 6).

– The imaging center sends images to the expert center for reading as in the pure teleradiology model (step 7).

Ideally the combined teleradiology and PACS model should be used in a healthcare center with the satellite imaging centers, multiple affiliated hospitals, and with backup radiology coverage between hospitals and imaging centers [4].

RIS and MDL system promises to bring a lot of economic efficiency and social networking in the work of community health, improve the quality of people's health care with approximately 90,000 cases per year can save about 25 % compared with the cost of the film in RIS (Huang 2004). FirstHealth RIS system allows for film budget cuts over 37 %, from approximately \$80,000 to \$50,000, director of FirstHealth images – Mike McCarthy has told in Computer World.

In this paper, we propose a model for Medical Diagnostic Laboratory together with the ER-diagram of MDL. We also describe issues of development LIS-RIS-PACS-Teleradiology, its integration into Hospital Information System (HIS) and present some results.

Related work

In Vietnam. With the strong development of the information technologies industry in Vietnam in recent years, most of the major hospitals in the country find the need to apply information technologies in a health sector, particularly in data management issues, medical imaging, video and medical needs of patients for the remote consultation. Starting from practical needs, many hospitals have invested to develop software systems and build Hospital Information Systems.

Since 2003, Viet-Duc Hospital has implemented projects for consultant surgical system online. This is part of a big project to build remote sensing centers of Viet-Duc Hospital. With this model, an experienced medical consultant of Viet-Duc Hospital can sit at his hospital to direct or to control a surgical case carrying out in other places by means of a camera and a remote monitor.

Foreseeing this need, a number of IT companies in Vietnam concentrate investment in the development of systems related to a medical field, including HIS and RIS systems. General research and development of RIS systems are still slow and not being invested properly although this is a core system and it is extremely important.

Transmission infrastructure in Vietnam is the next generation network NGN (Next Generation Network) with IP high-quality backbone network by the Posts and Telecommunications Corporation. It provides fiber optic transmission as technological security VPN (Virtual Private Network) which is a very good support for telemedicine applications. In addition, with the development of research and education network in Vietnam (VinaREN), high-speed connection organizations, research centers and universities in Vietnam have opened up opportunities for the new applications requiring more bandwidth, and there is a large amount of data which need to be exchanged and computed. Combined with the TEIN2 network (currently being upgraded to TEIN3 network), VinaREN can connect to the countries with great potential in science and technology such as France, Hong Kong and Taiwan, etc. This is the foundation network infrastructure to develop the high quality RIS and teleradiology systems.

In the world. In recent years, the term MDL, RIS and Telemedicine has appeared in more developed countries such as the United States, Canada, Switzerland, Japan. Telemedicine is the application of information technologies which conveniently, quickly and effectively bring health service, health education to remote users. These systems have gradually proved necessity and appropriate life needs of a growing society.

First PACS Project titled "Installation Site for Digital Imaging Network and Picture Archiving and Communication System" was implemented in U.S. in 1983 with the investment of U.S. Army and managed by The MITRE Corporation in 1986. According to estimates predicted in 2001, the United States has a cost to the healthcare system to nearly \$2 billion, not including the cost of equipment and other software applications.

In Asia, the First International Symposium on PACS and PHD (Personal Health Data), was held in Japan (7/1982) by the Japan Association of Medical Imaging Technology. Japan considers development of PACS as a national project. The national resources are distributed to many manufacturers and hospitals, universities.

In Europe, EuroPACS (Picture Archiving and Communication Systems in Europe) conferences are held every year since 1983. National Health Service (NHS) in the UK is one of the leading organizations in the research and development of PACS in Europe and also had a lot of success. According to the latest analysis from Frost and Sullivan indicates PACS market in Europe reached 535 million euros (679 million dollars) in 2009 and the figure is estimated to reach 1,065 million euros (1,353 million dollars) in 2016.

Due to the differences in operating conditions and environment, PACS are different in North America, Europe, and Asia. At first, the research and development of PACS in the U.S. set huge support from the government agencies and manufacturers. In European countries, the PACS developed through the support of local organizations and multinationals. The European research teams collaborate with major manufacturers, because most of the components of the PACS systems were developed in the U.S. and Japan [5]. RIS systems are also very popular in European countries and America. RIS are widely deployed in most hospitals, and also there are many components of RIS which have been developed as an open source and are available for users to learn and use. This is a favorable condition for us to learn and develop appropriate RIS, LIS and MDL systems in Vietnam.

Needs of local hospital. According Viet-Duc hospital, every year over 1,000 cases transferred to hospital died. In many cases, if the initial emergency handling is good and in time then the patient can live. Meanwhile, local hospitals are lack of equipment and qualified physicians are limited, which are not regularly updated. So how the physicians in central hospitals and in medical universities can support as well as consult for remote medical stations. Surgical teleconsultation system is an ideal solution.

Currently, almost all the general hospitals in Vietnam were equipped with diagnostic imaging equipment with new technologies such as Computed Tomography (CT), Magnetic Resonance (MR), Computed Radiography (CR), UltraSound (US). All devices are supported by the digital data standards (DICOM, JPEG, video). However, many hospitals still work with professional film images. They still spend a part of their budget for the purchase and storage of plastic film images.

The devices operate on an independent base and they are still not connected and have no exchange between departments with different expertise. Because of that it takes so much time and efforts to collect information about patient, to perform a general diagnostic or consultations. This causes a significant reduction in speed, flexibility, accuracy and convenience in handling cases of emergency or urgent surgery. This is an essential requirement in order to improve the quality of health care for people.

The process of receiving and processing patient data in traditional medical records are still complex and time-consuming in the examination and treatment. The issue of cost of film, film storage and replication is always a burden for hospitals. To retrieve patient records on traditional medical record, film records and learn the medical history of the patient is also a difficult issue for physicians. Therefore, solutions on integrating the LIS-RIS-PACS Teleradiology will be a good remedy to help address all the above issues.

LIS-RIS-PACS Architecture at general hospital

PACS Architecture. According to the general architecture, PACS system is divided into three main classes: Class Image Devices (modalities), Class PACS server systems and Class Workstation Applications (fig. 2).

The conventional imaging devices are digital X-ray (CR) cameras, Computed Tomography (CT), UltraSound (US) and Magnetic Resonance (MR) scanners. These devices are be able to provide medical images or video sequences.

PACS Server is a core component of the system responsible for three main functions:

- Collecting data from medical imaging devices immediately after patient admission photography pictures port (PACS GATEWAY).

– Organize storage and management of medical data and other relevant information of the patient.

– Provides, coordinates the application support the examination and treatment: information function to filter, display support functions, image processing and analysis, functional diagnosis support, functional support consultation.

The client application is a machine worked by physician and other medical staff. It will typically have high requirements on the display device (size, resolution, brightness, etc.) and it will help health officers to exploit the functionality offered by the system.

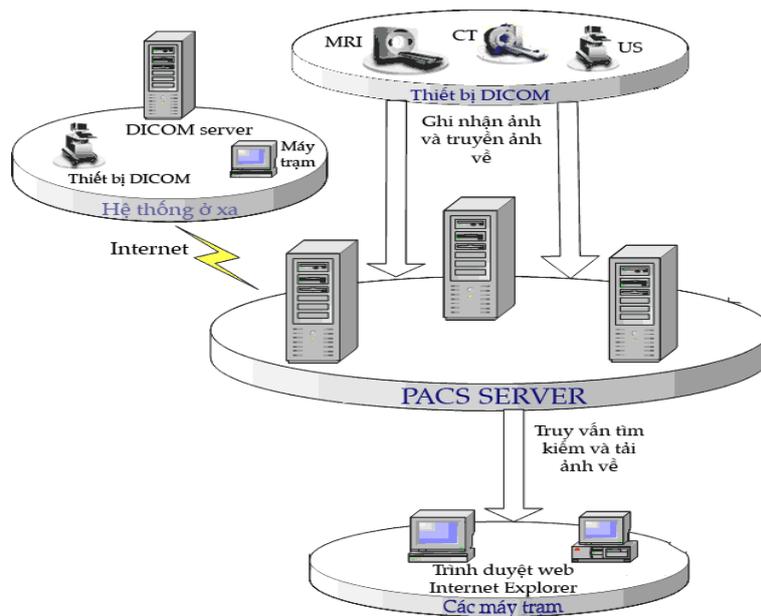


Fig. 2. Architectural model of PACS system

HIS-RIS-PACS Architecture. PACS is a system platform in the application system for the medical field. Typically, the PACS system will provide image data/video systems for medical rest. All three systems HIS, RIS and PACS need to be closely connected with each other, truly meet the requirements of the hospital and fully exploit the power that the number of medical devices can bring.

Therefore, PACS should be built as open system interfaces connected to the development of data standards in the field of medical standards like DICOM and HL7. This is useful when hospitals in developing countries synchronize three systems HIS-RIS-PACS together (fig. 3). Some hospitals develop first RIS-PACS or HIS, then continue development of remaining systems. This requires that the system must be able to communicate with each other easily, flexibly through common data standards.

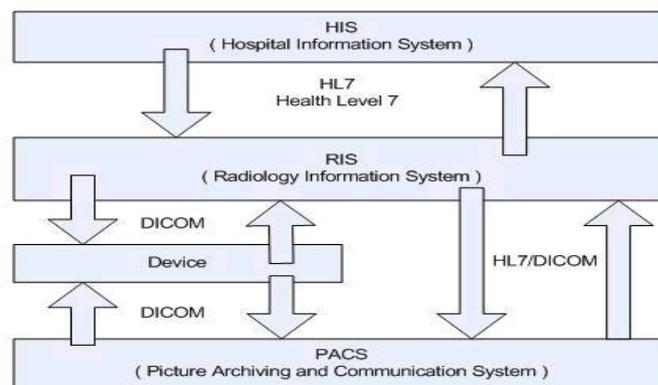


Fig. 3. The links model between HIS-RIS-PACS

Data Standards used for communication and representation. Three data standard most commonly used in the system as DICOM and HL7.

– Health Level 7 (HL7): The computer can only exchange data with each patient when they have a common communication protocol (communication environment and common language). It was developed in 1987 and HL7 has become a common protocol to be applied globally to share patient data.

– DICOM (The Digital Image and Communication in Medicine) is a standard that defines the format and rules of exchange of medical images and related information. Medical imaging is received from a device with different digital images such as CT (Computed Tomography), MR (Magnetic Resonance), US (UltraSound), NM (Nuclear Medicine). It creates a common language for communication and allows exchange of images and related information between medical devices and information systems in healthcare networks.

– LOINC (Logical Observation Identifiers Name, and Codes) is a codification for clinical laboratory values and common observation.

Flowchart of the workflow for deployment of the HIS-RIS-PACS system. PACS system was designed in conjunction with RIS to meet the work requirements of the medical staff from the stage of welcome patients, receiving and storing medical data (images, video) to the stage of support the display, processing, analysis and diagnosis of the health status of patients.

The RIS-PACS system is guaranteed to bring the convenient, fast processing for both medical staff and patients, along with historical information of patient. The biggest advantage when deploying the RIS system is the rapid speed of the response, convenience and the ability to assess the medical history of the patient. With this system, right after the hospital receive a patient, a physician immediately can review a medical history of the patient. Since then, the medical doctor gives the order to the Diagnostic Imaging Department. After photographing finished, the patient does not have to wait for receiving film and confirmation by doctors at the Department of Diagnostic Imaging. Patients do not need to bring the film back to the clinic or emergency department (fig. 4). At the clinic or emergency department, physicians can easily view, process and analyze images immediately after patients were taken. In particular, at this moment the doctors between the concerning departments can organize and conduct the diagnosis through remote consultation program. Soon there will be conclusion for the patient. This contributes to simplification of procedures for patients and shorten the process time and handle the medical emergencies.

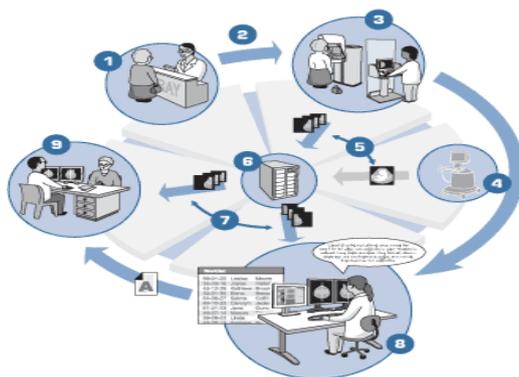


Fig. 4. Flowchart of workflow to deploying HIS-RIS-PACS systems

Medical diagnosis laboratory model. To provide a basis for the development of LIS which really meet the requirements of health care workers, a domain analysis for Electronic Health Records (EHR) has been undertaken in close cooperation of computer scientists with several domain experts [6].

Firstly, based on field studies in some hospitals, a generic hospital laboratory structure is derived and the relevant entities for MDL are identified. Secondly, hospital specific laboratory workflow problems are discussed. Sources of the information for construction of information model are document circulation and experts of laboratory. As well as in any other organization, passage of

documents through MDL is accompanied by the certain procedures of the coordination, the statement and signing of documents and the control over their origin [7].

A macro model of MDL functioning follows a certain sequence of events. First, during input, research orders and biomaterial samples are registered and brought into correspondence with each other. Next, analyses (a set of laboratory tests) are carried out automatically or manually. Then, the obtained results of these tests are passed to a requester. The following peculiarities can be outlined at this stage:

- Test results (and their dynamics) are of great diagnostic importance.
- There is significant document circulation between clinical departments and laboratories.
- There are a great number of tests to perform.
- There is an availability of efficient automatic analyzers, information from which can be transferred.
- There is a necessity to improve the reliability and quality of laboratory research.
- There is a great deal of routine work completed by laboratory employees.
- The necessity of laboratory operational statistics preparation and availability of scientific statistics.

All of these factors work together to propel the necessity to solve the problems of transferring and storing data, as well as the need to act responsibly to ensure the reliability and quality of publicly available laboratory research results. Therefore, the best solution to these problems is the use of modern IT technologies and facilities in laboratory activities.

Relations between Hospital Information System, Laboratory Information System and Radiology Information System are shown in fig. 5.

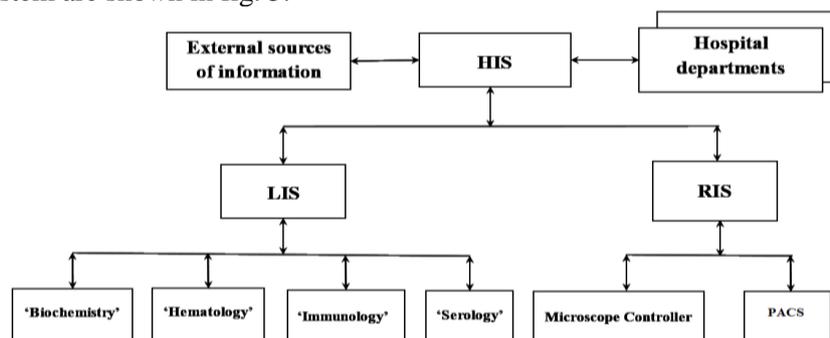


Fig. 5. Relations between Hospital Information System, Laboratory Information Systems and Radiology Information System

The ER-diagram of MDL. The database of MDL contains relevant information concerning entities and relationships in which the MDL is interested. A complete description of an entity or relationship may not be recorded in the MDL database. It is impossible (and, perhaps, unnecessary) to record every potentially available piece of information about entities and relationships [8]. From now on, we shall consider only the entities and relationships (and the information concerning them), which are to enter into the design of a database.

Entity Relationship Diagrams illustrate the logical structure of databases. An entity is an object or concept about which you want to store information. Relationships illustrate how two entities share information in the database structure.

On the basis of domain analysis, the high-level (not detailed) ER-diagram of MDL is presented in fig. 6. This diagram incorporates some of the important semantic information about the MDL. Its components are rectangles that represent entities and diamonds represent relationships.

An explanation of the model in fig. 6 is given as follows. Analysis (entity **ANALYSIS**) is a set of the laboratory tests (entity **TEST**). For example, biochemical blood analysis includes whole protein, albumin, glucose, etc.

Reference range of the test depends on the method of testing, patient's gender and age.

Result values of some tests can be verbal. Verbal values are collected in the entity **TERM**.

In some cases, entities can be self-linked. For example, tests can include other tests (relationship **component**).

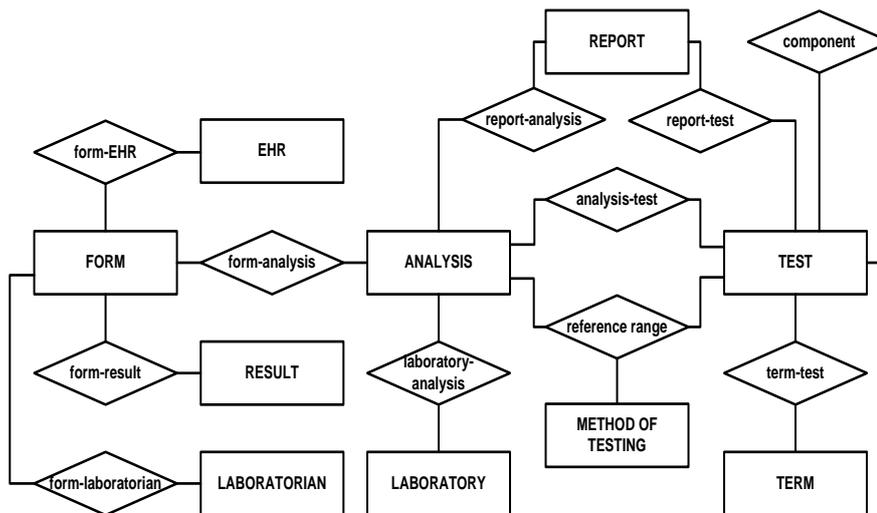


Fig. 6. The ER-diagram of Medical Diagnostic Laboratory

The entity **FORM** collects all information on analysis (patient ID, doctor, date of sampling, etc.). All patients’ analyses are contained in the entity **FORM**, which is linked to the entity **EHR** by relationship **form-EHR**. The entity **RESULT** is intended to store the results of laboratory tests. The list of laboratory employees is represented as the entity **LABORATORIAN**.

Representation of **TEST** as a separate entity lays in the basis of our approach [9]. According to this approach, the database scheme does not depend on quantity and structure of laboratory subdivisions and tests in MDL.

Implementation and results

Deployment at the common health hospitals. From the orientation of the system above, we want to introduce some typical functions of the system that we have developed and that are being implemented effectively in a number of hospitals.

TT	MÃ SỐ BN	HỌ VÀ TÊN BN	TUỔI	YÊU CẦU CHỤP	NGÀY KHÁM	GIỜ KHÁM	CHI TIẾT
	SA0610-093511	Cẩm Đào	4	SIÊU ÂM Siêu âm Tim...	22/06/2010	10.43.45	Chi tiết
	SA0610-093510	Hữu Thọ	1998	SIÊU ÂM Siêu âm Tuyế...	22/06/2010	10.43.45	Chi tiết
	SA0610-093509	Bệnh Nhân 509	15	SIÊU ÂM Siêu âm Phụ k...	22/06/2010	10.43.45	Chi tiết
	SA0610-093508	Bệnh Nhân 509	15	SIÊU ÂM Siêu âm Tuyế...	22/06/2010	10.43.45	Chi tiết
	SA0610-093508	Bệnh Nhân 508 102	16T	SIÊU ÂM Siêu âm Phụ k...	22/06/2010	10.43.45	Chi tiết
	SA0610-093508	Bệnh Nhân 508 102	16T	SIÊU ÂM Siêu âm Tuyế...	22/06/2010	10.43.45	Chi tiết
	SA0610-093507	Bệnh Nhân 507 101	11T	SIÊU ÂM Siêu âm Sản p...	22/06/2010	10.43.45	Chi tiết
	SA0610-093498	Rô Côn	18	SIÊU ÂM Siêu âm Đụng ...	22/06/2010	10.43.45	Chi tiết

Fig. 7. Program manager for examination and treatment at the Diagnosis Department

Program manager for examination and treatment at the Department of Diagnostic Imaging is shown in fig. 7. It connects information flows between: patient reception area, Examination Department, Emergency Department, Finance and Accounting Department, Diagnostic Imaging Department [10].

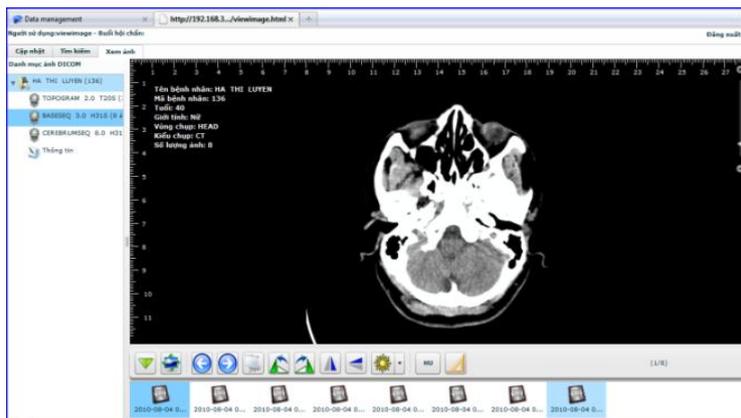


Fig. 8. Specialized software for medical image filtering and visualization

The software extracts patient information, search, view detailed photos/video of patient diagnosis (fig. 8). Moreover, the software also supports operations change the image size, image change direction, increase or decrease the contrast, changing the light/dark, HU factor analysis, measuring the size of the damaged region, recording results and conclusions.

The user interface of the software for management and tracking data, management of patient data related to diagnostic imaging, user management, management of information services for imaging diagnostics is depicted on the fig. 9.

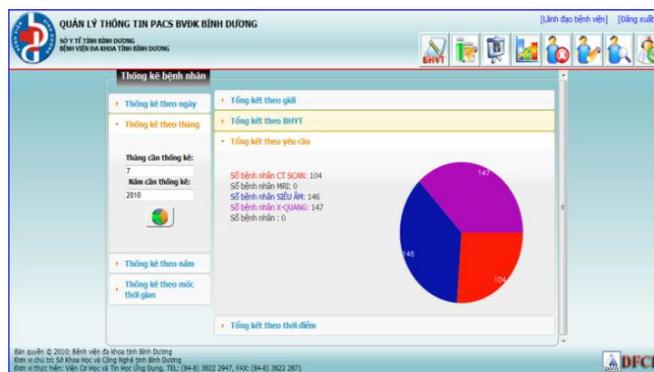


Fig. 9. The UI of the software for management and tracking patients' data

The software supports activities at the Department of Diagnostic Ultrasound images (fig. 10). It allows the storage, management, display the ultrasound images and video.



Fig. 10. The software for ultrasound consultations

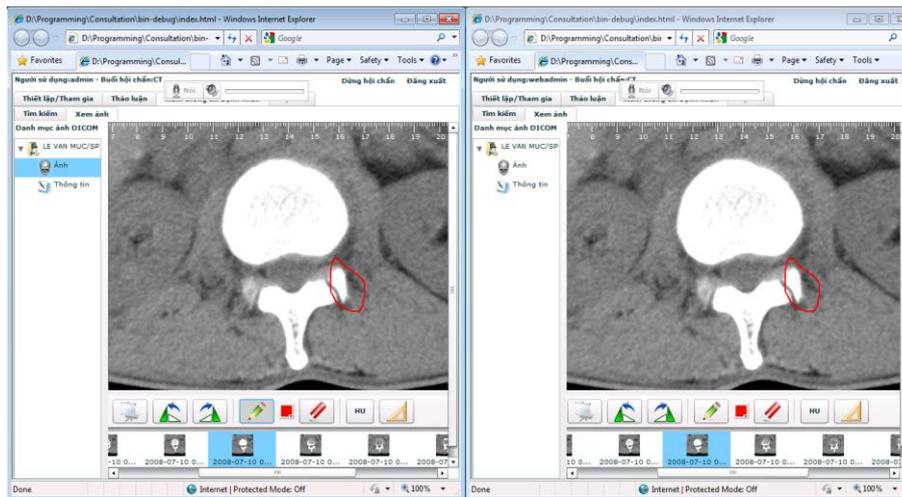


Fig. 11. The software for remote consultation program based on medical imaging

The software for organization and management of remote medical consultation based on medical images is presented in fig. 11. It allows to perform synchronization and coordination of activities during consultation. The end users of the consultation infrastructure can communicate with each other by means of chat window, voice communication and screen sharing (image synchronization with the tools to create focal points, objects localize etc.).

MDL and tuberculosis prevention. In tropical countries the most serious illness besides cancer is tuberculosis. Implementing Medical Diagnostics Laboratory on the Data Center will support all the hospitals which have not been implemented for their local PACS there. We have suggested the HCM City Public Health Association to survey all the hospitals which participated to the Tuberculosis Prevention Program then implemented Cloud PACS Server at the Data Center to receive all the images of Patients in remote hospitals (fig. 12) [11].

At present, there is an emphasis on sharing clinical and research data. Beyond the obvious reasons of facilitating patient care outside of single institution model, it is believed that new knowledge about diseases and their treatment will only come about through team science and pooling of data, computational resources, and domain expertise [12]. Imaging with its emergent role in all areas of healthcare, has been a particular so as to establish normalized datasets for research and validation studies [13].

PACS Server and PACS Gateway have been implemented at Pham Ngoc Thach Hospital to carry out the Tuberculosis Prevention Program of HCM City Public Health Association.

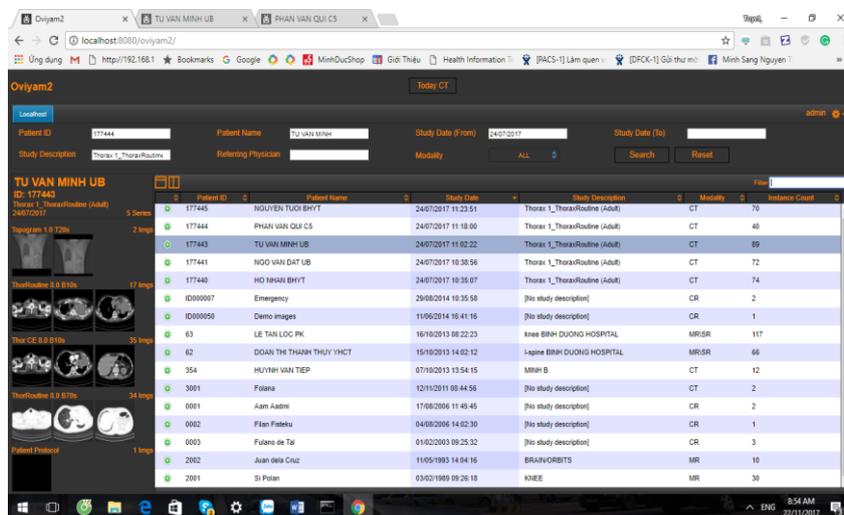


Fig. 12. The Medical doctor and the person in the TPP Program have an access to the database

The chest radiology is often the first examination to suggest pulmonary pathology. More advanced imaging techniques such as CT, MR, and ultrasound are typically used to better characterize respiratory pathology because they provide cross-sectional perspective and eliminate the superimposition of structures that is characteristic of projection imaging [14]. CT is the most commonly used advanced imaging technique to further characterize pulmonary parenchymal, pulmonary vascular, and pleural pathology, owing to its high spatial resolution and the high native contrast of aerated lung relative to soft tissue (fig. 13 and fig. 14).



Fig. 13. An example of a lungs CT archived in the database of PPT Program

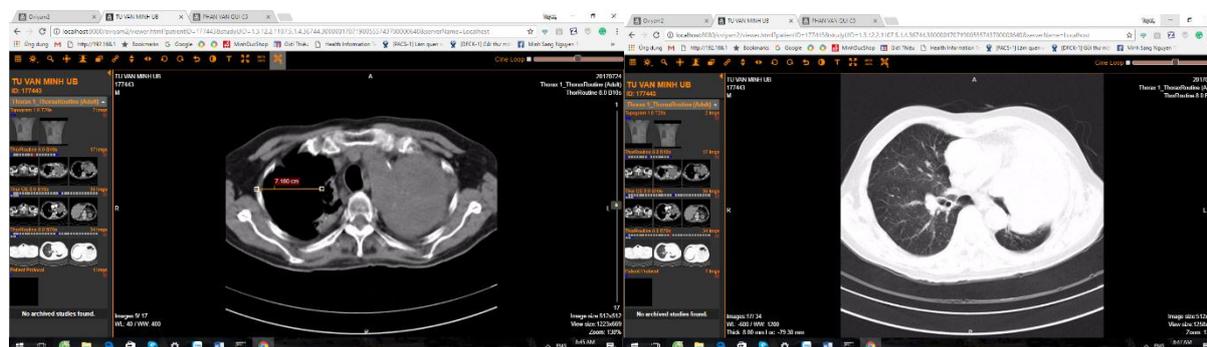


Fig. 14. Axial CT images in a patient with mild asthma show

Conclusion. Medical Imaging informatics is the rapidly evolving field combining biomedical informatics and image processing, developing and adapting core method in informatics to improve the usage and application of imaging in healthcare; and to derive new knowledge from image study. LIS-RIS-PACS-Teleradiology system improve the use of imaging throughout the medical care process. Along with the strong development of science and technology, particularly the application of information technology in practice, it can be said that better development is needed. The application of this system in hospitals will contribute greatly to improvement of the quality of healthcare in the community.

User applications of information technologies into a health sector are key directions which the Government and the Ministry of Health are very interested in. Accordingly, we continue strong investment in research and development, further improving LIS-RIS-PACS-Teleradiology system.

References

1. Bui Alex A. T., Morioka C. Information system and architectures. *Medical Imaging Informatics*. Springer, 2010, chap. 3.
2. Ablameyko S., Mozheyko D. Design and development of the public healthcare laboratory information system. *13th Intern. Convention and Fair Informatica*, 9–13 February 2009. Havana, 2009, pp. 2521–2557.

3. Mikhailov Y. Information computer technologies – an actual and inevitable step of perfection of laboratory diagnostics (by the example of creation and use of automated workgroup "Hematology"). *Clinical Laboratory Diagnostics*, 2001, vol. 7, pp. 25–32.
4. Huang H. K. Telemedicine and teleradiology. *PACS and Imaging Informatics*. John Wiley and Sons, 2010, chap. 15, pp. 454–455.
5. Tuyet D. V., Thi V. D., Thang T. C., Tuyen T. N. D., Huynh V. K., ... Hiep N. T. Extending application of PACS system and building a telemedicine system at Binh Duong General Hospital. *Proceedings of the First Asian Winter School on Intern. and Knowledge Engineering (AWSIKE 2014)*. Vung Tau, Vietnam, 2014, pp. 51–60.
6. Perreault L. E., Wiederhold G. System design and evaluation. *Medical Informatics: Computer Application in Health Care*. Addison-Wesley Publishing Company, 1990, chap. 5, pp. 151–177.
7. Guliev Ya. I., Khatkevitch M. I. Process and document in healthcare information systems. A tutorial review. *Proceedings of the Intern. Conf. Program Systems: Theory and Applications*. Pereslavl, 2004, pp. 169–187.
8. Chen P. The entity-relational model. Toward a unified view of data. *ACM TODS*, 1976, vol. 1, pp. 9–36.
9. Mozheyko D., Anishchanka A. Integrating LIS in clinical laboratories. *Healthcare IT Management*, 2006, vol. 1(1), pp. 26–29.
10. Tuyet D. V., Phu P. T., Lam T. V., Thinh P. H. Q., Huynh V. K., Nam N. P. K., Hiep N. T. Applying RFID technology in study, building patient information assessed system. *Proceedings of the First Asian Winter School on Intern. and Knowledge Engineering (AWSIKE 2014)*. Vung Tau, Vietnam, 2014, pp. 31–38.
11. Tuyet D.V., Khoa T. A., Tung N. T., Thang T. C., Hiep N. T. Applying IT in development of PACS for support diagnosis and research activity. *Proceeding of Health Science and Science and Technology Conf.* Binh Duong Province, 2013, pp. 91–96.
12. Tuyet D. V., Thi V. D., Thang T. C., Tuyen T. N. D., Huynh V. K., ... Hiep N. T. Extending application of PACS system and building telemedicine at Binh Duong and Dong Thap General Hospital. *Journal of Science and Technology*. Binh Duong University, 2014, pp. 49–63.
13. Masud M., Hossain S., Alamri A. Data interoperability and multimedia content management in e-health system. *IEEE Transactions on Information Technology in Biomedicine*, 2012, vol. 16, no. 6, pp. 1015–1023.
14. Aberle D., El-saden S., Abbona P., Gomez A., Motamedi K., ... Kangarloo H. A primer on imaging anatomy and physiology. *Medical Imaging Informatics*. Springer, 2010, chap. 2, pp. 54–55.

Information about the authors

Dao Van Tuyet – Post Graduate Student at Faculty of Applied Mathematics and Computer Science, Belarusian State University, Minsk, Republic of Belarus, Binh Duong University, Vietnam National Space Center of Vietnam Academy of Science and Technology (18, Hoang Quoc Viet, Hanoi, Vietnam). E-mail: dvtuyet@vnsc.org.vn

Sergey V. Ablameyko – D. Sc. (Technical Sciences), Academician of the National Academy of Science of Belarus, Professor at the Faculty of Mathematics and Mechanics, Belarusian State University, United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganova Str., 220012, Minsk, Republic of Belarus). E-mail: ablameyko@bsu.by

Информация об авторах

Дао Ван Туйет – аспирант Белорусского государственного университета, факультет прикладной математики и информатики, Университет Бинь Зьюнг, Вьетнамский национальный космический центр Вьетнамской академии наук и технологий (18, Хоанг Куок Вьет, Ханой, Вьетнам). E-mail: dvtuyet@vnsc.org.vn

Абламейко Сергей Владимирович – доктор технических наук, академик НАН Беларуси, профессор механико-математического факультета, Белорусский государственный университет, Объединенный институт проблем информатики Национальной академии наук Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: ablameyko@bsu.by

ISSN 1816-0301 (print)
УДК 004.31

Поступила в редакцию 30.01.2018
Received 30.01.2018

Е. В. Рыбенков, Н. А. Петровский

*Белорусский государственный университет
информатики и радиоэлектроники, Минск, Беларусь*

СИНТЕЗ FPGA-АРХИТЕКТУР БАНКОВ ФИЛЬТРОВ НА ОСНОВЕ БЛОЧНОЙ ЛЕСТНИЧНОЙ ФАКТОРИЗАЦИИ В АЛГЕБРЕ КВАТЕРНИОНОВ (ЧАСТЬ 1)

Аннотация. В настоящее время методологии проектирования систем на кристалле основываются на высокопараметризованных IP-компонентах (IP – intellectual property), которые для конкретного целевого приложения обеспечивают широкий диапазон регулировки затрат ресурсов, форматов данных арифметики с фиксированной запятой и производительности системы. В статье представлен систематический подход к синтезу FPGA-архитектур целочисленных обратимых параунитарных банков фильтров в алгебре кватернионов (Int-Q-ПУБФ) для трансформационного кодирования изображений по схеме L2L (lossless-to-lossy). Показывается, что базовым элементарным преобразованием банка фильтров является операция умножения кватернионов (Q-MUL). Блочная лестничная факторизация данной операции и распределенная арифметика на сумматорах положены в основу параметризуемого Q-MUL IP-компонента.

Ключевые слова: схема lossless-to-lossy, кватернионы, компрессия изображений, FPGA, банк фильтров, блочная лестничная факторизация

Для цитирования. Рыбенков, Е. В. Синтез FPGA-архитектур банков фильтров на основе блочной лестничной факторизации в алгебре кватернионов (часть 1) / Е. В. Рыбенков, Н. А. Петровский // Информатика. – 2018. – Т. 15, № 2. – С. 29–44.

E. V. Rybenkov, N. A. Petrovsky

Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

SYNTHESIS OF FPGA ARCHITECTURES OF BLOCK LIFTING-BASED FILTER BANKS IN QUATERNION ALGEBRA (PART 1)

Abstract. Nowadays the methodology for designing systems on a chip is based on highly parameterized IP components which provide a wide range of adjustment of costs in resources, fixed point arithmetic data formats and system performance for a specific target application. The article presents a systematic approach for synthesizing FPGA architectures of integer reversible paraunitary filter banks in quaternion algebra (Int-Q-PUBB) for L2L (lossless-to-lossy) image transformed encoding. It is shown that the basic elementary transformation of the filter bank is the operation of quaternion multiplication (Q-MUL), the block-lifting factorization of which and the distributed arithmetic on the adder are the basis of the parametrizable Q-MUL IP-component.

Keywords: lossless-to-lossy, quaternions, image compression, FPGA, filter bank, block-lifting factorization

For citation. Rybenkov E. V., Petrovsky N. A. Synthesis of FPGA architectures of block lifting-based filter banks in quaternion algebra (part 1). *Informatics*, 2018, vol. 15, no. 2, pp. 29–44 (in Russian).

Введение. В последние несколько десятилетий банки фильтров рассматриваются как наиболее эффективная техника компрессии мультимедиаданных. Они применяются в кодировании аудиосигналов, изображений и видеопотока в таких стандартах, как JPEG, JPEG2000, JPEG XR, MPEG и H.264/AVC [1]. Схема максимально децимированного M -канального банка фильтров состоит из систем анализа и синтеза, полифазная структура которых формулируется следующим образом [2, 3]:

$$\begin{aligned} [H_0(z) \ H_1(z) \ \dots \ H_{M-1}(z)]^T &= \mathbf{E}(z^M) \mathbf{e}(z)^T; \\ [F_0(z) \ F_1(z) \ \dots \ F_{M-1}(z)]^T &= \mathbf{e}(z) \mathbf{D}(z^M), \end{aligned} \quad (1)$$

где $\mathbf{e}(z) = [1 \ z^{-1} \ \dots \ z^{-(M-1)}]$, $H_k(z)$ и $F_k(z)$ – передаточные функции k -го канала банка фильтров анализа и синтеза соответственно; T (в степени) – знак транспонирования матрицы. Если матрица $\mathbf{E}(z)$ обратима, то полифазная матрица синтеза $\mathbf{D}(z)$ может быть выбрана как обратная матрица $\mathbf{E}(z)$, тогда достигается перфективная реконструкция:

$$\mathbf{D}(z)\mathbf{E}(z) = bz^{-l}\mathbf{I}, \quad b \neq 0, \quad l \geq 0. \quad (2)$$

Такой банк фильтров называется банком фильтров с перфективной реконструкцией входного сигнала или *биортогональным банком фильтров* (БОБФ). Если справедливо

$$\mathbf{E}^T(z^{-1})\mathbf{E}(z) = \mathbf{I}, \quad \mathbf{D}(z) = \mathbf{E}^T(z^{-1}), \quad (3)$$

то банк фильтров относится к специальному классу *параунитарных банков фильтров* (ПУБФ). Если длина импульсной характеристики фильтра прототипа больше числа каналов, то компоненты данных полифазных матриц являются полиномами переменной z . Двухканальный банк фильтров ($M = 2$) – это дискретное вейвлет-преобразование. Когда длина импульсной характеристики фильтра прототипа равна числу каналов M , полифазные матрицы $\mathbf{E}(z)$ и $\mathbf{D}(z)$ образуются из полиномов нулевого порядка – скалярных величин, тогда полифазные матрицы становятся эквивалентными стандартным преобразованиям. Например, восьмиканальный ($M = 8$) максимально децимированный ПУБФ ассоциируется с другим преобразованием, широко используемым в трансформационном кодировании изображений в стандарте JPEG, – дискретным косинусным преобразованием (ДКП).

Особый интерес в настоящее время представляют схемы, работающие по принципу L2L [4] и осуществляющие сжатие и восстановление мультимедиаданных как без потерь, так и с контролируемым внесением артефактов. Схема L2L унифицирует схемы кодирования изображений lossy (с потерями) и lossless (без потерь), когда требуются высокое качество и степень компрессии. Это обеспечивает возможность широкого применения схем кодирования изображений в соответствии с мультимедийными приложениями. Таким образом, задача создания универсального кодера изображений, работающего по схеме L2L, в настоящее время весьма актуальна.

Для кодирования изображения по схеме lossless требуется, чтобы трансформационное преобразование было обратимо и реализовано в целочисленной арифметике, т. е. преобразование соответствовало схеме обработки «целое-в-целое». Стандарты JPEG [5] и JPEG2000 [6] уже имеют режимы сжатия по схемам lossy и lossless: независимо осуществляются кодирование исходного изображения по схеме lossy и кодирование исходного изображения по схеме lossless. Такой подход предполагает выполнение трансформационного преобразования отдельно, что увеличивает вычислительную сложность кодера. Простейшим способом кодирования изображения по схеме L2L является применение целочисленных преобразователей, например дискретного вейвлет-преобразования 5/3 стандарта JPEG2000 или 4×8 иерархического расширенного биортогонального преобразования (HLBT) стандарта HD Photo (JPEG-XR) [7]. Однако данные трансформационные преобразователи проектировались для режима сжатия исходного изображения по схеме lossless и не соответствуют кодированию изображений по схеме lossy.

Целочисленные ДКП (IntDCT) [8, 9], которые реализованы на основе лестничной схемной параметризации, были предложены для трансформационного кодирования изображений. Тем не менее результаты [8, 9] не могут быть использованы для кодирования изображений по схеме lossless, потому что динамический диапазон представления коэффициентов лестничной схемной параметризации не контролируется и часто бывает очень большим, что требует

большой разрядности слова представления коэффициентов или реализации арифметики с фиксированной запятой переменного формата [10]. В работе [4] предложены структуры банков фильтров на основе лестничной схемной параметризации для кодирования изображения по принципу L2L, но не берется во внимание число операций округления результатов умножения в лестничной структуре, которые можно рассматривать как белый шум квантования. Число операций округления влияет на компактность энергии в субполосах банка фильтров, и получить высокую эффективность систем сжатия изображений по схеме L2L не представляется возможным. Следовательно, количество операций округления в целочисленном трансформационном преобразователе для схемы L2L компрессии изображений должно быть уменьшено в максимально возможной степени.

Таким образом, банки фильтров можно рассматривать как обобщенный преобразователь в системе трансформационного кодирования изображений, которые должны обладать свойством совершенной реконструкции, иметь линейную фазочастотную характеристику (ФЧХ), сохранять структурную регулярность и не вносить постоянную составляющую, достаточно просто отображаться на аппаратные и программные ресурсы (структурную регулярность, модульность, высокий вычислительный параллелизм, буферы памяти небольшой емкости, малую латентность и потребляемую мощность).

Максимально децимированный банк фильтров. Стандартная решетчатая факторизация ПУБФ с линейной ФЧХ. Для ПУБФ анализа с линейной ФЧХ и четного числа каналов M выбор параметров компонентов разложения решетчатой факторизации передаточной полифазной матрицы $\mathbf{E}(z)$ в (1) осуществляется следующим образом [11]:

$$\mathbf{E}(z) = \mathbf{G}_{N-1}(z)\mathbf{G}_{N-2}(z)\dots\mathbf{G}_1(z)\mathbf{E}_0 = \left(\prod_{i=N-1}^1 \mathbf{G}_i(z)\right)\mathbf{E}_0, \quad (4)$$

где за начальной матрицей

$$\mathbf{E}_0 = \frac{1}{\sqrt{2}}\Phi_0 \mathbf{W} \text{diag}(\mathbf{I}_{M/2}, \mathbf{J}_{M/2}) \quad (5)$$

следуют этапы факторизации

$$\mathbf{G}_i(z) = \frac{1}{2}\Phi_i \mathbf{W} \Lambda(z) \mathbf{W}, \quad i = 1, \dots, N-1, \quad (6)$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{I}_{M/2} & \mathbf{I}_{M/2} \\ \mathbf{I}_{M/2} & -\mathbf{I}_{M/2} \end{bmatrix}, \quad \Lambda(z) = \text{diag}(\mathbf{I}_{M/2}, z^{-1}\mathbf{I}_{M/2}).$$

Матрицы перекрестных суммирований \mathbf{W} и элементы задержки $\Lambda(z)$ фиксированные. Следовательно, все количество степеней свободы (неизвестных переменных) при синтезе ПУБФ с линейной ФЧХ определяется параметрами матрицы:

$$\Phi_i = \text{diag}(\mathbf{U}_i, \mathbf{V}_i), \quad (7)$$

где \mathbf{U}_i и \mathbf{V}_i – произвольные ортогональные матрицы размера $(M/2) \times (M/2)$. Кроме того, для упрощения обозначенной выше факторизации при $i > 0$ матрица \mathbf{U}_i может быть представлена единичной матрицей. Тогда справедливо выражение

$$\Phi_i = \text{diag}(\mathbf{I}_{M/2}, \mathbf{V}_i), \quad i > 0. \quad (8)$$

Среди ПУБФ с линейной ФЧХ необходимо выделить банки фильтров, у которых наблюдается попарная зеркальная симметрия частотных характеристик фильтров относительно $\pi/2$.

Для таких банков фильтров передаточные функции или импульсные характеристики анализирующих фильтров должны удовлетворять соответственно условиям

$$H_{M-1-k}(z) = \pm H_k(-z), \quad (9)$$

$$h_{M-1-k}(n) = \pm (-1)^n h_k(n),$$

где $k = (0, \dots, N-1)$, $n = (0, \dots, L-1)$, L – длина импульсной характеристики фильтра. Для четного числа каналов M ПУБФ с попарной зеркальной симметрией частотных характеристик фильтров может быть получен, если модифицировать факторизацию решетчатой структуры ПУБФ с линейной ФЧХ (4). Действительно, достаточно ассоциировать \mathbf{U}_i с \mathbf{V}_i в (7) и представить матрицу \mathbf{U}_i следующим образом [12]:

$$\begin{aligned} \mathbf{U}_i &= \mathbf{\Gamma} \mathbf{V}_i \mathbf{\Gamma}, \quad i = 0, \dots, N-2, \\ \mathbf{U}_{N-1} &= \mathbf{J}_{M/2} \mathbf{V}_{N-1} \mathbf{\Gamma}, \end{aligned} \quad (10)$$

где $\mathbf{J}_{M/2}$ – обратная единичная матрица; $\mathbf{\Gamma}$ – диагональная матрица, элементы которой определяются как $\gamma_{mm} = (-1)^{m-1}$, $m = 1, \dots, M/2$. Количество неизвестных переменных (степеней свободы) при синтезе данного банка фильтров сократилось. Из (10) видно, что параметры настройки банка фильтров определяются только матрицей \mathbf{V}_i . Следовательно, оптимизация коэффициентов упростилась, что явилось основной мотивацией разработки ПУБФ с линейной ФЧХ, у которых наблюдается попарная зеркальная симметрия частотных характеристик фильтров относительно $\pi/2$.

Полифазная передаточная матрица $\mathbf{D}(z)$ банка фильтров синтеза конструируется на основе обратного включения компонентов факторизации полифазной передаточной матрицы банка анализа:

$$\mathbf{D}(z) = \mathbf{E}_0 \left(\prod_{i=1}^{N-1} \mathbf{G}_i(z) \right). \quad (11)$$

При синтезе и реализации ПУБФ с линейной ФЧХ матрицы \mathbf{U}_i и \mathbf{V}_i в (7) и (10), как правило, параметризуются планарными вращениями Гивенса, каждое из которых соответствует одной степени свободы. Данный подход подобен QR-декомпозиции матриц и приводит к нескольким эквивалентным вычислительным схемам. Однако описанные выше факторизации теряют свойство параунитарности, когда матрицы \mathbf{U}_i и \mathbf{V}_i представляются с ограниченной точностью. Перфективная реконструкция не обеспечивается банком фильтров с решетчатыми структурами и фиксированной запятой, потому что наличия в разложении (4) даже одной неортогональной компоненты достаточно, чтобы условие перфективной реконструкции (3) не выполнялось [11]. Это обусловлено тем, что квантование компонент ортогональных матриц в общем случае приведет к изменению норм столбцов.

Гиперкомплексная факторизация ПУБФ анализа с линейной ФЧХ

Оператор умножения кватернионов – базовый модуль факторизации. Последние 20 лет в цифровой обработке сигналов гиперкомплексные числа – кватернионы – рассматриваются как новая парадигма, позволяющая осуществлять обработку сигналов непосредственно в многомерном домене. Кватернионы уже с успехом использовались в цифровой обработке сигналов, например в адаптивных фильтрах Калмана [13], оценке спектра по методу MUSIC [14], параунитарных банках фильтров Q-ПУБФ [15, 16] и других приложениях [17].

Алгебра кватернионов \mathbb{H} является ассоциативной некоммутативной четырехмерной алгеброй $\mathbb{H} = \{Q = q_1 + q_2i + q_3j + q_4k; q_1, q_2, q_3, q_4 \in \mathbb{R}\}$, где ортогональные мнимые части подчиняются следующим законам умножения: $i^2 = j^2 = k^2 = ijk = -1$, $ij = -ji = k$, $jk = -kj = i$, $ki = -ik = j$ [18]. Для выполнения операции умножения кватернионов в векторно-матричной алгебре кватернионы представляются в виде четырехмерного вектора, а сама операция умножения выполняется по правилу умножения вектора на матрицу. Поскольку операция умножения некоммутативна, выделяют «левый» оператор умножения $\mathbf{M}^+(Q)$ и «правый» $\mathbf{M}^-(Q)$ [16]:

$$P \cdot Q \neq Q \cdot P,$$

$$R = P \cdot Q = \mathbf{M}^+(P) \cdot \mathbf{Q} = \mathbf{M}^-(Q) \cdot \mathbf{P},$$

$$\mathbf{M}^+(P) = \begin{bmatrix} p_1 & -p_2 & -p_3 & -p_4 \\ p_2 & p_1 & -p_4 & p_3 \\ p_3 & p_4 & p_1 & -p_2 \\ p_4 & -p_3 & p_2 & p_1 \end{bmatrix}, \quad \mathbf{M}^-(Q) = \begin{bmatrix} q_1 & -q_2 & -q_3 & -q_4 \\ q_2 & q_1 & q_4 & -q_3 \\ q_3 & -q_4 & q_1 & q_2 \\ q_4 & q_3 & -q_2 & q_1 \end{bmatrix}. \quad (12)$$

Обе матрицы (12) ортогональные и имеют детерминант, равный +1, что позволяет отнести их к группе специальных матриц размерностью 4×4 , называемых $SO(4)$:

$$\mathbf{M}^\pm(Q)^{-1} = (1/|Q|) \cdot \mathbf{M}^\pm(Q)^T, \quad (13)$$

где $|Q| = \sqrt{q_1^2 + q_2^2 + q_3^2 + q_4^2}$ – норма кватерниона. Как следует из структур матриц, их свойства не зависят от значений компонентов кватерниона $\mathbf{M}^\pm(Q) \cdot \mathbf{M}^\pm(Q)^T = \mathbf{I}_4$ и, следовательно, не чувствительны к их модификации, например квантованию. Матрицы $\mathbf{M}^+(Q)$ и $\mathbf{M}^-(Q)$ соотносятся между собой как

$$\mathbf{M}^\mp(Q) = \mathbf{D}_C \mathbf{M}^\pm(Q)^T \mathbf{D}_C, \quad (14)$$

где $\mathbf{D}_C = \text{diag}(1, -\mathbf{I}_3)$ описывает в матричной нотации оператор гиперкомплексного сопряжения. Определив сопряженный кватернион $\bar{Q} = q_1 - q_2i - q_3j - q_4k$ в векторно-матричном виде как $\bar{Q} = \mathbf{D}_C \cdot \mathbf{Q}$, можно получить, что $\mathbf{M}^\pm(\bar{Q}) = \mathbf{M}^\pm(Q)^T$. Тогда выражение (14) переопределяется следующим образом: $\mathbf{M}^\mp(Q) = \mathbf{D}_C \cdot \mathbf{M}^\pm(\bar{Q}) \cdot \mathbf{D}_C$. Последнее равенство показывает, что результаты, полученные для умножения «левого» типа qx или «правого» типа xq , могут быть применимы к умножению на сопряженный кватернион. В частности, это доказывает, что вычислительная сложность операторов умножений кватернионов одинакова. Данный факт позволяет сосредоточить внимание на организации вычисления «левого» оператора умножения кватернионов $\mathbf{M}^+(Q)$.

В работе [19] утверждается, что для каждой ортогональной матрицы \mathbf{A} размерностью 4×4 существует уникальная (с точностью до знака) пара единичных кватернионов P и Q ($|P|=1$, $|Q|=1$), таких, что

$$\forall_{\mathbf{A} \in SO(4)} \exists_{P, Q \in \text{unit quat.}} \mathbf{A} = \mathbf{M}^+(P) \cdot \mathbf{M}^-(Q) = \mathbf{M}^-(Q) \cdot \mathbf{M}^+(P). \quad (15)$$

Авторы работ [15, 16] показали, что умножитель кватернионов является базовым блоком для ПУБФ, и доказали в [16], что четырех- и восьмиканальный варианты (4)–(6) могут быть модифицированы, когда матрицы U_i и V_i параметризуются операторами умножения кватернионов. При этом, в отличие от вращений Гивенса, сохраняется ортогональность (3), несмотря на квантование компонент матриц операторов умножения кватернионов. Таким образом, вариант реализации ПУБФ в арифметике с фиксированной запятой становится системой структурно без потерь.

Кватернионный восьмиканальный ПУБФ с линейной ФЧХ. Принимая во внимание (15), как показано в работах [16, 20], матрица (7) параметризуется ортогональными блочными матрицами операторов умножения кватернионов, сохраняющих свойство постоянной нормы столбцов, следующим образом:

$$\begin{aligned} \begin{bmatrix} U_i & \mathbf{0} \\ \mathbf{0} & V_i \end{bmatrix} &= \text{diag}\{M^-(S_i), M^-(S_i)\} \cdot \text{diag}\{M^-(\bar{R}_i), M^-(R_i)\} \times \\ &\times \text{diag}\{M^+(\bar{Q}_i), M^+(Q_i)\} \cdot \text{diag}\{M^+(P_i), M^+(P_i)\}, \\ \text{diag}\{M^+(Q_i), M^+(Q_i)\} W &= W \text{diag}\{M^+(Q_i), M^+(Q_i)\}, \\ \text{diag}\{M^+(Q_i), M^+(Q_i)\} \Lambda(z) &= \Lambda(z) \text{diag}\{M^+(Q_i), M^+(Q_i)\}. \end{aligned} \tag{16}$$

Следовательно, матрицы в (5) и (6) будут определены как

$$\begin{aligned} \Phi_0 &= \text{diag}\{M^-(\bar{R}_0), M^-(R_0)\} \cdot \text{diag}\{M^+(\bar{Q}_0), M^+(Q_0)\} \times \\ &\times \text{diag}\{M^-(S_0), M^-(S_0)\} \cdot \text{diag}\{M^+(P_0), M^+(P_0)\}, \\ \Phi_i &= \text{diag}\{M^-(\bar{R}_i), M^-(R_i)\} \cdot \text{diag}\{M^+(\bar{Q}_i), M^+(Q_i)\}, \quad i=1, \dots, N-1. \end{aligned} \tag{17}$$

Для $N = 3$ соответствующая структура Q-ПУБФ показана на рис. 1. К сожалению, сохранение Q-ПУБФ как структурно-ортогональной системы (системы без потерь) сопровождается многими избыточными вычислениями. Вычислительная сложность в три раза больше, чем у наиболее упрощенного варианта [21] (подробный анализ опущен для краткости). Тем не менее существуют очень эффективные VLSI- и FPGA-архитектурные решения реализации операторов умножения кватернионов [22, 23], что на практике делает разницу менее значительной.

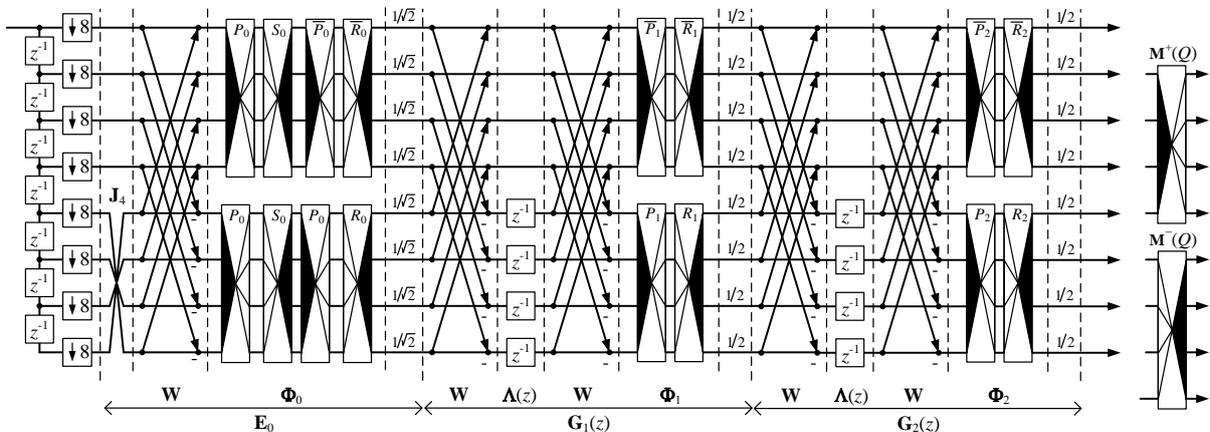


Рис. 1. Восьмиканальный Q-ПУБФ с линейной ФЧХ ($N = 3$)

Кватернионный восьмиканальный ПУБФ с попарно зеркальной симметрией частотных характеристик. Параметризация ортогональными блочными матрицами операторов умножения кватернионов восьмиканального ПУБФ с попарно зеркальной симметрией частотных характеристик (9) определяется как

$$\Phi_i = \begin{bmatrix} \Gamma \mathbf{V}_i \Gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_i \end{bmatrix} = \text{diag}\{\Gamma, \mathbf{I}\} \cdot \text{diag}\{\mathbf{M}^-(P_i), \mathbf{M}^-(P_i)\} \times \\ \times \text{diag}\{\mathbf{M}^+(Q_i), \mathbf{M}^+(Q_i)\} \cdot \text{diag}\{\Gamma, \mathbf{I}\}, \quad i = 0, \dots, N-2; \quad (18)$$

$$\Phi_{N-1} = \begin{bmatrix} \mathbf{J} \mathbf{V}_{N-1} \Gamma & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_{N-1} \end{bmatrix} = \text{diag}\{\mathbf{J}, \mathbf{I}\} \cdot \text{diag}\{\mathbf{M}^-(P_{N-1}), \mathbf{M}^-(P_{N-1})\} \times \\ \times \text{diag}\{\mathbf{M}^+(Q_{N-1}), \mathbf{M}^+(Q_{N-1})\} \cdot \text{diag}\{\Gamma, \mathbf{I}\}.$$

Структура, полученная таким образом для $N = 3$, изображена на рис. 2.

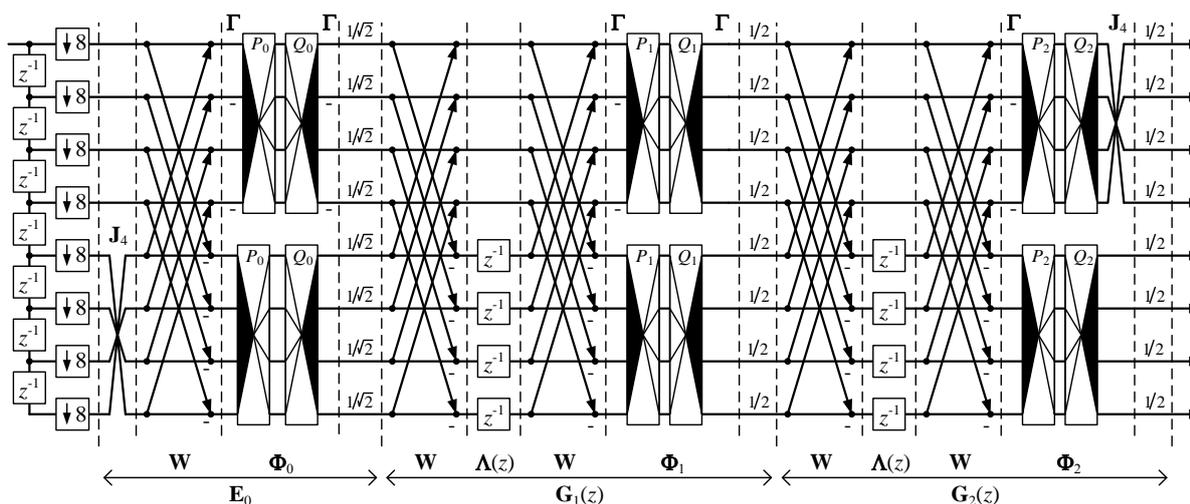


Рис. 2. Восьмиканальный Q-ПУБФ с попарно зеркальной симметрией частотных характеристик ($N = 3$)

Следует отметить, что блочные матрицы в (17) и (18) можно интерпретировать как матрицы умножения восьмимерных гиперкомплексных чисел, называемых октонионами [18]. Предложенный подход факторизации Q-ПУБФ дает возможность непосредственного выражения условия регулярности первого порядка в значениях коэффициентов решетчатой структуры Q-ПУБФ [20]:

$$Q_{N-1} = \pm \frac{1}{2} \overline{Q_{N-2}} \dots \overline{Q_0} \cdot c_1 \cdot \overline{P_0} \dots \overline{P_{N-1}} \cdot c_2, \quad (19)$$

где c_1 и c_2 – кватернионы вида $c_1 = 1 + i - j + k$, $c_2 = 0 + 0i + 0j + k$. Регулярность первого порядка позволяет локализовать постоянную составляющую на выходе банка фильтров, благодаря чему на обработанном изображении будет отсутствовать артефакт «шахматная доска». Например, для Q-ПУБФ с регулярностью первого порядка значения всех АЧХ-фильтров, кроме нулевого, стремятся к $-\infty$ на нулевой частоте [20].

Постановка задачи. Высокая вычислительная сложность Q-ПУБФ и требование работы в реальном масштабе времени для многих практических мультимедийных приложений часто обуславливают решение в виде специализированной аппаратуры. Благодаря развитию FPGA и языков описания структур [24] появилась возможность быстрого создания прототипа системы. Однако в каждой проблемной области требуется разработка предметной части структурного синтеза системы. В настоящее время методологии проектирования систем на FPGA основываются на IP-компонентах (IP – intellectual property) [25]. Высокопараметризованные IP-проекты характеризуются большей возможностью повторного использования в отличие от проектов, которые обычно оптимизированы под определенный критерий проектирования и, таким образом, являются слабопараметризованными. Такие параметры системы, как пропускная способность, затраты в ресурсах, формат данных в арифметике с фиксированной запятой, являются наиболее определяющими параметрами проекта. IP-компоненты, которые обеспечивают широкий диапазон регулировки данных параметров в процессе создания проекта FPGA, дают возможность выбора архитектуры системы, отвечающей требованиям конкретного целевого приложения.

Анализ факторизаций (16)–(18) показывает, что базовым элементарным преобразованием Q-ПУБФ является умножение кватернионов (Q-MUL), где один из сомножителей – кватернион-константа, т. е. кватернион с постоянными параметрами. При этом умножение кватернионов – ключевая операция, от эффективной реализации которой зависят характеристики всего преобразования. Следовательно, операция Q-MUL может быть выбрана в качестве параметризуемого IP-компонента. Прямое умножение матрицы на вектор потребует 16 умножений действительных чисел и 12 алгебраических сложений. В работе [26] показано, как можно вычислить продукт умножения кватернионов на основе восьми действительных умножений, но данная техника не подходит для случая, когда один операнд есть кватернион-константа. Алгоритм выполнения умножения кватернионов на арифметике с фиксированной запятой на основе лестничной структуры представлен в работе [27] (12 операций умножения на действительные числа), где основной компонент схемы – умножитель-накопитель (MAC). Аппаратная реализация MAC-операций на FPGA представляет собой специализированный блок DSP48. К достоинствам использования блоков DSP48 можно отнести универсальность и скорость работы. Однако ограниченное количество блоков DSP48 на FPGA не позволяет использовать данный подход при реализации Q-ПУБФ с длинной импульсной характеристикой или большим числом каналов. Представляют интерес вычислительные схемы умножителя кватернионов без использования устройств действительных умножений. Известно решение реализации умножителя кватернионов с использованием подхода распределенной арифметики на памяти [22], которое позволяет получать схемные решения, сбалансированные как по производительности, так и по потребляемой мощности. Однако для достижения высокой скорости и точности вычислений здесь необходим большой объем памяти и Q-ПУБФ не является обратимым преобразованием. CORDIC-лестничная параметризация позволяет интегрировать CORDIC-алгоритм «внутри» лестничной схемы умножителя, заменив действительные умножения на микровращения CORDIC-алгоритма: сложение и сдвиг. Параллельные и конвейерные структуры CORDIC-процессоров позволяют достигать высокую скорость вычислений [23], но основные их недостатки – большая латентность операции Q-MUL и затраты оборудования.

В настоящей работе предлагается гибкая технология быстрого прототипирования архитектур процессоров Int-Q-ПУБФ на основе FPGA, в основу которой положен Q-MUL IP-компонент оператора умножения кватернионов с использованием подхода распределенной арифметики на сумматорах DA-ADDER, обеспечивающий широкий диапазон регулировок параметров Int-Q-ПУБФ: амплитудных и фазочастотных характеристик, показателя эффективности кодирования, структурной регулярности, производительности, форматов данных арифметики с фиксированной запятой и затрат в ресурсах.

Блочные лестничные структурные преобразования оператора умножения кватернионов. Блочная лестничная параметризация операторов умножения кватернионов может быть представлена следующим образом [27]:

$$\mathbf{M}^+(\mathcal{Q}) = \pm \begin{bmatrix} \mathbf{C}(\mathcal{Q}) & -\mathbf{S}(\mathcal{Q}) \\ \mathbf{S}(\mathcal{Q}) & \mathbf{C}(\mathcal{Q}) \end{bmatrix} = \pm \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{F}(\mathcal{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{U}(\mathcal{Q})} \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{G}(\mathcal{Q}) & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{L}(\mathcal{Q})} \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{H}(\mathcal{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{V}(\mathcal{Q})}; \quad (20)$$

$$\mathbf{M}^-(\mathcal{Q}) = \pm \begin{bmatrix} \mathbf{C}(\mathcal{Q}) & -\mathbf{S}(\mathcal{Q}) \\ \mathbf{S}(\mathcal{Q}) & \mathbf{C}(\mathcal{Q}) \end{bmatrix} = \pm \begin{bmatrix} \mathbf{I}_2 & \mathbf{\Gamma}_2 \cdot \mathbf{H}(\mathcal{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{G}(\mathcal{Q}) \cdot \mathbf{\Gamma}_2 & \mathbf{I}_2 \end{bmatrix} \begin{bmatrix} \mathbf{I}_2 & \mathbf{\Gamma}_2 \cdot \mathbf{F}(\mathcal{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}; \quad (21)$$

$$\mathbf{C}(\mathcal{Q}) = \begin{bmatrix} q_1 & -q_2 \\ q_2 & q_1 \end{bmatrix}, \quad \mathbf{S}(\mathcal{Q}) = \begin{bmatrix} q_3 & q_4 \\ q_4 & -q_3 \end{bmatrix}. \quad (22)$$

Выражения (20) и (21) могут рассматриваться как расширение 2D-вращения на случай четырехмерного пространства. Для заданного кватерниона \mathcal{Q} и матрицы оператора умножения $\mathbf{M}^+(\mathcal{Q})$ определяется набор матричных выражений, которые могут быть решены однозначно для $\mathbf{F}(\mathcal{Q})$, $\mathbf{G}(\mathcal{Q})$, $\mathbf{H}(\mathcal{Q})$ при условии, что матрица $\mathbf{S}(\mathcal{Q})$ является несингулярной или, более определенно, ненулевой [27]:

$$\begin{aligned} \mathbf{F}(\mathcal{Q}) &= (\mathbf{C}(\mathcal{Q}) \mp \mathbf{I}_2) \mathbf{S}(\mathcal{Q})^{-1}, \\ \mathbf{G}(\mathcal{Q}) &= \mathbf{S}(\mathcal{Q}), \\ \mathbf{H}(\mathcal{Q}) &= \mathbf{S}(\mathcal{Q})^{-1} (\mathbf{C}(\mathcal{Q}) \mp \mathbf{I}_2). \end{aligned} \quad (23)$$

Поскольку матрицы $\mathbf{C}(\mathcal{Q})$, \mathbf{I}_2 , $\mathbf{S}(\mathcal{Q})$ являются ортогональными, то матрицы $\mathbf{F}(\mathcal{Q})$, $\mathbf{G}(\mathcal{Q})$, $\mathbf{H}(\mathcal{Q})$ также будут ортогональными исходя из свойств умножения ортогональных матриц [19]. Элементы матриц $\mathbf{F}(\mathcal{Q})$, $\mathbf{G}(\mathcal{Q})$, $\mathbf{H}(\mathcal{Q})$ представляют собой вещественные коэффициенты блочной лестничной параметризации. Однако динамический диапазон коэффициентов матриц $\mathbf{F}(\mathcal{Q})$, $\mathbf{G}(\mathcal{Q})$, $\mathbf{H}(\mathcal{Q})$ увеличивается, что неприемлемо для арифметики с фиксированной запятой. Приведение параметров множителя к динамическому диапазону $[-1; 1]$ может быть достигнуто, если оператор умножения кватерниона выбирать согласно правилу [27]

$$\mathbf{M}^\pm(\mathcal{Q}) = \begin{cases} \mathbf{P}_{post} \cdot \mathbf{M}^\pm(\tilde{\mathcal{Q}}) \cdot \mathbf{P}_{pre}, & \text{если } \det(\mathbf{P}) = 1, \\ \mathbf{P}_{post} \cdot \mathbf{M}^\mp(\tilde{\mathcal{Q}}) \cdot \mathbf{P}_{pre}, & \text{если } \det(\mathbf{P}) = -1, \end{cases} \quad (24)$$

где \mathbf{P} – матрица перестановки, которая связывает \mathcal{Q} и $\tilde{\mathcal{Q}} = \mathbf{P} \cdot \mathcal{Q}$. Матрицы \mathbf{P}_{pre} и \mathbf{P}_{post} являются матрицами пред- и постобработки, состоящей в перестановке и, возможно, в изменении знаков компонентов кватерниона \mathcal{Q} , например $\mathcal{Q} = q_1 + q_2i + q_3j + q_4k$ и $\tilde{\mathcal{Q}} = q_4 + q_2i - q_1j + q_3k$: $\mathcal{Q}\mathbf{x} = \mathbf{M}^\pm(\mathcal{Q})\mathbf{x} = \mathbf{P}_{post} \mathbf{M}^\pm(\tilde{\mathcal{Q}}) \mathbf{P}_{pre} \mathbf{x}$. Обе эти операции имеют одинаковую вычислительную сложность. На основе соотношения между левым и правым операторами умножения кватернионов (14), матриц перестановок \mathbf{P}_{pre} и \mathbf{P}_{post} , а также коэффициентов блочной лестничной параметризации $\mathbf{F}(\mathcal{Q})$, $\mathbf{G}(\mathcal{Q})$, $\mathbf{H}(\mathcal{Q})$ левого оператора умножения $\mathbf{M}^+(\mathcal{Q})$ определяются правила модификации операторов умножения $\mathbf{M}^+(\bar{\mathcal{Q}})$, $\mathbf{M}^-(\mathcal{Q})$ и $\mathbf{M}^-(\bar{\mathcal{Q}})$ (табл. 1).

Таблица 1

Правила модификации операторов умножения кватерниона $\mathbf{M}^\pm(Q)$

Целевой оператор умножения $\mathbf{M}^\pm(Q)$	Правило модификации оператора для \tilde{Q}
$\mathbf{M}^+(Q)$	$\tilde{\mathbf{P}}_{pre} = \mathbf{P}_{pre}; \tilde{\mathbf{P}}_{post} = \mathbf{P}_{post};$ $\tilde{\mathbf{F}} = \mathbf{F}; \tilde{\mathbf{G}} = \mathbf{G}; \tilde{\mathbf{H}} = \mathbf{H}$
$\mathbf{M}^+(\bar{Q})$	$\tilde{\mathbf{P}}_{pre} = \mathbf{P}_{post}^T; \tilde{\mathbf{P}}_{post} = \mathbf{P}_{pre}^T;$ $\tilde{\mathbf{F}} = -\mathbf{H}; \tilde{\mathbf{G}} = -\mathbf{G}; \tilde{\mathbf{H}} = -\mathbf{F}$
$\mathbf{M}^-(Q)$	$\tilde{\mathbf{P}}_{pre} = (\mathbf{P}_{post}^T) \mathbf{D}_c; \tilde{\mathbf{P}}_{post} = \mathbf{D}_c (\mathbf{P}_{pre}^T);$ $\tilde{\mathbf{F}} = -\mathbf{H}; \tilde{\mathbf{G}} = -\mathbf{G}; \tilde{\mathbf{H}} = -\mathbf{F}$
$\mathbf{M}^-(\bar{Q})$	$\tilde{\mathbf{P}}_{pre} = \mathbf{P}_{pre} \cdot \mathbf{D}_c; \tilde{\mathbf{P}}_{post} = \mathbf{D}_c \cdot \mathbf{P}_{post};$ $\tilde{\mathbf{F}} = \mathbf{F}; \tilde{\mathbf{G}} = \mathbf{G}; \tilde{\mathbf{H}} = \mathbf{H}$

Таким образом, структура обратимого оператора умножения кватернионов на основе блочной лестничной схемной параметризации представляет собой каскадное соединение блочных матриц $\mathbf{F}(Q)$, $\mathbf{G}(Q)$, $\mathbf{H}(Q)$ и их инверсное включение с отрицательным знаком (обратное преобразование – умножение на сопряженный кватернион \bar{Q}) [27]. При этом ошибки округления соответствующих результатов умножений взаимно компенсируются секциями прямого и обратного преобразований. Входной вектор \mathbf{x} проходит три ступени обработки $\mathbf{F}(Q)$, $\mathbf{G}(Q)$, $\mathbf{H}(Q)$ для формирования произведения $\mathbf{M}^+(Q) \cdot \mathbf{x}$ (рис. 3).

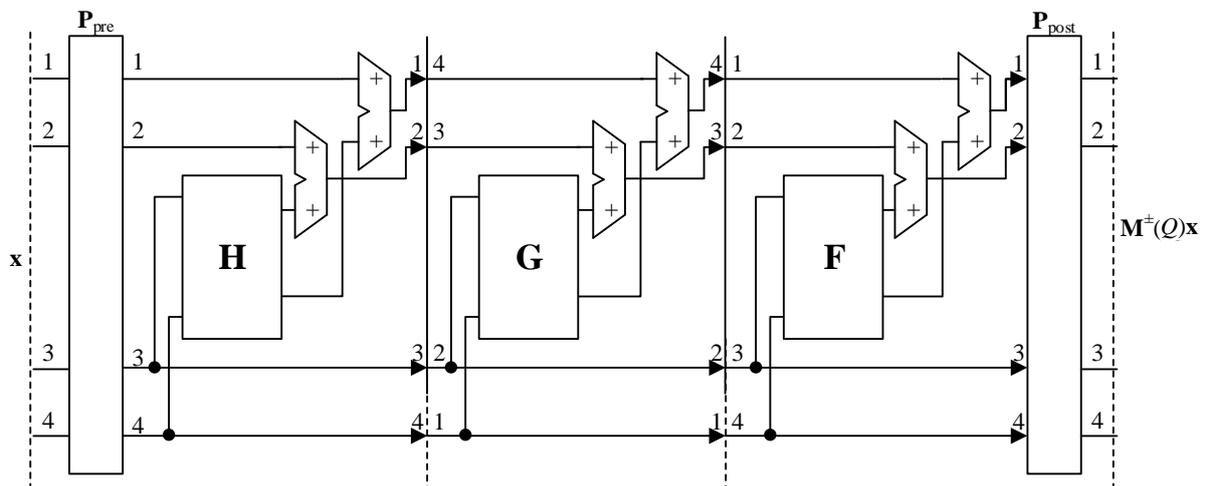


Рис. 3. Структурная схема Q-MUL IP-компонента для левого оператора умножения кватернионов (прямое преобразование)

Умножение в блоках \mathbf{H} , \mathbf{G} , \mathbf{F} может быть реализовано на основе 2D-CORDIC-алгоритмов [23] или распределенной арифметики на памяти (DA-ROM) [22]. Конвейерные структуры 2D-CORDIC-процессоров и параллельные реализации умножения в блоках на DA-ROM позволяют достигать высокой скорости вычислений, но основной их недостаток – большая латентность операции Q-MUL (2D-CORDIC) и затраты оборудования (2D-CORDIC и DA-ROM). Другим альтернативным подходом к реализации Q-MUL IP-компонента может быть распределенная

арифметика на сумматорах DA-ADDER [28]. В основе DA-ADDER лежат последовательные побитные операции над константным вектором коэффициентов при переменном входном векторе. DA-ADDER для уменьшения вычислений используют тот факт, что константный вектор содержит биты, равные нулю. Кроме того, можно выделить общие термы при суммировании [29].

Конвейерная схема Q-MUL IP-компонента. Параметризация. Процесс получения результата умножения входного кватерниона $\mathbf{x} = [x_1, x_2, x_3, x_4]^T$ на кватернион-константу Q разделяется на пять ступеней:

$$\begin{aligned} \mathbf{y} &= Q \cdot \mathbf{x} = \mathbf{M}^+(Q) \cdot \mathbf{x} = \mathbf{P}_{post} \cdot \mathbf{M}^+(\tilde{Q}) \cdot \mathbf{P}_{pre} \cdot \mathbf{x} = \\ &= \mathbf{P}_{post} \cdot \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{F}(\tilde{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{U}(\tilde{Q})} \cdot \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{G}(\tilde{Q}) & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{L}(\tilde{Q})} \cdot \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{H}(\tilde{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{V}(\tilde{Q})} \cdot \mathbf{P}_{pre} \cdot \mathbf{x}. \end{aligned} \quad (25)$$

Две ступени перестановки входных переменных \mathbf{x} и формирования результата умножения \mathbf{y} (матрицы \mathbf{P}_{pre} и \mathbf{P}_{post} соответственно) обусловлены модификацией исходного кватерниона константы Q и соответствующими модификациями матриц $\mathbf{V}(\tilde{Q})$, $\mathbf{L}(\tilde{Q})$, $\mathbf{U}(\tilde{Q})$. С целью упрощения блоков перестановки данных (матрицы \mathbf{P}_{pre} , \mathbf{P}_{post} в (25) и $\Gamma_{M/2}$ в (18)) операция изменения знака выносится и предлагается включить ее внутрь лестничной структуры, т. е. в соответствующие ступени Q-MUL IP-компонента $\mathbf{V}(\tilde{Q})$, $\mathbf{L}(\tilde{Q})$, $\mathbf{U}(\tilde{Q})$, выполнив выходные сумматоры каждой ступени конвейерной схемы (см. рис. 3) в виде сумматоров (вычитателей), управление которыми осуществляется соответствующими векторами знаков $\mathbf{SH} = [SH_1, SH_2]^T$, $\mathbf{SG} = [SG_1, SG_2]^T$, $\mathbf{SF} = [SF_1, SF_2]^T$:

$$\mathbf{y} = (\mathbf{B}_{post} \cdot \mathbf{L}) \cdot \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{F}(\tilde{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{U}(\tilde{Q})} \cdot \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ \mathbf{G}(\tilde{Q}) & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{L}(\tilde{Q})} \cdot \underbrace{\begin{bmatrix} \mathbf{I}_2 & \mathbf{H}(\tilde{Q}) \\ \mathbf{0} & \mathbf{I}_2 \end{bmatrix}}_{\mathbf{V}(\tilde{Q})} \cdot (\mathbf{R} \cdot \mathbf{B}_{pre}) \cdot \mathbf{x}, \quad (26)$$

где матрицы перестановок определяются как

$$\mathbf{B}_{pre} = \left| \mathbf{P}_{pre} \right|, \quad \mathbf{B}_{post} = \left| \mathbf{P}_{post} \right|, \quad (27)$$

а соответствующие матрицы знаков как

$$\mathbf{R} = \mathbf{P}_{pre} \cdot \mathbf{B}_{pre}^{-1} = \text{diag}(r_1, r_2, r_3, r_4), \quad r_i \in (-1, 1), \quad (28)$$

$$\mathbf{L} = \mathbf{B}_{post}^{-1} \cdot \mathbf{P}_{post} = \text{diag}(l_1, l_2, l_3, l_4), \quad l_i \in (-1, 1).$$

Таким образом, функциональная схема Q-MUL IP-компонента (рис. 3) оператора умножения кватернионов $\mathbf{M}^+(P_1)$ параметризована следующим набором параметров: матрицами перестановок \mathbf{B}_{pre} и \mathbf{B}_{post} , векторами знаков \mathbf{SH} , \mathbf{SG} , \mathbf{SF} , а также соответствующими структурами деревьев сумматоров блоков распределенной арифметики $\mathbf{F}(\tilde{Q})$, $\mathbf{G}(\tilde{Q})$, $\mathbf{H}(\tilde{Q})$ (рис. 4).

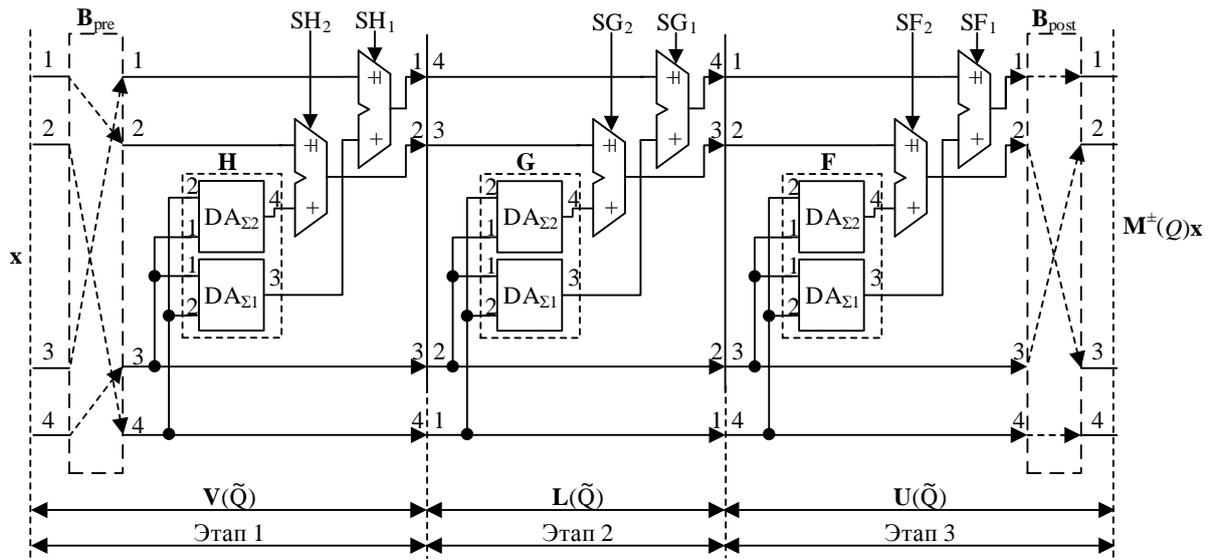


Рис. 4. Функциональная схема Q-MUL IP-компонента оператора $\mathbf{M}^+(\tilde{Q})$ умножения кватернионов

Продукт первого этапа обработки (рис. 4) согласно (26) определяется следующим образом:

$$\mathbf{y}_V = \begin{bmatrix} 1 & 0 & h_{11} & h_{12} \\ 0 & 1 & h_{21} & h_{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_1 & 0 & 0 & 0 \\ 0 & r_2 & 0 & 0 \\ 0 & 0 & r_3 & 0 \\ 0 & 0 & 0 & r_4 \end{bmatrix} \mathbf{B}_{pre} \cdot \mathbf{x} = \mathbf{V}(\tilde{Q}) \cdot \mathbf{R} \cdot \mathbf{B}_{pre} \cdot \mathbf{x}.$$

После умножения справа матрицы знаков \mathbf{R} на $\mathbf{H}(\tilde{Q})$ степени факторизации $\mathbf{V}(\tilde{Q})$ формируется вектор знаков $\mathbf{SH}(SH_1 = r_1, SH_2 = r_2)$ и определяется изменение знаков коэффициентов $\mathbf{H}(\tilde{Q})$:

$$\mathbf{y}_V = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r_3 & 0 \\ 0 & 0 & 0 & r_4 \end{bmatrix} \begin{bmatrix} SH_1 & 0 & r_3 \cdot h_{11} & r_4 \cdot h_{12} \\ 0 & SH_2 & r_3 \cdot h_{21} & r_4 \cdot h_{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{B}_{pre} \cdot \mathbf{x}.$$

Выходной продукт Q-MUL IP-компонента оператора $\mathbf{M}^+(\tilde{Q})$ формируется как результат умножения \mathbf{y}_L выхода второго этапа $\mathbf{L}(\tilde{Q})$ на $\mathbf{F}(\tilde{Q})$ степени факторизации $\mathbf{U}(\tilde{Q})$:

$$\mathbf{y} = \mathbf{B}_{post} \cdot \begin{bmatrix} l_1 & 0 & 0 & 0 \\ 0 & l_2 & 0 & 0 \\ 0 & 0 & l_3 & 0 \\ 0 & 0 & 0 & l_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & f_{11} & f_{12} \\ 0 & 1 & f_{21} & f_{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{y}_L = \mathbf{B}_{post} \cdot \mathbf{L} \cdot \mathbf{U}(\tilde{Q}) \cdot \mathbf{y}_L.$$

Компоненты $SF_1 = l_1$, $SF_2 = l_2$ вектора знаков \mathbf{SF} и изменения знаков коэффициентов $\mathbf{F}(\tilde{Q})$ определяются из выражения

$$\mathbf{y} = \mathbf{B}_{post} \cdot \begin{bmatrix} SF_1 & 0 & \left(\frac{l_1}{l_3}\right) \cdot f_{11} & \left(\frac{l_1}{l_4}\right) \cdot f_{12} \\ 0 & SF_2 & \left(\frac{l_2}{l_3}\right) \cdot f_{21} & \left(\frac{l_2}{l_4}\right) \cdot f_{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & l_3 & 0 \\ 0 & 0 & 0 & l_4 \end{bmatrix} \cdot \mathbf{y}_L. \quad (29)$$

Таким образом, результат умножения входного кватерниона \mathbf{x} на оператор умножения $\mathbf{M}^+(\tilde{Q})$ формируется как результат перемножения следующих матриц:

$$\mathbf{y} = \mathbf{B}_{post} \cdot \begin{bmatrix} SF_1 & 0 & \left(\frac{l_1}{l_3}\right) \cdot f_{11} & \left(\frac{l_1}{l_4}\right) \cdot f_{12} \\ 0 & SF_2 & \left(\frac{l_2}{l_3}\right) \cdot f_{21} & \left(\frac{l_2}{l_4}\right) \cdot f_{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & l_3 & 0 \\ 0 & 0 & 0 & l_4 \end{bmatrix} \cdot \mathbf{L}(\tilde{Q}) \times \quad (30)$$

$$\times \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r_3 & 0 \\ 0 & 0 & 0 & r_4 \end{bmatrix} \cdot \begin{bmatrix} SH_1 & 0 & r_3 \cdot h_{11} & r_4 \cdot h_{12} \\ 0 & SH_2 & r_3 \cdot h_{21} & r_4 \cdot h_{22} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \mathbf{B}_{pre} \cdot \mathbf{x}.$$

Компоненты SG_1 и SG_2 вектора знаков \mathbf{SG} , а также изменение знаков коэффициентов $\mathbf{G}(\tilde{Q})$ находятся в результате следующей модификации этапа $\mathbf{L}(\tilde{Q})$ разложения (26):

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & l_3 & 0 \\ 0 & 0 & 0 & l_4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ g_{11} & g_{12} & 1 & 0 \\ g_{21} & g_{22} & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & r_3 & 0 \\ 0 & 0 & 0 & r_4 \end{bmatrix} = \quad (31)$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ l_3 \cdot g_{11} & l_3 \cdot g_{12} & l_3 \cdot r_3 & 0 \\ l_4 \cdot g_{21} & l_4 \cdot g_{22} & 0 & l_4 \cdot r_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ l_3 \cdot g_{11} & l_3 \cdot g_{12} & SG_1 & 0 \\ l_4 \cdot g_{21} & l_4 \cdot g_{22} & 0 & SG_2 \end{bmatrix},$$

где $SG_1 = l_3 \cdot r_3$ и $SG_2 = l_4 \cdot r_4$. В табл. 2 приведены формулы расчета изменения знаков лестничных коэффициентов для множителя кватернионов $\mathbf{M}^+(\tilde{Q})$. Аналогичным образом можно получить правила изменения знаков матриц разложения $\mathbf{V}(\tilde{Q})$, $\mathbf{L}(\tilde{Q})$, $\mathbf{U}(\tilde{Q})$ для правого операнда умножения кватернионов $\mathbf{M}^-(Q)$ и матрицы $\mathbf{\Gamma}_{M/2}$ (18).

Таблица 2

Правила изменения знаков коэффициентов лестничной структуры оператора умножения кватернионов $\mathbf{M}^+(Q)$

Оператор умножения	Лестничные коэффициенты умножителя кватернионов
$\mathbf{M}^+(Q)$	$\mathbf{SH} = [SH_1, SH_2]^T = [r_1, r_2]; \mathbf{SG} = [SG_1, SG_2]^T = [l_3 \cdot r_3, l_4 \cdot r_4]; \mathbf{SF} = [SF_1, SF_2]^T = [l_1, l_2],$ $\tilde{f}_{11} = l_1 \cdot f_{11}; \tilde{f}_{21} = l_2 \cdot f_{21}; \tilde{f}_{12} = l_1 \cdot f_{12}; \tilde{f}_{22} = l_2 \cdot f_{22};$ $\tilde{g}_{11} = l_3 \cdot g_{11}; \tilde{g}_{21} = l_4 \cdot g_{21}; \tilde{g}_{12} = l_3 \cdot g_{12}; \tilde{g}_{22} = l_4 \cdot g_{22};$ $\tilde{h}_{11} = r_3 \cdot h_{11}; \tilde{h}_{21} = r_3 \cdot h_{21}; \tilde{h}_{12} = r_4 \cdot h_{12}; \tilde{h}_{22} = r_4 \cdot h_{22}$
$\Gamma_4 \cdot \mathbf{M}^+(Q)$	$\mathbf{SH} = [SH_1, SH_2]^T = [r_1, r_2]; \mathbf{SG} = [SG_1, SG_2]^T = [l_3 \cdot r_3, -l_4 \cdot r_4]; \mathbf{SF} = [SF_1, SF_2]^T = [l_1, -l_2];$ $\tilde{f}_{11} = l_1 \cdot f_{11}; \tilde{f}_{21} = -l_2 \cdot f_{21}; \tilde{f}_{12} = l_1 \cdot f_{12}; \tilde{f}_{22} = -l_2 \cdot f_{22};$ $\tilde{g}_{11} = l_3 \cdot g_{11}; \tilde{g}_{21} = -l_4 \cdot g_{21}; \tilde{g}_{12} = l_3 \cdot g_{12}; \tilde{g}_{22} = -l_4 \cdot g_{22};$ $\tilde{h}_{11} = r_3 \cdot h_{11}; \tilde{h}_{21} = r_3 \cdot h_{21}; \tilde{h}_{12} = r_4 \cdot h_{12}; \tilde{h}_{22} = r_4 \cdot h_{22}$
$\mathbf{M}^+(Q) \cdot \Gamma_4$	$\mathbf{SH} = [SH_1, SH_2]^T = [r_1, -r_2]; \mathbf{SG} = [SG_1, SG_2]^T = [l_3 \cdot r_3, -l_4 \cdot r_4]; \mathbf{SF} = [SF_1, SF_2]^T = [l_1, l_2];$ $\tilde{f}_{11} = l_1 \cdot f_{11}; \tilde{f}_{21} = l_2 \cdot f_{21}; \tilde{f}_{12} = l_1 \cdot f_{12}; \tilde{f}_{22} = l_2 \cdot f_{22};$ $\tilde{g}_{11} = l_3 \cdot g_{11}; \tilde{g}_{21} = l_4 \cdot g_{21}; \tilde{g}_{12} = l_3 \cdot g_{12}; \tilde{g}_{22} = l_4 \cdot g_{22};$ $\tilde{h}_{11} = r_3 \cdot h_{11}; \tilde{h}_{21} = r_3 \cdot h_{21}; \tilde{h}_{12} = -r_4 \cdot h_{12}; \tilde{h}_{22} = -r_4 \cdot h_{22}$

Заключение. Предложенная функциональная схема Q-MUL IP-компонента является универсальной, поскольку изменение векторов знаков \mathbf{SH} , \mathbf{SG} , \mathbf{SF} , матриц \mathbf{B}_{post} , \mathbf{B}_{pre} и коэффициентов блока распределенной арифметики на сумматорах \mathbf{H} , \mathbf{G} , \mathbf{F} позволяет, не меняя структуру умножителя $\mathbf{M}^+(Q)$, вычислять все используемые модификации операторов умножителей кватернионов в Int-Q-ПУБФ. Таким образом, внесение операции изменения знаков входного кватерниона внутрь блочной лестничной структуры позволило унифицировать VHDL-описание структуры Q-MUL IP-компонента.

Список использованных источников

1. Rao, K. Techniques and Standards for Image, Video, and Audio Coding / K. Rao, J. Hwang. – Prentice Hall, 1996. – 563 p.
2. Vaidyanathan, P. P. Multirate Systems and Filter Banks / P. P. Vaidyanathan. – Englewood Cliffs, NJ : Prentice Hall, 1993. – 911 p.
3. Крот, А. М. Теория анализа и синтеза бэнк-фильтров и их применение / А. М. Крот, В. О. Кудрявцев // Успехи современной радиоэлектроники. – 1999. – № 2. – С. 3–17.
4. Suzuki, T. Lifting-based paraunitary filterbanks for lossy/lossless image coding / T. Suzuki, Y. Tanaka, M. Ikehara // Proc. of Intern. Conf. EUSIPCO. – Florence, 2006. – Vol. 6.
5. Pennebaker, W. B. JPEG: Still Image Compression Standard / W. B. Pennebaker, J. L. Mitchell. – N. Y. : Van Nostrand Reinhold, 1993. – 639 p.
6. Taubman, D. JPEG2000: Image Compression Fundamentals, Standards, and Practice / D. Taubman, M. Marcellin. – Boston : Kluwer Academic Publishers, 2002. – 777 p.
7. ISO/IEC 29199-2:2012. Information technology – JPEG XR image coding system. – Part 2: Image coding specification. – 2012.
8. Chen, Y. J. Integer discrete cosine transform (IntDCT) / Y. J. Chen, S. Oraintara, T. Nguyen // Proc. of 2nd Intern. Conf. ICICS. – Sydney, 1999. – Vol. 99.
9. Liang, J. Fast multiplierless approximations of the DCT with the lifting scheme / J. Liang, T.D. Tran // IEEE Trans. Signal Process. – 2001. – Vol. 49, no. 12. – P. 3032–3044.
10. Petrovsky, A. Dynamic reconfigurable on the lifting steps wavelet packet processor with frame-based psychoacoustic optimized time-frequency tiling for real-time audio applications / A. Petrovsky, M. Rodionov, Al. Petrovsky // Design and Architectures for Digital Signal Processing. – Vienna : InTech, 2013. – P. 3–30.

11. Lattice structure for regular paraunitary linear-phase filterbanks and m -band orthogonal symmetric wavelets / S. Orantara [et. al.] // *IEEE Trans. Signal Process.* – 2001. – Vol. 49, no. 11. – P. 2659–2672.
12. Gan, L. A simplified lattice factorization for linear-phase paraunitary filter banks with pairwise mirror image frequency responses / L. Gan, K.-K. Ma // *IEEE Trans. Circuits Syst. II.* – 2004. – Vol. 51, no. 1. – P. 3–7.
13. Choukroun, D. Novel quaternion Kalman filter / D. Choukroun, I. Bar-Itzhack, Y. Oshman // *IEEE Trans. Aerosp. Electron. Syst.* – 2006. – Vol. 42, no. 1. – P. 174–190.
14. Miron, S. Quaternion-MUSIC for vector-sensor array processing / S. Miron, N. Le Bihan, J. Mars // *IEEE Trans. Signal Process.* – 2006. – Vol. 54, no. 4. – P. 1218–1229.
15. Parfieniuk, M. Quaternionic building block for paraunitary filter banks / M. Parfieniuk, A. Petrovsky // *Proc. of the 12th European Signal Processing Conf. (EUSIPCO).* – Austria, Vienna, 2004. – P. 1237–1240.
16. Парфенюк, М. Параунитарные банки фильтров на основе алгебры кватернионов: теория и применение / М. Парфенюк, А. А. Петровский // *Цифровая обработка сигналов.* – 2008. – № 1. – С. 22–36.
17. Karney, C. Quaternions in molecular modeling / C. Karney // *J. Molecular Graphics and Modelling.* – 2007. – Vol. 25, no. 5. – P. 595–604.
18. Кантор, И. Л. Гиперкомплексные числа / И. Л. Кантор, А. С. Солодовников. – М. : Наука, 1973. – 145 с.
19. Baker, H. G. Quaternions and orthogonal 4×4 real matrices / H. G. Baker [Electronic resource]. – 1996. – Mode of access: <http://archive.gamedev.net/archive/reference/articles/article428.html>. – Date of access: 05.04.2018.
20. Parfieniuk, M. Inherently lossless structures for eight- and sixchannel linear-phase paraunitary filter banks based on quaternion multipliers / M. Parfieniuk, A. Petrovsky // *Signal Process.* – 2010. – Vol. 90. – P. 1755–1767.
21. Gan, L. A simplified lattice factorization for linear-phase perfect reconstruction filter bank / L. Gan, K.-K. Ma // *IEEE Signal Process. Lett.* – 2001. – Vol. 8, no. 7. – P. 207–209.
22. Petrovsky, N. Low read-only memory distributed arithmetic implementation of quaternion multiplier using split matrix approach / N. Petrovsky, A. Stankevich, A. Petrovsky // *Electronics Letters.* – 2014. – Vol. 50, no. 24. – P. 1809–1811.
23. Петровский, Н. А. CORDIC-техника для фиксированного угла вращения в операции умножения кватернионов / Н. А. Петровский, А. В. Станкевич, А. А. Петровский // *Информатика.* – 2015. – № 4(48). – С. 85–108.
24. Бибило П. Н. Основы языка VHDL : учеб. пособие. Изд. 6-е. – М. : Книжный дом «Либроком», 2014. – 328 с.
25. A systematic approach for synthesizing VLSI architectures of lifting-based filter banks and transforms / R. Bartholoma [et al.] // *IEEE Trans. Circuits Syst. I.* – 2008. – Vol. 55, no. 7. – P. 1939–1952.
26. The complexity of the quaternion product : Rep. TR 75-245 / Cornell University. – 1975. – Mode of access: <http://www.theworld.com/~sweetser/quaternions/ps/cornellctr75-245.pdf>. – Date of access: 05.04.2018.
27. Parfieniuk, M. Quaternion multiplier inspired by the lifting implementation of plane rotations / M. Parfieniuk, A. Petrovsky // *IEEE Trans. Circuits Syst. I.* – 2010. – Vol. 57, no. 10. – P. 2708–2717.
28. Petrovsky, N. A. Design and implementation of reversible integer quaternionic paraunitary filter banks on adder-based distributed arithmetic / N. A. Petrovsky, E. V. Rybenkov, A. A. Petrovsky // *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA 2017).* – Poznan, 2017. – P. 17–22.
29. Chang, T. S. New distributed arithmetic algorithm and its application to IDCT / T. S. Chang, C. Chen, C. W. Jen // *IEE Proceedings – Circuits, Devices and Systems.* – 1999. – Vol. 146, no. 4. – P. 159–163.

References

1. Rao K., Hwang J. *Techniques and Standards for Image, Video, and Audio Coding.* Prentice Hall, 1996, 563 p.
2. Vaidyanathan P. P. *Multirate Systems and Filter Banks.* Englewood Cliffs, NJ, Prentice Hall, 1993, 911 p.
3. Krot A. M., Kudrjavcev V. O. Teoriya analiza i sinteza bjenk-fil'trov i ih primeneniye [Theory of analysis and synthesis filter banks and their application]. *Uspеhi sovremennoy radiojelektroniki [The Successes of Modern Radio Electronics]*, 1999, no. 2, pp. 3–17 (in Russian).
4. Suzuki T., Tanaka Y., Ikehara M. Lifting-based paraunitary filterbanks for lossy/lossless image coding. *Proc. of Intern. Conf. EUSIPCO.* Florence, 2006, vol. 6.
5. Pennebaker W. B., Mitchell J. L. *JPEG: Still Image Compression Standard.* N. Y., Van Nostrand Reinhold, 1993, 639 p.
6. Taubman D., Marcellin M. *JPEG2000: Image Compression Fundamentals, Standards, and Practice.* Boston, Kluwer Academic Publishers, 2002, 777 p. doi: 10.1007/978-1-4615-0799-4
7. ISO/IEC 29199-2:2012. Information technology – JPEG XR image coding system. – Part 2: Image coding specification. – 2012.
8. Chen Y. J., Orantara S., Nguyen T. Integer discrete cosine transform (IntDCT). *Proc. of 2nd Intern. Conf. ICICS.* Sydney, 1999, vol. 99.
9. Liang J., Tran T. D. Fast multiplierless approximations of the DCT with the lifting scheme. *IEEE Trans. Signal Process.*, 2001, vol. 49, no. 12, pp. 3032–3044. doi: 10.1109/78.969511
10. Petrovsky A., Rodionov M., Petrovsky A. I. Dynamic reconfigurable on the lifting steps wavelet packet processor with frame-based psychoacoustic optimized time-frequency tiling for real-time audio applications. *Design and Architectures for Digital Signal Processing.* Vienna, InTech, 2013, pp. 3–30. doi: 10.5772/51604
11. Orantara S., Tran T. D., Heller P. N., Nguyen T. Q. Lattice structure for regular paraunitary linear-phase filterbanks and m -band orthogonal symmetric wavelets. *IEEE Trans. Signal Process.*, 2001, vol. 49, no. 11, pp. 2659–2672. doi: 10.1109/78.960413

12. Gan L., Ma K.-K. A simplified lattice factorization for linear-phase paraunitary filter banks with pairwise mirror image frequency responses. *IEEE Trans. Circuits Syst. II*, 2004, vol. 51, no. 1, pp. 3–7. doi: 10.1109/TCSII.2003.821515
13. Choukroun D., Bar-Itzhack I., Oshman Y. Novel quaternion Kalman filter. *IEEE Trans. Aerosp. Electron. Syst.*, 2006, vol. 42, no. 1, pp. 174–190. doi: 10.1109/TAES.2006.1603413
14. Miron S., Bihan Le N., Mars J. Quaternion-MUSIC for vector-sensor array processing. *IEEE Trans. Signal Process.*, 2006, vol. 54, no. 4, pp. 1218–1229. doi: 10.1109/TSP.2006.870630
15. Parfieniuk M., Petrovsky A. Quaternionic building block for paraunitary filter banks. *Proc. of the 12th European Signal Processing Conf. (EUSIPCO)*. Austria, Vienna, 2004, pp. 1237–1240.
16. Parfieniuk M., Petrovsky A. Paraunitarynye banki fil'trov na osnove algebry kvaternionov: teoriya i primeneniye [Quaternion based paraunitary filter banks: theory and applications]. *Cifrovaja obrabotka signalov [Journal of Digital Signal Processing]*, 2008, no. 1, pp. 22–36 (in Russian).
17. Karney C. Quaternions in molecular modeling. *J. Molecular Graphics and Modelling*, 2007, vol. 25, no. 5, pp. 595–604.
18. Kantor I. L., Solodovnikov A. S. Giperkompleksnye chisla. *Hypercomplex Algebra*. Moscow, Science Publ., 1973, 145 p. (in Russian).
19. Baker H. G. *Quaternions and orthogonal 4x4 real matrices*, 1996. Available at: <http://archive.gamedev.net/archive/reference/articles/article428.html> (accessed 05.04.2018).
20. Parfieniuk M., Petrovsky A. Inherently lossless structures for eight- and sixchannel linear-phase paraunitary filter banks based on quaternion multipliers. *Signal Process.* 2010, vol. 90, pp. 1755–1767. doi: 10.1016/j.sigpro.2010.01.008
21. Gan L., Ma K.-K. A simplified lattice factorization for linear-phase perfect reconstruction filter bank. *IEEE Signal Process. Lett.*, 2001, vol. 8, no. 7, pp. 207–209. doi: 10.1109/97.928679
22. Petrovsky N., Stankevich A., Petrovsky A. Low read-only memory distributed arithmetic implementation of quaternion multiplier using split matrix approach. *Electronics Letters*, 2014, vol. 50, no. 24, pp. 1809–1811. doi:10.1049/el.2014.1775
23. Petrovsky N. A., Stankevich A. V., Petrovsky A. A. CORDIC-tehnika dlja fiksirovannogo ugla vrashhenija v operacii umnozhenija kvaternionov [CORDIC-techniques for fixed angle of rotation in multiplying operation of quaternions]. *Informatika [Informatics]*, 2015, no. 4(48), pp. 85–108 (in Russian).
24. Bibilo P. N. Osnovy jazyka VHDL: uchebnoe posobie. Izd. 6-e. *Fundamentals of VHDL: Textbook. Ed. 6th*. Moscow, Knizhnyj dom «Librokom» Publ., 2014, 328 p. (in Russian).
25. Bartholoma R., Greiner T., Kesel F., Rosenstiel W. A systematic approach for synthesizing VLSI architectures of lifting-based filter banks and transforms. *IEEE Trans. Circuits Syst. I*, 2008, vol. 55, no. 7, pp. 1939–1952.
26. The complexity of the quaternion product: Rep. TR 75-245 / Cornell University. – 1975. Available at: <http://www.theworld.com/~sweetser/quaternions/ps/cornellcstr75-245.pdf> (accessed 05.04.2018).
27. Parfieniuk M., Petrovsky A. Quaternion multiplier inspired by the lifting implementation of plane rotations. *IEEE Trans. Circuits Syst. I*, 2010, vol. 57, no. 10, pp. 2708–2717. doi: 10.1109/TCSI.2010.2046259
28. Petrovsky N. A., Rybenkov E. V., Petrovsky A. A. Design and implementation of reversible integer quaternionic paraunitary filter banks on adder-based distributed arithmetic. *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA 2017)*. Poznan, 2017, pp. 17–22. doi: 10.23919/SPA.2017.8166830
29. Chang T. S., Chen C., Jen C. W. New distributed arithmetic algorithm and its application to IDCT. *IEE Proceedings – Circuits, Devices and Systems*, 1999, vol. 146, no. 4, pp. 159–163. doi: 10.1049/ip-cds:19990537

Информация об авторах

Рыбенков Евгений Викторович – ассистент кафедры электронно-вычислительных систем, Белорусский государственный университет информатики и радиоэлектроники (ул. П. Бровки, 6, 220013, Минск, Республика Беларусь). E-mail: rybenkov@bsuir.by

Петровский Николай Александрович – доцент кафедры электронно-вычислительных систем, Белорусский государственный университет информатики и радиоэлектроники (ул. П. Бровки, 6, 220013, Минск, Республика Беларусь). E-mail: nick.petrovsky@bsuir.by

Information about the authors

Eugene V. Rybenkov – Assistant Professor, Department of Computer Engineering, Belarusian State University of Informatics and Radioelectronics (6, P. Brovki Str., 220013, Minsk, Republic of Belarus). E-mail: rybenkov@bsuir.by

Nick A. Petrovsky – Associate Professor, Ph. D. (Computer Science), Department of Computer Engineering, Belarusian State University of Informatics and Radioelectronics (6, P. Brovki Str., 220013, Minsk, Republic of Belarus). E-mail: nick.petrovsky@bsuir.by

ISSN 1816-0301 (print)
УДК 004.912:025.045

Паступіла ў рэдакцыю 19.01.2018
Received 19.01.2018

С. І. Лысы¹, Г. Р. Станіславенка², Ю. С. Гецэвіч¹

¹*Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі
навук Беларусі, Мінск, Беларусь*

²*Цэнтр даследаванняў беларускай культуры, мовы і літаратуры Нацыянальнай акадэміі
навук Беларусі, Мінск, Беларусь*

АЎТАМАТЫЗАВАНАЯ ГЕНЕРАЦЫЯ АЛФАВІТНА-ПРАДМЕТНАГА ПАКАЗАЛЬНІКА УНІВЕРСАЛЬНОЙ ДЗЕСЯТКОВАЙ КЛАСІФІКАЦЫІ НА БЕЛАРУСКАЙ МОВЕ

Анотацыя. Універсальная дзесятковая класіфікацыя (УДК) з'яўляецца міжнароднай класіфікацыйнай сістэмай, якая адпавядае найбольш істотным патрабаванням да класіфікацый: міжнароднасць, універсальнасць, пашыральнасць. Табліцы УДК былі перакладзены і апублікаваны цалкам ці часткова на больш чым 40 мовах, а выкарыстоўваецца УДК прыкладна ў 130 краінах свету. На тэрыторыі Беларусі УДК дзейнічае на працягу апошніх 50 гадоў. Аднак толькі ў 2016 г. з'явілася афіцыйнае выданне УДК на беларускай мове. Алфавітна-прадметны паказальнік (АПП), які складае больш за чвэрць выдання, быў падрыхтаваны пры дапамозе алгарытму, які аўтаматызаваў працэс яго стварэння.

У артыкуле падрабязна апісваецца падыход да аўтаматызаванага стварэння АПП беларускамоўнага выдання УДК. Разглядаюцца патрабаванні да АПП, параўноўваюцца фарматы АПП, выкарыстаныя ў выданнях УДК розных краін. Таксама апісваюцца электронныя лінгвістычныя рэсурсы, неабходныя для аўтаматызаванай генерацыі АПП, прыводзіцца падрабязны пакрокавы алгарытм. У якасці падцверджання працаздольнасці і карэктнасці працы дадзенага алгарытму прыводзіцца распрацаваны праграмны прататып сістэмы генерацыі АПП УДК. Дадзены прататып быў непасрэдным чынам выкарыстаны пры стварэнні АПП першага выдання Універсальнай дзесятковай класіфікацыі на беларускай мове.

Ключавыя словы: апрацоўка электронных тэкстаў, сістэмы класіфікацыі, беларуская мова, Універсальная дзесятковая класіфікацыя, алфавітна-прадметны паказальнік

Для цытавання. Лысы, С. І. Аўтаматызаваная генерацыя алфавітна-прадметнага паказальніка Універсальнай дзесятковай класіфікацыі на беларускай мове / С. І. Лысы, Г. Р. Станіславенка, Ю. С. Гецэвіч // Информатика. – 2018. – Т. 15, № 2. – С. 45–54.

S. I. Lysy¹, H. R. Stanislavenka², Yu. S. Hetsevich¹

¹*The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus*

²*Center for the Belarusian Culture, Language and Literature Researches of the National Academy
of Sciences of Belarus, Minsk, Belarus*

AUTOMATED ALPHABETIC SUBJECT INDEX GENERATION FOR UNIVERSAL DECIMAL CLASSIFICATION IN BELARUSIAN

Abstract. Universal Decimal Classification (UDC) is an international classification system that corresponds to the essential classification requirements: internationalism, universality, expansiveness. UDC Tables were translated and published fully or partially in more than 40 languages. UDC is used in nearly 130 countries. In Belarus UDC has been used for the last 50 years. But the official edition of UDC in the Belarusian language was released only in 2016. An alphabetic subject index (ASI) that makes more than a quarter of the edition was prepared with the help of an algorithm, which automated its creation.

In the main part of the article an approach to automated ASI generation for Belarusian UDC edition is described in detail. Authors review ASI requirements, compare ASI formats used in different countries. Electronic linguistic resources needed for the automated generation of ASI are described in the article. The detailed step-by-step algorithm is provided. As confirmation of efficiency and correctness of the algorithm, a program prototype of ASI generation system was developed and described in the article. This prototype was directly used in ASI creation for the first edition of UDC in the Belarusian language.

Keywords: electronic texts processing, classification systems, the Belarusian language, Universal Decimal Classification, alphabetic-subject index

For citation. Lysy S. I., Stanislavenka H. R., Hetsevich Yu. S. Automated alphabetic subject index generation for Universal Decimal Classification in Belarusian. *Informatics*, 2018, vol. 15, no. 2, pp. 45–54 (in Belarusian).

Уводзіны. У XIX ст. бельгійскія бібліёграфы Поль Атлэ (фр. *Paul Otlet*) і Анры Лафантэн (фр. *Henri La Fontaine*) дапрацавалі «Дзесятковую класіфікацыю» Мэлвіла Дзьюі (англ. *Melvil Dewey*), якую апошні ствараў для Бібліятэкі Кангрэса ў ЗША, і заклалі пачатак развіцця Універсальнай дзесятковай класіфікацыі [1].

На сённяшні дзень распрацаваная бельгійцамі класіфікацыя адпавядае найбольш істотным патрабаванням да класіфікацый – міжнароднасці і ўніверсальнасці. Міжнароднасць падцвярджаецца тым фактам, што табліцы УДК былі перакладзены і апублікаваны цалкам ці часткова на больш чым 40 мовах, а выкарыстоўваецца УДК прыкладна ў 130 краінах свету. Універсальнасць класіфікацыі заключаецца ў дзесятковым прынцыпе яе пабудовы: дзяленні аднаго класа на дзесяць (ці менш) падкласаў. Дзесятковы прынцып дазваляе практычна неабмежавана пашыраць яе шляхам дадання новых лічбаў да ўжо існуючых, не ламаючы ўсёй сістэмы ў цэлым, і з дапамогай шматлікіх сродкаў і прыёмаў індэксавання выкарыстоўваць УДК для сістэматызацыі і наступнага пошуку разнастайных крыніц інфармацыі – ад вузкатэматычных выданняў спецыяльнай дакументацыі да вялікіх галіновых і шматгаліновых даведачна-інфармацыйных фондаў. Гэта дазваляе наблізіцца да рашэння праблемы пошуку дадзеных у вялікім інфармацыйным патоку і іх сістэматызацыі.

На тэрыторыі Беларусі УДК была ўведзена ў якасці адзінай сістэмы класіфікацыі для тэхнічных бібліятэк і органаў навукова-тэхнічнай інфармацыі ў 1963 г. [2]. На працягу наступных 30 гадоў канкурэнцыю УДК стваралі такія сістэмы класіфікацыі, як ББК (Бібліятэчна-бібліяграфічная класіфікацыя) і АКЛ (Адзіная класіфікацыя літаратуры). Аднак пасля распаду СССР і атрымання Беларуссю незалежнасці УДК замацавалася ў якасці асноўнай інфармацыйна-пошукавай сістэмы ў нашай краіне. У 1993 г. Нацыянальная кніжная палата Беларусі для ўнутранага карыстання пераклала на беларускую мову і выдала скарачаныя табліцы УДК [2]. Нягледзячы на тое, што больш за дзве сотні ўстаноў у Беларусі выкарыстоўвалі і выкарыстоўваюць у сваёй працы УДК, да 2016 г. афіцыйнага выдання на беларускай мове не існавала. Аўтары дадзенага артыкула ў супрацоўніцтве з Нацыянальнай бібліятэкай Беларусі ўдзельнічалі ў падрыхтоўцы выдання перакладу табліц УДК на беларускую мову [3] і, у прыватнасці, займаліся стварэннем алфавітна-прадметнага паказальніка да УДК, які з'яўляецца неад'емнай яго часткай і важным элементам у працэсе пошуку інфармацыі па табліцах УДК.

З улікам вялікага аб'ёму табліц УДК і іх перыядычнага дапаўнення і ўдасканалення хуткае і якаснае стварэнне АПП, а таксама яго дынамічная актуалізацыя могуць быць магчымымі толькі пры дапамозе сучасных інфармацыйных тэхналогій. Па гэтых прычынах было вырашана распрацаваць алгарытм, які дазваляе з масіву запісаў УДК аўтаматызавана згенераваць масіў адпаведнасцяў «тэрміналагічная адзінка – каардынацыйны код/індэкс». Такім чынам, перад аўтарамі паўстала задача распрацоўкі алгарытму пераўтварэння тэкстаў табліц УДК у АПП. Ён традыцыйна ўяўляе сабой спіс слоў, адсартаваных у алфавітным парадку разам з індэксамі УДК, паводле якіх можна адшукаць той ці іншы тэрмін у табліцах УДК.

Дадзены алгарытм і распрацаваны на яго аснове прататып сістэмы генерацыі АПП пакліканы паспрыяць у дасягненні наступных мэт:

- стварэння першага выдання табліц УДК на беларускай мове;
- папаўнення электронных лінгвістычных рэсурсаў спецыялізаванай і вузкатэрміналагічнай лексікай, якая выкарыстоўваецца ва УДК;
- удасканалення алгарытмаў апрацоўкі электронных тэкстаў вялікіх аб'ёмаў на беларускай мове.

Паняцце алфавітна-прадметнага паказальніка і патрабаванні да яго. Алфавітна-прадметны паказальнік з'яўляецца важным складнікам сістэмы УДК і арыентаваны на выкананне каардынацыйнай і пошукавай функцыі. АПП выкарыстоўваецца для спрашчэння і паскарэння тэрміналагічнага і прадметнага пошуку ў межах асноўных табліц УДК у працэсе індэксавання літаратуры для забеспячэння максімальнай уніфікацыі і адзінства індэксавання.

АПП уяўляе сабой масіў адпаведнасцяў «тэрміналагічная адзінка – каардынацыйны код/індэкс» (мал. 1). Кожная адпаведнасць уяўляе сабой асобную прадметную рубрыку.

Noord-Polen - plaats Polen (438.5)	Samojeedse talen - taal =511.2
Noordpooleilanden, Europese - plaats (984)	San Juan (provincie) - plaats Argentinië (825.2)
Noordpoolgebied, Amerikaans - plaats (987)	San Luis (provincie) - plaats Argentinië (825.5)
Noordpoolgebied, Russisch - plaats (985)	San Marino (Republiek) - plaats Europa (454.4)
Noordpoolgebieden (98)	sancties - economisch recht 346.9
Noords skiën - sport 796.922	sandelhout - botanie 582.728
Noordse talen	sandwiches - koken - huishoudkunde 641.84
- linguïstiek 811.113	Sanima - taal =855.72
- literatuur 821.113	sanitair
- taal =113	- bouwnijverheid 696
Noordse volkeren, gebieden van - plaats (368)	- huishoudkunde 644.6
noordwest - plaats (1-16)	sanitair in gebouwen - gezondheidstechniek 628.6
noordwest Polen - plaats Polen (438.4)	sanitair papier - papierindustrie 676.25
Noordwest-China - plaats China (514)	Sanskriet
noordwestelijke Grote Oceaan - plaats (265.5)	- literatuur 821.211
Nooristani - taal =23	- taal =211
Noors	Santa Catarine (deelstaat) - plaats Brazilië (816.4)
- literatuur 821.113.5	Santa Cruz (provincie) - plaats Argentinië (828.8)
- taal =113.5	Santa Fe (provincie) - plaats Argentinië (821.6)
Noorse - linguïstiek 811.113.5	Santalales - botanie 582.728

Мал. 1. Фрагменты алфавітна-прадметнага паказальніка галандскага выдання УДК [4]

Варта адзначыць, што фармат галандскага АПП не з'яўляецца адзіным магчымым. Напрыклад, А. А. Сербін у артыкуле [5] прыводзіць наступны спіс спосабаў падачы прадметных рубрык:

1. Каардынацыйная рубрыка тлумачальнага характару, дзе тлумачэнне тэрміну асвятляецца з пункту гледжання вобласці прымянення.

Прыклад: *Ацэталі (хімія) 547–316*

2. Каардынацыйная рубрыка простаі формы семантычнага прадстаўлення.

Прыклад: *Электроніка 621.3*

3. Каардынацыйная рубрыка аспектнага характару, у якой указваецца аспект прымянення дадзенага тэрміну.

Прыклад:

Кракадзілы:

– *(Заалогія) 598.14*

– *(Палеазаалогія) 568.14*

– *(Паляванне) 639.14*

4. Каардынацыйна-арыентаваная рубрыка з аспектна-прадметнай дэталізацыяй, дзе асвятляецца аспектнасць і прадметнасць аспекта дадзенага тэрміну.

Прыклад:

Тэрмадынаміка

– *Атмасферы (метэаралогія) 551.511.33*

– *Біялагічных працэсаў (біяфізіка) 577.31*

– *Зямлі (геафізіка) 550.36*

– *Хімічная (фіз. хімія) 544.3*

Па разглядзе фарматаў АПП, выкарыстаных у галандскім [4] і рускім выданнях УДК, для беларускамоўнага выдання было вырашана выкарыстоўваць каардынацыйныя рубрыкі аспектнага характару для слоў, якія адносяцца адначасова да некалькіх абласцей прымянення, ва ўсіх жа астатніх выпадках – падаваць каардынацыйныя рубрыкі тлумачальнага характару (мал. 2).

Жалезабетон (<i>будаўніцтва</i>) 693.5	Снарад
Жалоба (<i>антрапалогія</i>) 393.7	(<i>фізіка</i>) 531.55
Жаніцьба (<i>антрапалогія</i>) 392.5	(<i>тэхналогіі</i>) 623.4, 623.46, 629.76, 629.762
Жанр	Снег
(<i>мастацтва</i>) 7.041	(<i>геалагічныя навукі</i>) 551.322, 551.578.4, 556.12
(<i>кіно</i>) 791.2, 791.22, 791.223	(<i>тэхналогіі</i>) 624.14
(<i>літаратура</i>) 82-1/-9, 82-94, 82-95	(<i>прамысловасць</i>) 685.36
Жанчына	Снегіровыя сойкі (<i>заалогія</i>) 598.293
(<i>антрапалогія</i>) 392.6	Снэпшот (<i>фатаграфія</i>) 77.055
(<i>медыцына</i>) 618.16, 618.17, 618.18	Совападобныя (<i>заалогія</i>) 598.27, 598.279, 598.279.4
Жарганізм (<i>мовазнаўства</i>) 81`373.48	Сода (<i>хімія</i>) 661.3
Жарт (<i>антрапалогія</i>) 398.25, 398.6	Сойка (<i>заалогія</i>) 598.293
Жаўтушка (<i>заалогія</i>) 598.296	Сок (<i>хімія</i>) 547.914, 663.8, 663.81
Жах (<i>кіно</i>) 791.221.9	Сокалападобныя (<i>заалогія</i>) 598.27, 598.279, 598.279.3
Жвачныя	Солад (<i>хімія</i>) 663.032
(<i>заалогія</i>) 599.735	Соль
(<i>сельская гаспадарка</i>) 636.2	(<i>хімія</i>) 54-38, 549.451.1, 549.89, 661.74, 661.8, 664.41
Жвір (<i>будаўніцтва</i>) 691.22	(<i>геалагічныя навукі</i>) 552.53, 553.63, 553.78
	(<i>прыкладное мастацтва</i>) 745.56

Мал. 2. Фрагменты алфавітна-прадметнага паказальніка беларускага выдання УДК [3]

Для карэктнай працы алгарытму аўтаматызаванай генерацыі АПП і паслядоўнага выканання патрабаванняў да АПП было неабходна распрацаваць адмысловыя лінгвістычныя рэсурсы.

Электронныя лінгвістычныя рэсурсы для аўтаматызаванай генерацыі АПП. АПП з'яўляецца алфавітна ўпарадкаваным спісам паняццяў, якія сустракаюцца ў класах асноўных табліц УДК. Гэтыя паняцці прыводзяцца ў пачатковай форме разам з кодам адпаведных класаў УДК. Для карэктнага функцыянавання алгарытму аўтаматызаванай генерацыі АПП было неабходна атрымаць або распрацаваць адмысловыя лінгвістычныя і тэматычныя рэсурсы, сярод якіх:

- спіс стоп-класаў (класаў УДК, якія не выкарыстоўваюцца пры фарміраванні АПП);
- спіс стоп-слоў (слоў, прысутных у табліцах УДК, якія не павінны выкарыстоўвацца пры фарміраванні АПП);
- спіс адпаведнасцяў «код класа – дамен»;
- слоўнік лексікі, які змяшчае ўсе словы, якія сустракаюцца ва УДК, з іх пачатковымі формамі, пазнакамі націску і граматычнымі характарыстыкамі.

У той час, як спіс стоп-класаў быў атрыманы ад Кансорцыума УДК (*UDC Consortium*) (мал. 3, а), астатнія тры рэсурсы былі адмыслова распрацаваны. У спіс стоп-слоў увайшлі словы, якія маюць агульны характар і не валодаюць асаблівай тэрміналагічнай спецыфікай. Гэта такія словы, як «асноўны», «агульны», «другі», «кожны» і інш. Спіс адпаведнасцяў «код класа – дамен» прызначаны для фарміравання рубрык аспектнага і тлумачальнага характару, якія прадугледжваюць указанне вобласці прымянення таго ці іншага тэрміну. Дадзены спіс таксама фарміраваўся ўручную і складаўся са скарачанага кода класа і адпаведнай вобласці прымянення, або дамена (мал. 3, б).

7.049	72 (архітэктура)
7.06	73 (пластычнае мастацтва)
7.08	74 (прыкладное мастацтва)
7.094	75 (жывапіс)
711.1	76 (графіка)
72.01/.05	77 (фатаграфія)
725	78 (музыка)
725.22	79 (забавы)
725.89	80 (мовазнаўства)
726.9	81 (мовазнаўства)
728.8	82 (літаратура)
736	

а)

б)

Мал. 3. Фрагменты спісаў: а) стоп-класаў; б) адпаведнасцяў «код класа – дамен»

Для атрымання электроннага слоўніка лексікі УДК былі здзейснены наступныя крокі:

1. Складанне спісу ўнікальных слоў асноўных табліц УДК з іх кантэкстамі пры дапамозе вэб-сэрвіса «Падлік частотнасці слоў» (<http://corpus.by/WordFrequencyCounter>, <http://ssrlab.by/1457>).

2. Вылучэнне слоў, якія адсутнічаюць у даступных электронных граматычных слоўніках беларускай мовы. Для гэтай мэты быў выкарыстаны інтэрнэт-сэрвіс «Праверка правапісу» (<http://corpus.by/SpellChecker>, <http://ssrlab.by/3334>).

3. Ручное складанне спісу адпаведнасцяў «слова – пачатковая форма» для слоў, адсутнічаючых у наяўных граматычных слоўніках (мал. 4, а).

4. Ручная пастаноўка націскаў у складзеным спісе (мал. 4, б).

5. Удасканаленне слоўніка: генерацыя парадыгмаў слоў па пачатковай і некалькіх ускосных формах з атрыманням граматычных характарыстык (<http://corpus.by/WordParadigmGenerator>, <http://ssrlab.by/5047>) (мал. 4, в).

2248	Жанчына-вайсковец	3357	кулі=к-саро+ка	5044	+
2249	Жанчыны-вайскоўцы	3358	кулікі=-саро+кі	5045	Во=дападрыхто+ўка_NNIFO
2250	+	3359	+	5046	Во=дападрыхто+ўкі_NNIFG
2251	Жанчына-маці	3360	куліна+рна	5047	Во=дападрыхто+ўцы_NNIFD
2252	Жанчыны-маці	3361	+	5048	Во=дападрыхто+ўку_NNIFA
2253	+	3362	культу=рна-выхава+ўчы	5049	Во=дападрыхто+ўкаі_NNIFI
2254	Жанчына-юрыст	3363	культу=рна-выхава+ўчае	5050	Во=дападрыхто+ўкаю_NNIFS
2255	Жанчыны-юрысты	3364	+	5051	Во=дападрыхто+ўцы_NNIFR
2256	+	3365	культу=рна-мо+ўны	5052	Во=дападрыхто+ўкі_NNIFPO
2257	жаўтушка	3366	культу=рна-мо+ўныя	5053	Во=дападрыхто+вак_NNIFPG
2258	жаўтушкі	3367	культу=рна-мо+ўных	5054	Во=дападрыхто+ўкам_NNIFPD
2259	+	3368	+	5055	Во=дападрыхто+ўкі_NNIFPA
2260	жужаль	3369	кумі+н	5056	Во=дападрыхто+ўкамі_NNIFPI
2261	жужалі	3370	+	5057	Во=дападрыхто+ўках_NNIFPR
2262	+	3371	кумкава+т	5058	+
2263	Жужуй	3372	+	5059	Во=жыкападо+бны_JJMO
2264	+	3373	куні=цападо+бны	5060	Во=жыкападо+бнага_JJMG
2265	жук-алень	3374	куні=цападо+бныя	5061	Во=жыкападо+бнаму_JJMD
2266	жукі-алені	3375	+	5062	Во=жыкападо+бны_JJMA
2267	+	3376	ку=рападо+бны	5063	Во=жыкападо+бнага_JJMU
2268	жук-насарог	3377	ку=рападо+бныя	5064	Во=жыкападо+бным_JJMI
2269	жукі-насарогі	3378	+	5065	Во=жыкападо+бным_JJMR
				5066	Во=жыкападо+бная_JJFO

а)

б)

в)

Мал. 4. Фрагменты распрацаваных слоўнікаў: а) адпаведнасці «слова – пачатковая форма»;

б) адпаведнасці «слова – пачатковая форма» з указаннем націскаў;

в) поўныя парадыгмы слоў з указаннем націскаў і граматычных характарыстык

Атрыманыя электронныя лінгвістычныя рэсурсы зрабілі магчымай не толькі тэарэтычную распрацоўку алгарытму аўтаматызаванай генерацыі АПП УДК на беларускай мове, але і яго практычную рэалізацыю ў адмысловым праграмным прататыпе.

Алгарытм аўтаматызаванага стварэння алфавітна-прадметнага паказальніка УДК на беларускай мове. Алгарытм дае магчымасць сфарміраваць АПП УДК па тэкстам яе табліц і арыентаваны на працу з беларускім выданнем УДК, але аддзеленасць лінгвістычных і тэматычных рэсурсаў, якія выкарыстоўваюцца ў алгарытме, ад самога алгарытму дазваляе адаптаваць яго для працы з УДК на іншых мовах. Дадзены алгарытм заснаваны на падыходах, апісаных аўтарамі ў артыкулах [6, 7].

Уваходныя дадзеныя алгарытму: поўны тэкст асноўных табліц УДК T_{UDC} .

Рэсурсы алгарытму:

– тэкставы файл F_{sc} , які змяшчае спіс стоп-класаў;

– тэкставы файл F_{sw} , які змяшчае спіс стоп-слоў;

– тэкставы файл F_{dic} , які змяшчае характэрную для УДК лексіку з указаннем пэўных граматычных характарыстык;

– тэкставы файл F_{dom} , які змяшчае спіс адпаведнасцяў «код класа – дамен».

Уваход:

Крок 1. Загрузка рэсурсаў. Адбываецца загрузка файлаў F_{sc} , F_{sw} , F_{dic} , F_{dom} з адмысловымі рэсурсамі ў памяць камп'ютара, фарміруюцца адпаведныя спісы.

Крок 1.1. Фарміраванне спісу стоп-класаў. Адбываецца загрузка файла са спісам стоп-класаў F_{sc} . Фарміруецца спіс $L_{sc} = \langle sc_1, \dots, sc_A \rangle$, дзе sc_a – a -ты стоп-клас, $a = 1, \dots, A$.

Крок 1.2. Фарміраванне спісу стоп-слоў. Адбываецца загрузка файла са спісам стоп-слоў F_{sw} . Фарміруецца спіс $L_{sw} = \langle sw_1, \dots, sw_B \rangle$, дзе sw_b – b -е стоп-слова, $b = 1, \dots, B$.

Крок 1.3. Фарміраванне спецыялізаванага слоўніка. Адбываецца загрузка файла-слоўніка F_{dic} . Фарміруецца спіс $L_{dic} = \langle \langle w_1, wa_1, wc_1, wi_1 \rangle, \dots, \langle w_C, wa_C, wc_C, wi_C \rangle \rangle$, дзе w_c – c -е слова слоўніка, wa_c – націск c -га слова слоўніка, wc_c – катэгорыя c -га слова слоўніка, wi_c – пачатковая форма c -га слова слоўніка, $c = 1, \dots, C$.

Крок 1.4. Фарміраванне спісу прыналежнасці класаў даменам. Адбываецца загрузка файла F_{dom} . Фарміруецца спіс $L_{dom} = \langle \langle cl_1, dom_1 \rangle, \dots, \langle cl_D, dom_D \rangle \rangle$, дзе cl_d – d -ты клас спісу, dom_d – дамен, які адпавядае d -му класу спісу, $d = 1, \dots, D$.

Крок 2. Фарміраванне спісу класаў УДК. Уваходны тэкст T_{UDC} разбіваецца на асобныя запісы – класы УДК; у кожным запісе вылучаецца код класа і апісанне класа. Такім чынам на аснове ўваходнага тэксту T_{UDC} фарміруецца спіс $L_{UDC} = \langle Cl_1, \dots, Cl_N \rangle$, дзе $Cl_n = \langle Not_n, Cap_n \rangle$, Not_n – n -ты код класа, Cap_n – n -е апісанне класа, $n = 1, \dots, N$.

Крок 3. Апрацоўка спісу класаў УДК. Ствараецца спіс адпаведнасцяў «слова – мноства кодаў класаў» L_{res} , у які будучы заносіцца вынікі апрацоўкі. Кожны элемент Cl_n спісу L_{UDC} праходзіць крокі 3.1–3.6.

Крок 3.1. Фільтрацыя паводле спісу стоп-класаў. Адбываецца вызначэнне, ці прыналежаць код класа Not_n спісу стоп-класаў L_{sc} . Калі прыналежнасць выяўлена, то адбываецца пераход да наступнага элементу Cl_{n+1} і кроку 3.1 (пры $n = N$ – да кроку 4). Іначай – да кроку 3.2.

Крок 3.2. Вылучэнне слоў. У апісанні класа Cap_n вылучаюцца ўсе сімвальныя паслядоўнасці, якія адпавядаюць шаблону будовы слова Pt_w . Выкарыстоўваючы сінтаксіс рэгулярных выказаў PCRE (<http://www.pcre.org/original/doc/html/pcrepattern.html>), дадзены шаблон можна прадставіць наступным чынам:

$$Pt_w = [set1][set1set2]^*$$

дзе $set1$ – мноства сімвалаў, з якіх можа пачынацца слова, $set2$ – мноства сімвалаў, з якіх можа складацца, але не можа пачынацца слова, $set1set2 = set1 \cup set2$. У склад мноства $set1$ уваходзяць літары беларускага алфавіту, у склад мноства $set2$ – злучок, апостраф, сімвалы націскаў і інш.

Вылучаныя ў апісанні класа Cap_n словы заносзяцца ў спіс $W_{cap} = \langle wrd_1, \dots, wrd_M \rangle$, дзе wrd_m – m -е вылучанае слова, $m = 1, \dots, M$.

Крок 3.3. Нармалізацыя слоў. Кожнае слова wrd_m са спісу W_{cap} прыводзіцца да нармалізаванай электроннай формы (слова прыводзіцца да ніжняга рэгістра; адбываецца ўніфікацыя апострафаў, замена літары «ў» на «у» ў пачатку слова і інш.). Вынікам дадзенай апрацоўкі з'яўляецца спіс нармалізаваных слоў $W'_{cap} = \langle wrd'_1, \dots, wrd'_M \rangle$, дзе wrd'_m – m -е нармалізаванае слова, $m = 1, \dots, M$.

Крок 3.4. Фільтрацыя паводле спісу стоп-слоў. Для кожнага слова са спісу W'_{cap} адбываецца вызначэнне, ці прыналежаць яно спісу стоп-слоў L_{sw} . Калі прыналежнасць не выяўлена, то слова заносіцца ў спіс $W''_{cap} = \langle wrd''_1, \dots, wrd''_K \rangle$, дзе wrd''_k – k -е дапушчальнае слова, $k = 1, \dots, K$.

Крок 3.5. Фільтрацыя паводле часціны мовы і прывядзенне да пачатковай формы. Для кожнага слова са спісу W''_{cap} адбываецца вызначэнне часціны мовы wc і пачатковай формы wi паводле спісу характэрнай для УДК лексікі L_{dic} . Калі бягучае слова з'яўляецца назоўнікам, прыметнікам або дзеепрыметнікам, то яго пачатковая форма wi заносіцца ў спіс $W'''_{cap} = \langle wrd'''_1, \dots, wrd'''_J \rangle$, дзе wrd'''_j – j -е дапушчальнае слова ў пачатковай форме, $j = 1, \dots, J$.

Крок 3.6. Занясенне ў выніковы спіс. Кожнае слова са спісу W'''_{cap} заносіцца ў выніковы спіс L_{res} . Калі ў выніковым спісе L_{res} бягучае слова яшчэ не сустракаецца, то слова дадаецца ў выглядзе новага элемента, у адпаведнасць якому ставіцца масіў з адным элементам – кодам

апрацоўваемага класа Not_n . Калі ж слова ўжо занесена ў выніковы спіс L_{res} , то код апрацоўваемага класа Not_n заносіцца ў масіў кодаў класаў, адпаведных апрацоўваемаму слову.

Крок 4. Прысвойванне даменаў. Адбываецца перабор элементаў спісу L_{res} . У кожным элеменце L_{res} да кожнага кода класа вызначаецца тэматычны дамен dom паводле спісу L_{dom} . Коды класаў, дамены якіх супадаюць, групуюцца ў адзін масіў.

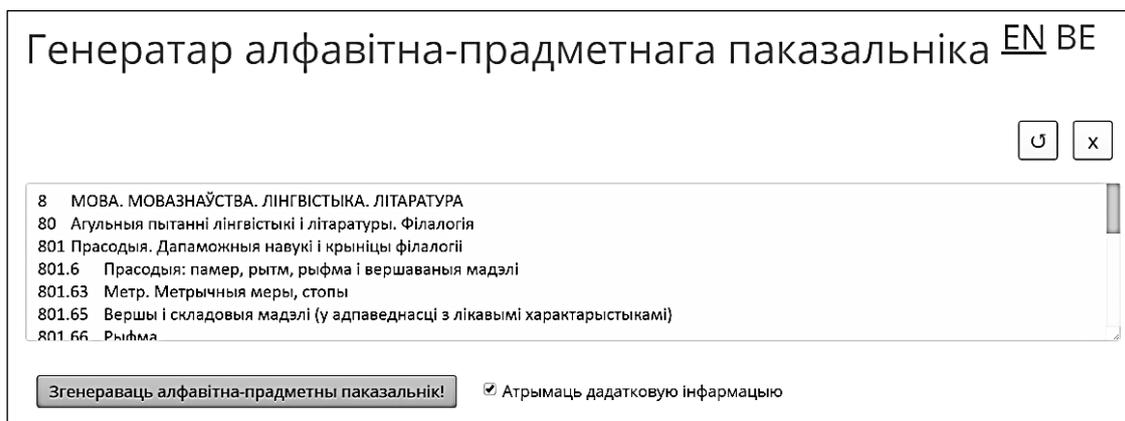
Такім чынам, вынікаем з'яўляецца спіс $L_{res} = \langle \langle wrd_1, ent_1 \rangle, \dots, \langle wrd_R, ent_R \rangle \rangle$, дзе wrd_r – r -е слова выніковага спісу, $r = 1, \dots, R$; у сваю чаргу, $ent_r = \langle \langle dom_1, Lnot_1 \rangle, \dots, \langle dom_s, Lnot_s \rangle \rangle$, дзе dom_s – s -ты дамен, да якога адносіцца слова wrd_r , $s = 1, \dots, S$. І ўрэшце $Lnot_s = \langle not_1, \dots, not_T \rangle$, дзе not_t – t -ты код класа, адпаведнага дамену dom_s , да якога адносіцца слова wrd_r , $t = 1, \dots, T$.

Крок 5. Фарміраванне выніку. Адбываецца сартыроўка выніковага спісу L_{res} паводле беларускага алфавіту і прывядзенне да тэкставага фармату з неабходным фарматаваннем. Вынік T_{ASI} выводзіцца на экран і захоўваецца ў адпаведным файле на серверы.

Канец алгарытму.

У выніку выканання апісанага алгарытму на аснове табліц УДК на беларускай мове фарміруецца АПП. Варта адзначыць, што прапанаваны алгарытм з'яўляецца прыдатным для пашырэння на іншыя мовы, бо лінгвістычныя і тэматычныя рэсурсы, выкарыстаныя ў ім, адзелены ад алгарытму і могуць быць распрацаваны для іншай мовы.

Прататып сістэмы генерацыі алфавітна-прадметнага паказальніка УДК. Для апрабачы, тэставання і нагляднасці працаздольнасці апісанага вышэй алгарытму быў распрацаваны прататып сістэмы генерацыі АПП УДК у форме вэб-сэрвіса «Генератар алфавітна-прадметнага паказальніка», які даступны для вольнага выкарыстання ў Інтэрнэце па адрасе <http://corpus.by/AlphabeticalSubjectIndexGenerator>. Дадзены сэрвіс дае магчымасць канвертаваць тэкст табліц УДК у АПП (мал. 5).



Мал. 5. Інтэрнэт-сэрвіс «Генератар алфавітна-прадметнага паказальніка»

Для карэктнай працы вэб-сэрвіса табліцы УДК павінны быць пададзены ў наступным фармаце: класы УДК адзелены адзін ад аднаго пераводам радка, у межах аднаго запісу код класа адзелены ад апісання класа табуляцыяй (мал. 5). Сэрвіс дае магчымасць карыстальніку не толькі генераваць АПП на падставе ўведзенага тэксту, але і адсочваць пэўную дадатковую інфармацыю датычна працэсу апрацоўкі. Для гэтага патрэбна адзначыць пункт «Атрымаць дадатковую інфармацыю», сэрвіс прадставіць інфармацыю пра словы, якія не атрымалася апрацаваць з-за адсутнасці той ці іншай неабходнай інфармацыі ў базе, а таксама падасць спіс знойдзеных ва ўваходным тэксце амографіў. У табліцы прадстаўлены прыклад выніковага АПП, згенераванага на падставе фрагменту табліц УДК пры дапамозе вэб-сэрвіса «Генератар алфавітна-прадметнага паказальніка».

Прыклад працы вэб-сэрвіса «Генератар алфавітна-прадметнага паказальніка»

Фрагмент асноўных табліц УДК	АПП, згенераваны на аснове фрагмента УДК
8 МОВА. МОВАЗНАЎСТВА. ЛІНГВІСТЫКА. ЛІТАРАТУРА	А
80 Агульныя пытанні лінгвістыкі і літаратуры. Філалогія	Адпаведнасць (<i>мовазнаўства</i>) 801.65
801 Прасодыя. Дапаможныя навукі і крыніцы філалогіі	В
801.6 Прасодыя: памер, рытм, рыфма і вершаваныя мадэлі	Верш (<i>мовазнаўства</i>) 801.65, 801.67
801.63 Метр. Метрычныя меры, стопы	Вершаваны (<i>мовазнаўства</i>) 801.6
801.65 Вершы і складовыя мадэлі (у адпаведнасці з лікавымі характарыстыкамі)	З
801.66 Рыфма	Зборнік (<i>мовазнаўства</i>) 801.8
801.67 Стансы, строфы, куплеты, вершы (у паэме)	К
801.7 Дапаможныя філалагічныя дысцыпліны	Крыніца (<i>мовазнаўства</i>) 801.8
801.8 Філалагічныя і лінгвістычныя крыніцы. Зборнікі тэкстаў	Куплет (<i>мовазнаўства</i>) 801.67
	Л
	Лікавы (<i>мовазнаўства</i>) 801.65
	Лінгвістыка 8, 80
	Лінгвістычны (<i>мовазнаўства</i>) 801.8
	Літаратура 8, 80
	М
	Мадэль (<i>мовазнаўства</i>) 801.6, 801.65
	Мера (<i>мовазнаўства</i>) 801.63
	Метр (<i>мовазнаўства</i>) 801.63
	Метрычны (<i>мовазнаўства</i>) 801.63
	Мова 8
	Мовазнаўства 8
	П
	Памер (<i>мовазнаўства</i>) 801.6
	Паэма (<i>мовазнаўства</i>) 801.67
	Прасодыя (<i>мовазнаўства</i>) 801.6
	Пытанне 80
	Р
	Рытм (<i>мовазнаўства</i>) 801.6
	Рыфма (<i>мовазнаўства</i>) 801.6, 801.66
	С
	Складовы (<i>мовазнаўства</i>) 801.65
	Станс або Стансы (<i>мовазнаўства</i>) 801.67
	Стапа (<i>мовазнаўства</i>) 801.63
	Страфа (<i>мовазнаўства</i>) 801.67
	Т
	Тэкст (<i>мовазнаўства</i>) 801.8
	Ф
	Філалагічны (<i>мовазнаўства</i>) 801.8
	Філалогія 80
	Х
	Характарыстыка (<i>мовазнаўства</i>) 801.65

Прыклад працы прататыпа сістэмы генерацыі АПП на падставе табліц УДК на беларускай мове дэманструе працаздольнасць распрацаваных алгарытмаў. Карэктнасць працы алгарытму была правярана рэдакцыйнай калегіяй выдання УДК на беларускай мове. Праграмная рэалізацыя дапамагла значна скараціць час на выданне, звёўшы працу рэдакцыйнай калегіі па распрацоўцы АПП да працэсу вычыткі і карэктывоўкі.

Заклучэнне. У артыкуле прапанавана ідэя магчымасці фарміравання алфавітна-прадметных паказальнікаў для розных структураваных сістэм праз выкарыстанне алгарытмаў апрацоўкі электронных тэкстаў машынным чынам. Гэтая ідэя праілюстравана распрацаваным і апісаным у артыкуле алгарытмам аўтаматызаванай генерацыі АПП да беларускага выдання УДК. У аснову алгарытму былі пакладзены метады і падыходы да аўтаматызаванай апрацоўкі

влялікіх аб'ёмаў электронных тэкстаў і распрацоўкі лінгвістычных і тэматычных рэсурсаў, шырока прадстаўлены на інтэрнэт-платформе апрацоўкі тэкстаў і маўлення www.Corpus.by.

На аснове апісанага ў артыкуле алгарытму быў распрацаваны прататып сістэмы генерацыі АПП, які знайшоў непасрэднае прымяненне ў падрыхтоўцы макета АПП першага беларускамоўнага выдання УДК [3].

Спіс выкарыстаных крыніц

1. McIlwaine, I. C. *The Universal Decimal Classification: a guide to its use* / I. C. McIlwaine. – The Hague : UDC Consortium, 2007. – 278 p.
2. Інструментарый індэксатара і яго прымяненне ў бібліятэках Беларусі / Нацыянальная бібліятэка Беларусі ; склад. С. А. Пугачова. – Мінск : Нацыянальная бібліятэка Беларусі, 2016. – 191 с.
3. Універсальная дзесятковая класіфікацыя: звыш 10 000 асноўных і дапаможных класаў / Аб'яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі, Нацыянальная бібліятэка Беларусі ; рэдкал.: Ю. С. Гецэвіч [і інш.]. – Мінск : Нацыянальная бібліятэка Беларусі, 2016. – 370 с.
4. *Universele Decimale Classificatie: beknopte nederlandse uitgave* / ed. G. J. A. Riesthuis. – The Hague : UDC Consortium, 2013. – 412 p.
5. Сербин, О. О. Подготовка УДК на украинском языке / О. О. Сербин // Библиосфера. – 2013. – № 2. – С. 69–73.
6. Станіслаўка, Г. Р. Выкарыстанне камп'ютарна-лінгвістычных сродкаў для перакладу ўніверсальнай дзесятковай класіфікацыі дамена «тэатр» з англійскай на беларускую мову і генерацыя алфавітна-прадметнага паказальніка / Г. Р. Станіслаўка, Ю. С. Гецэвіч, С. І. Лысы // Актуальныя пытанні германскай філалогіі і лінгвідыктыкі : матэрыялы XX Міжнароднага навука-практ. канф. / Брэст. гос. ун-т імя А. С. Пушкіна ; рэдкал.: Е. Г. Сальнікова [і др.]. – Брэст, 2016. – С. 264–266.
7. Станіслаўка, А. Г. Этапы падготовки першага выдання УДК на беларускім мове / А. Г. Станіслаўка, С. І. Лысы, Ю. С. Гецэвіч // Інфармацыя ў сучасным свеце : докл. Міжнароднага канф., Москва, 25–26 кастрычніка 2017 г. / ВІНІТІ РАН. – М., 2017. – С. 297–303.

References

1. McIlwaine I. C. *The Universal Decimal Classification: a guide to its use*. The Hague, UDC Consortium, 2007, 278 p.
2. Puhachova S. A. (ed.) *Instrumentaryj indeksatara i jaho prymaneniennje ŭ biblijatekach Bielarusi. Toolkit of indexers and its use in the libraries of Belarus*. Minsk, National Library of Belarus, 2016, 191 p. (in Belarusian).
3. Hetsevich Yu. S., Puhachova S. A., Stanislavenka H. R., Kuzminich T. V., Narejka A., Hetsevich S. A. (eds.) *Univiersalnaja dziesiatkovaja klasifikacyja: zvyš 10 000 asnoŭnych i dapamožnych klasaŭ. Universal Decimal Classification: more than 10 000 main and auxiliary classes*. Minsk, National Library of Belarus, 2016, 370 p. (in Belarusian).
4. Riesthuis G. J. A. (ed.) *Universele Decimale Classificatie: beknopte nederlandse uitgave*. The Hague, UDC Consortium, 2013, 412 p. (in Dutch).
5. Sierbin O. O. *Podhotovka UDK na ukrainskom jazykie. Preparation of UDC in Ukrainian*. *Bibliosfera*, 2013, no. 2, pp. 69–73 (in Russian).
6. Stanislavenka H. R., Hetsevich Yu. S., Lysy S. I. *Vykarystannie kamp'jutarna-linhvistyčnych srodkaŭ dla pierakladu ŭniviersalnaj dziesiatkovaj klasifikacyi damiena «teatr» z anhlijskaj na bielaruskuju movu i hienieracyja alfavitna-pradmetnaha pakazalnika [Using computational linguistic tools for translation of Universal Decimal Classification («theater» domain) from English into Belarusian and generation of Alphabetical Subject Index]. Aktualnyje voprosy hiermanskoj filolohii i linhvodidaktiki [Relevant questions of German philology and linguistics : XX International Scientific and Practical Conference Proceedings]*. Brest, 2016, pp. 264–266 (in Belarusian).
7. Stanislavenka H. R., Lysy S. I., Hetsevich Yu. S. *Etapy podhotovki piervocho izdaniia UDK na bielaruskom jazykie [Stages of preparation of the first edition of UDC in Belarusian]. Informaciia v sovriemiennom mirie : doklady Mieždunarodnoj konferencii [Information in the Modern World: Proceedings of the International Conference]*. Moscow, 2017, pp. 297–303 (in Russian).

Інфармацыя пра аўтараў

Лысы Станіслаў Іосіфавіч – аспірант, Аб’яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі (вул. Сурганова, 6, 220012, Мінск, Рэспубліка Беларусь). E-mail: stanislau.lysy@gmail.com

Станіславенка Ганна Рыгораўна – аспірант, Цэнтр даследаванняў беларускай культуры, мовы і літаратуры Нацыянальнай акадэміі навук Беларусі (вул. Сурганова, 1, 220072, Мінск, Рэспубліка Беларусь). E-mail: hanna.stanislaivenka@gmail.com

Гецэвіч Юрый Станіслававіч – кандыдат тэхнічных навук, Аб’яднаны інстытут праблем інфарматыкі Нацыянальнай акадэміі навук Беларусі (вул. Сурганова, 6, 220012, Мінск, Рэспубліка Беларусь). E-mail: yuras.hetsevich@newman.bas-net.by

Information about the authors

Stanislau I. Lysy – PhD student, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surhanava Str., 220012, Minsk, Republic of Belarus). E-mail: stanislau.lysy@gmail.com

Hanna R. Stanislavenka – PhD student, Center for the Belarusian Culture, Language and Literature Researches of the National Academy of Sciences of Belarus (1, Surhanava Str., 220072, Minsk, Republic of Belarus). E-mail: hanna.stanislaivenka@gmail.com

Yuryj S. Hetsevich – PhD (Engineering), The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surhanava Str., 220012, Minsk, Republic of Belarus). E-mail: yuras.hetsevich@newman.bas-net.by

ISSN 1816-0301 (print)
УДК 004

Поступила в редакцию 02.02.2018
Received 02.02.2018

Б. А. Залесский, Ф. С. Троцкий

*Объединенный институт проблем информатики Национальной академии
наук Беларуси, Минск, Беларусь*

ПАРАЛЛЕЛЬНАЯ ВЕРСИЯ ДЕТЕКТОРА ЭКСТРЕМАЛЬНЫХ ОСОБЫХ ТОЧЕК ИЗОБРАЖЕНИЙ

Аннотация. Рассматривается параллельная версия детектора особых (ключевых, характерных) точек экстремумов, применяемых для описания, анализа и сравнения изображений с помощью локальных дескрипторов, которые вычисляются в окрестностях найденных точек. Для задания ориентации дескрипторов предлагается использовать локальные гистограммы ориентированного градиента. В версии, предназначенной для выполнения на программно-аппаратной архитектуре CUDA, учтена специфика графических процессоров фирмы NVIDIA, что позволило ускорить вычисление экстремальных особых точек на несколько порядков. Вычисление неориентированных экстремальных особых точек изображения FullHD-размера на бюджетной видеокарте занимает 5–6 мс, ориентированных – 11–12 мс.

Ключевые слова: изображения, особые точки, детектор, алгоритм, параллельная версия, локальный дескриптор, CUDA

Для цитирования. Залесский, Б. А. Параллельная версия детектора экстремальных особых точек изображений / Б. А. Залесский, Ф. С. Троцкий // Информатика. – 2018. – Т. 15, № 2. – С. 55–63.

B. A. Zalesky, Ph. S. Trotski

*The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus*

PARALLEL VERSION OF DETECTOR OF EXTREMAL KEY POINTS ON IMAGES

Abstract. The article presents a parallel version of the detector of extremal key points, which are used to describe, analyze and compare digital images by local descriptors. Local descriptors are determined in neighborhoods of the extremal key points. The orientation of the descriptors are found with aid of Histograms of Oriented Gradient. The specificity of the parallel architecture of NVIDIA graphics cards has been taken into account in the developed version, oriented to the implementation on CUDA. It accelerated the calculation of the extremal key points by several orders. Computation of the not oriented extremal key points for images of the FullHD-size on the budget graphics card takes 5–6 ms. The oriented extremal key points are computed within 11–12 ms.

Keywords: images, key points, detector, algorithm, parallel version, local descriptor, CUDA

For citation. Zalesky B. A., Trotski Ph. S. Parallel version of detector of extremal key points on images. *Informatics*, 2018, vol. 15, no. 2, pp. 55–63 (in Russian).

Введение. Многие актуальные задачи анализа, сравнения и распознавания изображений решаются с помощью сопоставления их локальных характеристик, вычисленных в окрестностях особых (ключевых) точек. Следуя сложившейся терминологии, будем называть алгоритмы, находящие особые точки, *детекторами*, а алгоритмы, вычисляющие локальные характеристики в каждой найденной особой точке, – *дескрипторами*. Характеристики особых точек сравниваются с помощью *алгоритмов поиска соответствий* (англоязычный термин *matchers*). При использовании особых точек для решения задач сравнения или распознавания изображений сначала применяется один из детекторов, затем дескриптор и, наконец, алгоритм поиска соответствий.

Одни из первых детекторов, успешно примененных для решения задач компьютерного зрения, были предложены в 1999 г. в алгоритмах SIFT и HoG [1, 2]. В первом алгоритме используется детектор ключевых точек, основанный на анализе гессиана разности сглаженных

в разной степени копий исходного изображения, а во втором ключевыми считаются все точки регулярной решетки. Спустя год был опубликован алгоритм SURF, детектор которого находит ключевые точки с помощью приближенного гессияна, вычисляемого с помощью интегрального изображения [3]. Использование интегрального изображения позволило значительно ускорить процесс вычисления особых точек по сравнению с SIFT и HoG.

Затем появились FAST [4], STAR [5], BRISK [6], KAZE [7], AKAZE [8] и другие алгоритмы, в состав которых входят различные версии детекторов особых точек. Сравнительный анализ детекторов выполнен в работе [9]. Здесь следует отметить появившийся несколько лет назад дескриптор LATCH [10], который строит информативные локальные характеристики, обеспечивающие надежное сравнение изображений, с помощью ранее предложенных детекторов.

Среди недостатков известных детекторов следует отметить их чувствительность к изменениям яркости и контраста изображений, а также нередкие случаи неравномерного распределения особых точек на изображениях. (Области изображения без особых точек не могут быть обнаружены или распознаны с помощью алгоритмов поиска соответствий.)

Ранее одним из авторов [11] были предложены детекторы особых точек, основанные на поиске локальных экстремумов функций – характеристик непреобразованных изображений. Предложенные детекторы менее чувствительны к изменениям яркости изображений, найденные ими особые точки распределены на изображении более равномерно.

В настоящей статье рассматривается один из предложенных в [11] экстремальных детекторов, основанный на вычислении локальных максимумов оконных дисперсий изображения, который для краткости будем обозначать σ^2 -детектор. Формальное описание σ^2 -детектора приведено в разд. 1. Данный детектор находит на изображении примерно столько же особых точек, сколько находят другие экстремальные моментные детекторы [11], однако он проще реализуется программно, а его вычисление занимает меньше времени.

Указанный σ^2 -детектор может быть вычислен на CPU (central processing unit) персонального компьютера с помощью интегрального изображения [3], однако процессорная версия недостаточно быстра для работы с видеопоследовательностями размером FullHD и HD в режиме реального времени.

В настоящей работе предлагается параллельная версия алгоритма вычисления σ^2 -детектора, предназначенная для выполнения на программно-аппаратной архитектуре CUDA. В ней учтена специфика графических процессоров GPU (graphics processing unit) фирмы NVIDIA, что позволило ускорить вычисление детектора на несколько порядков. Описывается также кратко CUDA-версия вычисления ориентации особых точек относительно изображения с помощью локальных гистограмм ориентированного градиента. Приводятся результаты сравнения быстродействия предложенной параллельной версии алгоритма вычисления σ^2 -детектора с быстродействием других детекторов, в том числе реализованных на CUDA.

1. Определение σ^2 -детектора и его вычисление на CPU. Для формального определения σ^2 -детектора рассмотрим полутоновое изображение $\mathbf{I} = \{I(\mathbf{p})\}$, $\mathbf{p} \in S$, с множеством пикселей $S = \{(x, y)\}$, $x = 0, \dots, m-1$, $y = 0, \dots, n-1$, квадратные окрестности $O_k(\mathbf{p})$ с центром в пикселях \mathbf{p} размером $k \times k$, а также локальные (оконные) дисперсии

$$\sigma^2 = \sigma^2(\mathbf{p}) = \beta_2(\mathbf{p}) - m^2(\mathbf{p}), \quad m(\mathbf{p}) = k^{-2} \sum_{\mathbf{g} \in O_k(\mathbf{p})} I(\mathbf{g}) \quad \text{и} \quad \beta_2(\mathbf{p}) = k^{-2} \sum_{\mathbf{g} \in O_k(\mathbf{p})} I^2(\mathbf{g}). \quad (1)$$

Экстремальный σ^2 -детектор выделяет пиксел \mathbf{p} изображения как особую точку, если существует окрестность $O_\ell(\mathbf{q})$ размером $\ell \times \ell$ (значения k и ℓ выбираются соответствующими размерам локальных признаков на изображении, в большинстве экспериментов выбирались $k = 5$, $\ell = 4$) с центром в \mathbf{q} , содержащая \mathbf{p} , для которой значение $\sigma^2(\mathbf{p})$ является максимумом значений σ^2 в этой окрестности, причем выполняются следующие условия:

Условие 1. $\sigma^2(\mathbf{p}) > \sigma^2(\mathbf{v})$ хотя бы для одного пиксела $\mathbf{v} \in O_\ell(\mathbf{q})$.

Условие 2. Решение \mathbf{p} находится от центра окрестности \mathbf{q} на расстоянии, не превосходящем наперед заданное число τ ($0 < \tau < \frac{\ell}{2}$).

Замечание. Условие 1 равносильно тому, что $\sigma^2(\mathbf{p})$ не является константой в окрестности $O_\ell(\mathbf{q})$, а условие 2 предотвращает попадание экстремальной особой точки на границу окрестности $O_\ell(\mathbf{q})$. При нарушении хотя бы одного из условий найденные точки могут оказаться неинформативными. Нетрудно показать, что отсутствие условия 1 приводит к тому, что все внутренние точки изображения постоянной яркости стали бы особыми, хотя на самом деле в этом случае они не обладают индивидуальными локальными признаками. При невыполнении условия 2 все внутренние точки изображения, яркости которого лежат на некоторой наклонной плоскости, являются особыми.

Вместо условия 1 можно использовать другие условия, предотвращающие появление неинформативных особых точек, например условие, чтобы $\sigma^2(\mathbf{p})$ был строгим и поэтому единственным максимумом в $O_\ell(\mathbf{q})$, и т. д.

Формально экстремальные особые точки образуют множество

$$E = E_{k,\ell,\tau}(\mathbf{I}) = \{ \mathbf{p} \in S \mid \exists \mathbf{q} \in S : \mathbf{p} \in O_\ell(\mathbf{q}), \|\mathbf{p} - \mathbf{q}\| \leq \tau, \forall \mathbf{u} \in O_\ell(\mathbf{q}), \sigma^2(\mathbf{p}) \geq \sigma^2(\mathbf{u}), \exists \mathbf{v} \in O_\ell(\mathbf{q}) : \sigma^2(\mathbf{p}) > \sigma^2(\mathbf{v}) \}. \quad (2)$$

В некоторых случаях при практическом применении предложенного σ^2 -детектора множество особых точек E приходится прореживать – оставлять из нескольких расположенных близко равноценных особых точек одну, например, случайным образом или по какому-либо дополнительному критерию. Процедура прореживания особых точек не занимает много времени и используется при работе с большинством известных детекторов.

Предложенный σ^2 -детектор может быть вычислен на CPU с помощью двух интегральных изображений, одно из которых построено по яркостям, а второе – по квадратам яркостей исходного изображения.

Для вычисления оконной дисперсии на CPU можно также использовать метод «бегущей строки» для быстрого вычисления оконных сумм, но даже в этом случае время вычисления процессорной версии детектора будет слишком большим для работы с видеопоследовательностями.

В большинстве случаев локальные характеристики особых точек, вычисляемые дескрипторами, неинвариантны к повороту изображения, поэтому для каждой найденной точки нужно задавать каким-либо образом ее ориентацию, которая используется при построении характеристик. Несмотря на то что, по сути, ориентация особой точки является одной из ее характеристик, многие программные реализации детекторов определяют не только координаты таких точек, но и направления, относительно которых потом дескрипторы вычисляют характеристики. Ниже кратко описывается быстрая реализация на CUDA алгоритма процедур нахождения ориентации особых точек на основе поиска максимума локальных гистограмм ориентированного градиента.

2. Вычисление σ^2 -детектора на GPU. При описании параллельной версии алгоритма вычисления σ^2 -детектора на CUDA будет удобнее индексировать все прямоугольные окна на изображении не центральным пикселом, а координатами их левого верхнего угла. Для этого будем использовать вместо обозначения $O_k(\mathbf{p})$ обозначение O_p или $O_{(x,y)}$.

Сначала изложим кратко суть алгоритма, после чего приведем пошаговое описание его реализации.

Стандартные операции выделения памяти процессора, копирования входных данных в память GPU и результатов из памяти GPU, очистки памяти процессора и GPU не будут включены в состав алгоритма, хотя в разд. 4 приведены отдельно их длительности. Следует пом-

нить, что указанные операции являются относительно медленными – они занимают в общей сложности несколько миллисекунд, что сравнимо с выполнением нескольких шагов алгоритма, поэтому при его многократном применении, например, для работы с видеопоследовательностями желательно избегать повторяющихся выделений и очисток памяти и использовать однажды определенные массивы.

При организации параллельных вычислений на GPU одна из основных задач – минимизация числа конфликтов одновременного доступа к видеопамяти.

Идея предложенного алгоритма состоит в представлении сумм значений яркостей (квадратов яркостей) изображения в окне O_p детектора в виде сумм значений яркостей (квадратов яркостей) его столбцов, а затем сложения полученных сумм. Такой подход имеет смысл в силу того, что сумма значений одного столбца может использоваться при подсчете сумм яркостей (и квадратов яркостей) нескольких окон O_p . Для уменьшения числа конфликтов одновременного доступа к видеопамяти изображение разбивается на горизонтальные полосы, лежащие друг от друга на расстоянии k , равном высоте окна, и сначала считаются суммы значений яркостей $k \times 1$ -столбцов изображения, первые элементы которых начинаются на выделенных горизонтальных полосах (рис. 1).

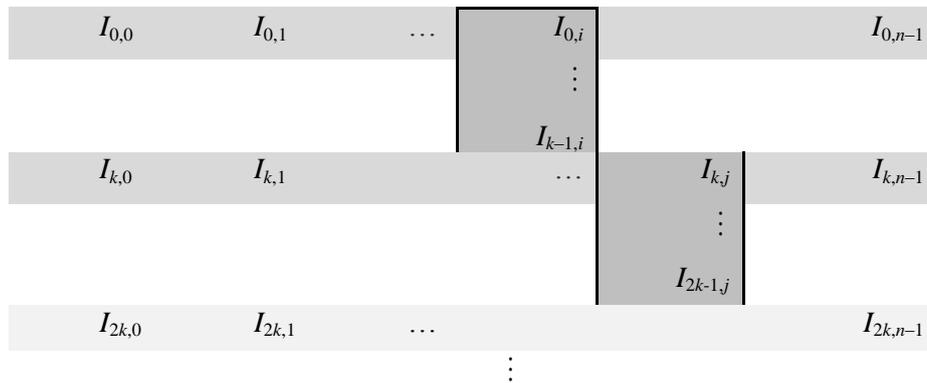


Рис. 1. Схема вычисления сумм значений яркостей в столбцах (изображены два столбца с начальными элементами на направляющих, индексы i, j – номера столбцов, k – номер строки изображения, n – длина изображения)

Затем методом бегущей строки вычисляются суммы всех остальных столбцов высотой k , не лежащих на выделенных полосах. Если обозначить через $c_{(x,y)} = \sum_{j=0}^{k-1} I_{(x,y+j)}$ (через $C_{(x,y)} = \sum_{j=0}^{k-1} I_{(x,y+j)}^2$) сумму (сумму квадратов) значений яркостей $k \times 1$ -столбца изображения с верхним пикселем $\mathbf{p} = (x, y)$, то получим выражения

$$c_{(x,y+1)} = c_{(x,y)} + I_{(x,y+k)} - I_{(x,y)},$$

$$C_{(x,y+1)} = C_{(x,y)} + I_{(x,y+k)}^2 - I_{(x,y)}^2. \quad (3)$$

Значения $c_{(x,y)}$ и $C_{(x,y)}$ записываются в массивы, размер которых равен размеру исходного изображения. Изображения c и C транспонируются, и для них повторяются предыдущие вычисления, в которых $I_{(x,y)}$ заменено на $c_{(x,y)}$, а $I_{(x,y)}^2$ – на $C_{(x,y)}$.

Далее вычисляется транспонированная матрица Σ^T с элементами

$$\tilde{\sigma}_{(x,y)}^2 = \sum_{u=0}^{k-1} C_{(x,u)} - \left(\sum_{u=0}^{k-1} c_{(x,u)} \right)^2 / k^2, \quad (4)$$

которая отличается от матрицы выборочных оконных дисперсий на множитель k^2 (очевидно, что координаты (x, y) максимумов двух матриц совпадают).

Для вычисления множества экстремальных ключевых точек в общем случае необходимо найти максимумы дисперсий внутри всех $\ell \times \ell$ -квадратов $O_{(x,y)}$, а затем проверить условие близости координат найденных локальных максимумов к центрам окрестностей, присутствующее в формуле (2).

Оптимизация вычисления на GPU локальных экстремумов во всех пикселах достаточна сложна, а сами вычисления весьма ресурсоемки, поэтому для ускорения алгоритма и упрощения его реализации предлагается разбить пикселы изображения на непересекающиеся квадраты. Внутри этих квадратов требуется найти пикселы локальных максимумов и выбрать в качестве особых точек те, которые удалены от центров квадратов на расстояние, не превосходящее заданное. После этого необходимо сместить все квадраты, разбивающие S , на половину их стороны по вертикали и горизонтали и найти остальные особые точки внутри них.

Запишем алгоритм вычислений на GPU:

Шаг 1. Вычислить и сохранить в 1D-массиве суммы $c_{(x,y)}$ значений яркостей и суммы $C_{(x,y)}$ квадратов яркостей столбцов, первые элементы которых лежат на горизонтальных направляющих с y -координатами вида $y = \mu k$, где $\mu = 0, \dots, [m/k] - 1$.

Шаг 2. Вычислить суммы всех остальных столбцов $c_{(x,y)}$ и $C_{(x,y)}$ методом бегущей строки по формуле (3) и сохранить результат в 1D-массивах, представляющих собой построчную запись матриц \mathbf{c} и \mathbf{C} .

Шаг 3. Транспонировать матрицы \mathbf{c} и \mathbf{C} , используя их 1D-представление. (Смысл шагов 3 и 4 объясняется в разд. 3.)

Шаг 4. Вычислить и сохранить в 1D-массиве значения $\tilde{\sigma}_{(x,y)}^2$ на горизонтальных направляющих с y -координатами вида $y = \mu k$, где $\mu = 0, \dots, [n/k] - 1$, используя 1D-представления матриц \mathbf{c}^T и \mathbf{C}^T .

Шаг 5. Вычислить методом бегущей строки и сохранить в 1D-массиве значения $\tilde{\sigma}_{(x,y)}^2$.

Шаг 6. Найти координаты (x_{\max}, y_{\max}) локальных максимумов $\tilde{\sigma}_{(x,y)}^2$ в окрестностях $O_{\ell,(x,y)}$ методом редукции.

Шаг 7. Сохранить в качестве особых точек те (x_{\max}, y_{\max}) , для которых $\|(x_{\max}, y_{\max}) - (x + \ell/2, y + \ell/2)\| < \tau$, где τ – наперед заданный порог, удовлетворяющий неравенству $0 < \tau < \frac{\ell}{2}$.

Шаг 8. Вычислить градиент $\mathbf{G}(\mathbf{I})$ изображения \mathbf{I} , используя один из известных дискретных дифференциальных операторов, например Собеля, Превитт, Робертса или Щара.

Шаг 9. Построить гистограмму оконного ориентированного градиента каждой особой точки и задать в качестве ее ориентации номер координаты максимального значения гистограммы.

3. Особенности реализации алгоритма на CUDA. Как известно, архитектура параллельных вычислений CUDA, разработанная фирмой NVIDIA, предназначена для проведения массивных параллельных вычислений на графических процессорах. Вычисления на графическом процессоре осуществляются путем запуска ядра GPU, при вызове которого указывается

число требуемых блоков, а также число нитей, используемых в каждом из задействованных блоков. Каждый задействованный блок выполняется на одном потоковом мультипроцессоре (SM). Количество одновременно выполняющихся операций (а следовательно, и время вычислений) зависит от числа мультипроцессоров и их характеристик. На бюджетных видеокартах имеется от 5 до 20 SM. Например, на GeForce GTX 1050 и 1060 имеется 10 SM, а на GTX1080 – 20.

Для эффективного параллельного выполнения операций нитями каждого блока требуется избегать конфликтов одновременного доступа нитей блока к ячейкам памяти GPU с учетом специфики обращения к ней варпами. Вышесказанное объясняет использование метода бегущей строки для вычисления детектора, а также способа вычисления среднего, дисперсии и ориентированного градиента по столбцам на шагах 1–4 алгоритма. Проведенные эксперименты показали, что вычисление оконной суммы значений яркости изображений размера FullHD на видеокарте GeForce GTX 1050 методом бегущей строки при выполнении суммирования по столбцам занимает 0,22 мс, а при суммировании по строкам – 2,21 мс. Операция транспонирования матрицы размером FullHD на той же видеокарте заняла 0,57 мс.

Таким образом, предлагаемый способ вычисления оконных сумм изображения суммированием значений $c_{(x,y)}$, $C_{(x,y)}$ методом бегущей строки по столбцам, затем транспонированием матрицы полученных сумм и, наконец, вычислением результата повторным суммированием транспонированной матрицы по столбцам потребовал в два раза меньше времени, чем суммирование промежуточных значений, полученных на шаге 2 алгоритма, по строке (без транспонирования).

Вычисление максимальных значений $\tilde{\sigma}^2_{(x,y)}$ в окрестностях $O_{\ell,(x,y)}$ выполнялось методом редукции [12, с. 75].

4. Результаты экспериментов. Для оценки быстродействия предложенной параллельной версии алгоритма вычисления экстремального σ^2 -детектора были проведены эксперименты с полутоновыми изображениями разных размеров на персональном компьютере, оборудованном видеокартой NVIDIA GeForce GTX 1050, имеющей 2 Гбайта видеопамати, и ноутбуке с видеокартой NVIDIA GeForce GTX 1060 с 6 Гбайтами видеопамати.

В таблице приведены средние длительности 50 повторных вычислений экстремальных особых точек с использованием для хранения исходных изображений неблокированной памяти, выделенной функцией malloc, и блокированной, выделенной функцией cudaHostAlloc (сНА).

Результаты экспериментов

Размер полутонового изображения	Количество найденных точек	Время вычисления, мс							
		NVIDIA GeForce GTX 1050				NVIDIA GeForce GTX 1060			
		с ориентацией		без ориентации		с ориентацией		без ориентации	
malloc	сНА	malloc	сНА	malloc	сНА	malloc	сНА		
1920×1080	10979	17,90	17,22	7,65	7,03	11,41	10,84	5,97	5,31
1600×850	7362	11,88	11,33	5,12	4,61	8,01	7,67	4,23	3,93
1280×720	4787	8,09	7,69	3,50	3,12	5,37	5,03	2,84	2,48
720×480	1613	3,55	3,59	1,53	1,51	2,56	2,34	1,37	1,19

Время выделения памяти на видеокарте не включено в результаты вычислений, так как при повторном использовании алгоритма целесообразно использовать однажды выделенную память. Время копирования данных из памяти компьютера в видеопамать GPU NVIDIA GeForce GTX 1060 изображения размером 1920×1080 составляет 0,67 мс для блокированной памяти и 1,22 мс – для неблокированной. Время копирования результатов вычислений – найденных характеристик – составляет примерно 0,15 мс.

Ниже приведены результаты сравнения предложенного детектора экстремальных особых точек и известного детектора SURF (реализованного на CUDA в OpenCV 3.3), дающего наиболее информативные особые точки на рассмотренных изображениях в сравнении с алгоритмами ORB, SIFT и AKAZE. Оба алгоритма применялись для установления соответствий между особыми точками пар изображений, найденными предложенным экстремальным детектором

и CUDA-версией детектора SURF. Для оценки быстродействия использовались только те пары изображений, которые были корректно совмещены друг с другом на основе особых точек, найденных обоими алгоритмами.

Нахождение экстремальных точек FullHD-изображения размером 1920×1080 на CPU Intel Core i7-6700 с использованием метода бегущей строки для быстрого вычисления оконных сумм заняло более 400 мс. Для изображений размером 720×480 это время равнялось 54 мс.

Приведем результаты вычислительных экспериментов для GPU NVIDIA GeForce GTX 1050. Вычисление особых точек на полутоновых изображениях природного ландшафта (рис. 2) размером FullHD с помощью программной реализации экстремального детектора на CUDA заняло около 17 мс (для каждого изображения), а с помощью CUDA-версии детектора SURF, реализованной в OpenCV 3.3, – примерно 54 мс. Для изображений лесного ландшафта размером 720×480 время вычисления особых точек с помощью реализации предложенного алгоритма и CUDA-версии детектора SURF заняло 4,5 и 13 мс соответственно.

Вычисление особых точек изображения размером 1920×1080 на видеокарте NVIDIA GeForce GTX 1060 предложенным детектором и детектором CUDA-версии SURF занимает соответственно 12,3 и 33,8 мс.



Рис. 2. Соответствие изображений, установленное с помощью экстремальных особых точек, найденных σ^2 -детектором

На рис. 3 приведены соответствия двух изображений, установленные предложенным детектором. Примечательно, что известные алгоритмы SIFT, SURF, ORB и AKAZE не нашли на меньшем изображении ни одной особой точки.



Рис. 3. Экстремальные особые точки, вычисленные σ^2 -детектором, и найденные соответствия; в качестве шаблона использовалось меньшее изображение (а) и большее изображение (б)

Заклучение. Особенностью предложенного детектора является возможность его применения для слабоконтрастных изображений, когда известные детекторы, основанные на вычислении градиента, не находят достаточного количества особых точек. Особые точки, вычисленные экстремальным детектором, в большинстве случаев распределяются по изображению более равномерно (по сравнению с точками, найденными другими детекторами).

Предложенная параллельная версия экстремального детектора особых точек позволяет проводить вычисления в режиме реального времени на бюджетных видеокартах, таких, например, как NVIDIA GeForce GTX 1050. Во многих случаях данная версия оказалась не только устойчивее, но и в несколько раз быстрее CUDA-версий известных детекторов. Например, вычисление особых точек вместе с их ориентацией на изображении размером 720×480 с помощью экстремального детектора занимает всего 2,3÷2,6 мс, что дает возможность выполнить другие функции обработки изображений, не выходя из режима реального времени.

Вычисления в предложенной параллельной реализации организованы так, чтобы уменьшить, насколько возможно, число конфликтов одновременного доступа к ячейкам видеопамяти.

Список использованных источников

1. Lowe, D. Object recognition from local scale invariant features / D. Lowe // Proc. of the 7th IEEE Intern. Conf. on Computer Vision (ICCV). – Corfu, 1999. – P. 1150–1157.
2. Dalah, N. Histograms of oriented gradients for human detection / N. Dalah, B. Triggs // Proc. of the IEEE Computer Society Conf. on Computer Vision and Pattern Recognition (CVPR 2005). – San Diego, 2005. – Vol. 1. – P. 886–893.
3. Bay, H. SURF: speeded up robust features / H. Bay, T. Tuytelaars, L. Van Gool // Proc. of the 9th European Conf. on Computer Vision (ECCV). – Austria, Graz, 2006. – Vol. 3951, pt. 1. – P. 404–417.
4. Rosten, E. Faster and better: a machine learning approach to corner detection / E. Rosten, T. Drummond // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2010. – Vol. 32, no. 1. – P. 105–119.
5. Agrawal, M. CenSurE: center surround extremas for realtime feature detection and matching / M. Agrawal, K. Konolige, M. R. Blas // Lecture Notes in Computer Science. – 2008. – Vol. 5305. – P. 102–115.
6. Leutenegger, S. BRISK: binary robust invariant scalable keypoints / S. Leutenegger, M. Chli, R. Y. Siegwart // Proc. of the 13th IEEE Intern. Conf. on Computer Vision (ICCV). – Barcelona, 2011. – P. 2548–2555.
7. Alcantarilla, P. KAZE features / P. Alcantarilla, A. Bartoli, J. Davison // Proc. of the 12th European Conf. on Computer Vision (ECCV). – Firenze, 2012. – P. 214–227.
8. Alcantarilla, P. Fast explicit diffusion for accelerated features in nonlinear scale spaces / P. Alcantarilla, J. Nuevo, A. Bartoli // Proc. of the 24th British Machine Vision Conf. (BMVC). – Bristol, 2013. – P. 11–21.
9. Comparative assessment of point feature detectors and descriptors in the context of robot navigation / A. Schmidt [et al.] // J. of Automation, Mobile Robotics and Intelligent Systems. – 2013. – Vol. 7, no. 1. – P. 11–20.
10. Levi, G. Learned arrangements of three patch codes / G. Levi, T. Hassner // IEEE Winter Conf. on Applications of Computer Vision (WACV). – NY, USA, 2016. – P. 33–42.
11. Залесский, Б. А. Детекторы экстремальных особых точек на изображениях / Б.А. Залесский // Доклады Национальной академии наук Беларуси. – 2017. – № 5(61). – С. 37–41.
12. Сандерс, Д. Технология CUDA в примерах: введение в программирование графических процессоров / Д. Сандерс, Э. Кэндрот. – М.: ДМК Пресс, 2013. – 234 с.

References

1. Lowe D. Object recognition from local scale invariant features. *Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV)*. Corfu, 1999, pp. 1150–1157.
2. Dalah N., Triggs B. Histograms of oriented gradients for human detection. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*. San Diego, 2005, vol. 1, pp. 886–893.
3. Bay H., Tuytelaars T., Van Gool L. SURF: speeded up robust features. *Proceedings of the 9th European Conference on Computer Vision (ECCV)*. Austria, Graz, 2006, vol. 3951, pt. 1, pp. 404–417.
4. Rosten E., Drummond T. Faster and better: a machine learning approach to corner detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, vol. 32, no. 1, pp. 105–119.
5. Agrawal M., Konolige K., Blas M. R. CenSurE: center surround extremas for realtime feature detection and matching. *Lecture Notes in Computer Science*, 2008, vol. 5305, pp. 102–115.
6. Leutenegger S., Chli M., Siegwart R. Y. BRISK: binary robust invariant scalable keypoints. *Proceedings of the 13th IEEE International Conference on Computer Vision (ICCV)*. Barcelona, 2011, pp. 2548–2555.
7. Alcantarilla P., Bartoli A., Davison J. KAZE features. *Proceedings of the 12th European Conference on Computer Vision (ECCV)*. Firenze, 2012, pp. 214–227.
8. Alcantarilla P., Nuevo J., Bartoli A. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *Proceedings of the 24th British Machine Vision Conference (BMVC)*. Bristol, 2013, p. 11–21. doi: 10.5244/C.27.13

9. Schmidt A., Kraft M., Fularz M., Domagała Z. Comparative assessment of point feature detectors and descriptors in the context of robot navigation. *Journal of Automation, Mobile Robotics and Intelligent Systems*, 2013, vol. 7, no. 1, pp. 11–20.
10. Levi G., Hassner T. Learned arrangements of three patch codes. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, NY, USA, March, 2016, p. 33–42. doi: 10.1109/WACV.2016.7477723
11. Zaleskiy B. A. Detektoryi ekstremalnykh osobiykh tochek na izobrazheniyah [Detectors of extremal key points on images]. *Doklady Natsionalnoy akademii nauk Belarusi [Reports of the National Academy of Sciences of Belarus]*, 2017, no. 5(61), pp. 37–41 (in Russian).
12. Sanders D., Kendrot E. Tehnologiya CUDA v primerah: vvedenie v programmirovaniye graficheskikh protsessirov. *CUDA by Example: An Introduction to General-Purpose GPU Programming*. Moscow, DMK Press Publ., 2013, 234 p. (in Russian).

Информация об авторах

Залесский Борис Андреевич – доктор физико-математических наук, заведующий лабораторией обработки и распознавания изображений, Объединенный институт проблем информатики Национальной академии наук Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: zalesky@newman.bas-net.by

Троцкий Филипп Сергеевич – младший научный сотрудник, Объединенный институт проблем информатики Национальной академии наук Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: trotskiphilipp@gmail.com

Information about the authors

Boris A. Zalesky – Dr. Sc. (Physics and Mathematics), Head of Laboratory of Image Processing and Recognition, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: zalesky@newman.bas-net.by

Philip S. Trotski – Junior researcher, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: trotskiphilipp@gmail.com

ISSN 1816-0301 (print)

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ**MATHEMATICAL MODELING**

УДК 517.958:537.8

Поступила в редакцию 08.01.2018
Received 08.01.2018**В. Т. Ерофеенко***Учреждение БГУ «НИИ прикладных проблем математики
и информатики», Минск, Беларусь***МОДЕЛИ БАЗИСНЫХ ЭЛЕКТРОМАГНИТНЫХ ВОЛН В БИИЗОТРОПНОЙ
СРЕДЕ С ПРОСТРАНСТВЕННОЙ ДИСПЕРСИЕЙ**

Аннотация. Разрабатывается математическая модель, описывающая распространение монохроматических электромагнитных волн в биизотропной матрице, содержащей сферические частицы из материалов с пространственной дисперсией. Моделирование основано на преобразовании интегро-дифференциальных уравнений для биизотропных сред к дифференциальным уравнениям электродинамики, содержащим электрическую и магнитную поляризации в операторном виде с дифференциальным оператором Лапласа. В рамках разработанной модели аналитически строится полная система базисных плоских, сферических и цилиндрических электромагнитных полей, распространяющихся в однородной биизотропной среде с пространственной дисперсией.

Ключевые слова: математические модели, интегро-дифференциальная модель, электромагнитные поля, плоские поля, сферические поля, цилиндрические поля, биизотропные среды, пространственная дисперсия

Для цитирования. Ерофеенко, В. Т. Модели базисных электромагнитных волн в биизотропной среде с пространственной дисперсией / В. Т. Ерофеенко // Информатика. – 2018. – Т. 15, № 2. – С. 64–73.

V. T. Erofeenko*Establishment of BSU "Research Institute of Applied
Mathematics and Informatics", Minsk, Belarus***MODELS OF BASIC ELECTROMAGNETIC WAVES IN THE BI-ISOTROPIC
MEDIUM WITH SPACE DISPERSION**

Abstract. A mathematical models describing the propagation of monochromatic electromagnetic waves in the bi-isotropic medium with space dispersion is being developed. Modeling is based on the transformation of integro-differential equations for bi-isotropic medium to the differential equations electrodynamics, containing electrical and magnetic polarizations in operator form with differential Laplace operator. Within the developed model it is analytically constructed a full system of basic plane, spherical and cylindrical electromagnetic fields extending in the homogeneous bi-isotropic medium with space dispersion.

Keywords: mathematical models, electromagnetic fields, plane fields, spherical fields, cylindrical fields, integro-differential model, bi-isotropic medium, space dispersion

For citation. Erofeenko V. T. Models of basic electromagnetic waves in the bi-isotropic medium with space dispersion. *Informatics*, 2018, vol. 15, no. 2, pp. 64–73 (in Russian).

Введение. В последнее десятилетие активно ведутся теоретические и прикладные исследования метаматериалов, биизотропных, бианизотропных и киральных материалов [1–5], а также электродинамических устройств из них [6–8]. Особое внимание уделяется исследованию плоских экранов, выполненных из биизотропных композитов, которые подвергаются воздействию плоских электромагнитных волн [9–11], полей дипольных источников [12, 13], пучков электромагнитных волн [14] и нестационарных полей [15]. Для прикладных исследований важ-

ным также является класс материалов с пространственной дисперсией [16–20]. Построение адекватных математических моделей, описывающих распространение электромагнитных волн в таких средах, представляет значительный научный интерес. Математические модели материалов, обладающих одновременно свойствами биизотропных сред и сред с пространственной дисперсией, не исследованы.

В настоящей статье предложена дифференциальная модель биизотропной среды с пространственной дисперсией, которая позволяет аналитическими методами строить распространяющиеся в среде электромагнитные поля в декартовых, сферических и цилиндрических координатах.

Интегро-дифференциальная модель биизотропной среды с пространственной дисперсией. Рассмотрим пространство R^3 с декартовой системой координат $Oxyz$, заполненное однородной биизотропной матрицей с материальными параметрами: $\varepsilon = \varepsilon_r \varepsilon_0$, $\mu = \mu_r \mu_0$ – диэлектрическая и магнитная проницаемости; $G = G_r/c$, $Z = Z_r/c$ – параметры биизотропности; ε_0 , μ_0 – электрическая и магнитная постоянные; c – скорость света в вакууме. В матрице случайным образом распределены частицы четырех типов с радиусами R_j ($j = 1, 2, 3, 4$), заполненные материалом с пространственной дисперсией. Такую композитную среду из матрицы с частицами будем называть биизотропной средой с пространственной дисперсией (БСПД). Комплексные амплитуды \vec{E}, \vec{H} монохроматического электромагнитного поля с временной зависимостью $\exp(-i\omega t)$ в такой среде подчиняются уравнениям

$$\operatorname{rot} \vec{H} = -i\omega(\varepsilon \vec{E} + G \vec{H} + \vec{P}), \operatorname{rot} \vec{E} = i\omega(\mu \vec{H} + Z \vec{E} + \vec{m}), \quad (1)$$

где ω – круговая частота поля.

Электрическая и магнитная поляризации определяются объемными интегралами по пространственным переменным $\vec{r}_0 = (x_0, y_0, z_0)$:

$$\vec{P}(\vec{r}) = \vec{P}^{(1)}(\vec{r}) + \vec{P}^{(2)}(\vec{r}) = \frac{\varepsilon_0}{V_1} \int_{D_{1M}} K_1(|\vec{r} - \vec{r}_0|/R_1) \vec{E}(\vec{r}_0) d\vec{r}_0 + \frac{1}{cV_3} \int_{D_{3M}} K_3(|\vec{r} - \vec{r}_0|/R_3) \vec{H}(\vec{r}_0) d\vec{r}_0; \quad (2)$$

$$\vec{m}(\vec{r}) = \vec{m}^{(1)}(\vec{r}) + \vec{m}^{(2)}(\vec{r}) = \frac{\mu_0}{V_2} \int_{D_{2M}} K_2(|\vec{r} - \vec{r}_0|/R_2) \vec{H}(\vec{r}_0) d\vec{r}_0 + \frac{1}{cV_4} \int_{D_{4M}} K_4(|\vec{r} - \vec{r}_0|/R_4) \vec{E}(\vec{r}_0) d\vec{r}_0, \quad (3)$$

где $D_{jM} = \{|\vec{r}_0|/|\vec{r} - \vec{r}_0| < R_j\}$ – шар радиуса R_j , описанный вокруг точки $M(x, y, z)$, точка

$M_0(x_0, y_0, z_0) \in D_{jM}$, $\vec{r} = (x, y, z)$; $V_j = \frac{4}{3} \pi R_j^3$; $K_j(p)$ – заданные функции, определяющие характер пространственной дисперсии, $j = 1, 2, 3, 4$.

Дифференциальная модель биизотропной среды с пространственной дисперсией. Интегро-дифференциальные уравнения (1)–(3), описывающие распространение электромагнитных волн в БСПД, преобразуем к дифференциальной модели среды. Такое преобразование при моделировании используется, когда радиусы частиц в матрице $R_j \ll \lambda_{\text{мат}}$, где $\lambda_{\text{мат}}$ – длина волн в материале матрицы [16, с. 170]. Методику преобразований и итоговые дифференциальные уравнения модели сформулируем в виде теоремы.

Теорема 1. С точностью до величин третьего порядка малости система интегро-дифференциальных уравнений (1)–(3) эквивалентна системе дифференциальных уравнений

$$\begin{aligned} \operatorname{rot} \vec{H} &= -i\omega(\varepsilon_{\Pi} \vec{E} + G_{\Pi} \vec{H} + P_1 \Delta \vec{E} + P_2 \Delta \vec{H}), \\ \operatorname{rot} \vec{E} &= i\omega(\mu_{\Pi} \vec{H} + Z_{\Pi} \vec{E} + m_1 \Delta \vec{H} + m_2 \Delta \vec{E}), \end{aligned} \quad (4)$$

где Δ – оператор Лапласа,

$$\begin{aligned}\varepsilon_{\Pi} &= \varepsilon_0 \left(\varepsilon_r + 3 \int_0^1 K_1(p) p^2 dp \right), \quad G_{\Pi} = \frac{1}{c} \left(G_r + 3 \int_0^1 K_3(p) p^2 dp \right), \\ P_1 &= \frac{\varepsilon_0}{2} R_1^2 \int_0^1 K_1(p) p^4 dp, \quad P_2 = \frac{1}{2c} R_3^2 \int_0^1 K_3(p) p^4 dp, \\ \mu_{\Pi} &= \mu_0 \left(\mu_r + 3 \int_0^1 K_2(p) p^2 dp \right), \quad Z_{\Pi} = \frac{1}{c} \left(Z_r + 3 \int_0^1 K_4(p) p^4 dp \right), \\ m_1 &= \frac{\mu_0}{2} R_2^2 \int_0^1 K_2(p) p^4 dp, \quad m_2 = \frac{1}{2c} R_4^2 \int_0^1 K_4(p) p^4 dp.\end{aligned}$$

Доказательство. Компоненты векторов $\vec{E}(\vec{r}_0) = E_x(\vec{r}_0)\vec{e}_x + E_y(\vec{r}_0)\vec{e}_y + E_z(\vec{r}_0)\vec{e}_z$, $\vec{H}(\vec{r}_0) = H_x(\vec{r}_0)\vec{e}_x + H_y(\vec{r}_0)\vec{e}_y + H_z(\vec{r}_0)\vec{e}_z$ разложим в ряды Тейлора в окрестности точки M . При этом ограничимся слагаемыми ряда до второго порядка включительно, пренебрегая величинами третьего порядка малости:

$$\begin{aligned}E_{\alpha}(\vec{r}_0) &= E_{\alpha}(\vec{r}) + \frac{\partial E_{\alpha}(\vec{r})}{\partial x}(x_0 - x) + \frac{\partial E_{\alpha}(\vec{r})}{\partial y}(y_0 - y) + \frac{\partial E_{\alpha}(\vec{r})}{\partial z}(z_0 - z) + \\ &+ \frac{1}{2} \left[\frac{\partial^2 E_{\alpha}(\vec{r})}{\partial x^2}(x_0 - x)^2 + \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial y^2}(y_0 - y)^2 + \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial z^2}(z_0 - z)^2 \right] + \\ &+ \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial x \partial y}(x_0 - x)(y_0 - y) + \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial x \partial z}(x_0 - x)(z_0 - z) + \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial y \partial z}(y_0 - y)(z_0 - z),\end{aligned}\tag{5}$$

$$\begin{aligned}H_{\alpha}(\vec{r}_0) &= H_{\alpha}(\vec{r}) + \frac{\partial H_{\alpha}(\vec{r})}{\partial x}(x_0 - x) + \frac{\partial H_{\alpha}(\vec{r})}{\partial y}(y_0 - y) + \frac{\partial H_{\alpha}(\vec{r})}{\partial z}(z_0 - z) + \\ &+ \frac{1}{2} \left[\frac{\partial^2 H_{\alpha}(\vec{r})}{\partial x^2}(x_0 - x)^2 + \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial y^2}(y_0 - y)^2 + \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial z^2}(z_0 - z)^2 \right] + \\ &+ \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial x \partial y}(x_0 - x)(y_0 - y) + \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial x \partial z}(x_0 - x)(z_0 - z) + \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial y \partial z}(y_0 - y)(z_0 - z),\end{aligned}$$

где $\alpha = x, y, z$.

Вычислим компоненты $P_{\alpha}^{(1)}$, $P_{\alpha}^{(2)}$ электрической поляризации (2). Подставляя (5) в (2), получим

$$\begin{aligned}P_{\alpha}^{(1)}(\vec{r}) &= \varepsilon_0 \left(I_0^{(1)} E_{\alpha}(\vec{r}) + I_1^{(1)} \frac{\partial E_{\alpha}(\vec{r})}{\partial x} + I_2^{(1)} \frac{\partial E_{\alpha}(\vec{r})}{\partial y} + I_3^{(1)} \frac{\partial E_{\alpha}(\vec{r})}{\partial z} + \right. \\ &+ I_{11}^{(1)} \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial x^2} + I_{22}^{(1)} \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial y^2} + I_{33}^{(1)} \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial z^2} + I_{12}^{(1)} \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial x \partial y} + I_{13}^{(1)} \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial x \partial z} + I_{23}^{(1)} \frac{\partial^2 E_{\alpha}(\vec{r})}{\partial y \partial z} \left. \right),\end{aligned}\tag{6}$$

$$\begin{aligned}P_{\alpha}^{(2)}(\vec{r}) &= \frac{1}{c} \left(I_0^{(3)} H_{\alpha}(\vec{r}) + I_1^{(3)} \frac{\partial H_{\alpha}(\vec{r})}{\partial x} + I_2^{(3)} \frac{\partial H_{\alpha}(\vec{r})}{\partial y} + I_3^{(3)} \frac{\partial H_{\alpha}(\vec{r})}{\partial z} + \right. \\ &+ I_{11}^{(3)} \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial x^2} + I_{22}^{(3)} \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial y^2} + I_{33}^{(3)} \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial z^2} + I_{12}^{(3)} \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial x \partial y} + I_{13}^{(3)} \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial x \partial z} + I_{23}^{(3)} \frac{\partial^2 H_{\alpha}(\vec{r})}{\partial y \partial z} \left. \right).\end{aligned}$$

Интегралы, входящие в разложения (6), вычислим аналитически:

$$\begin{aligned}
 I_0^{(j)} &= \frac{1}{V_j} \int_{D_M} K_j(p_j) d\vec{r}_0 = \kappa_j = 3 \int_0^1 K_j(p) p^2 dp, \\
 I_1^{(j)} &= \frac{1}{V_j} \int_{D_M} K_j(p) (x_0 - x) d\vec{r}_0 = 0, \quad I_2^{(j)} = \frac{1}{V_j} \int_{D_M} K_j(p) (y_0 - y) d\vec{r}_0 = 0, \\
 I_3^{(j)} &= \frac{1}{V_j} \int_{D_M} K_j(p) (z_0 - z) d\vec{r}_0 = 0, \\
 I_{11}^{(j)} &= \frac{1}{2V_j} \int_{D_M} K_j(p) (x_0 - x)^2 d\vec{r}_0 = g_j = \frac{1}{2} R_j^2 \int_0^1 K_j(p) p^4 dp, \\
 I_{22}^{(j)} &= \frac{1}{2V_j} \int_{D_M} K_j(p) (y_0 - y)^2 d\vec{r}_0 = g_j, \quad I_{33}^{(j)} = \frac{1}{2V_j} \int_{D_M} K_j(p) (z_0 - z)^2 d\vec{r}_0 = g_j, \\
 I_{12}^{(j)} &= \frac{1}{V_j} \int_{D_M} K_j(p_j) (x_0 - x) (y_0 - y) d\vec{r}_0 = 0, \quad I_{13}^{(j)} = \frac{1}{V_j} \int_{D_M} K_j(p_j) (x_0 - x) (z_0 - z) d\vec{r}_0 = 0, \\
 I_{23}^{(j)} &= \frac{1}{V_j} \int_{D_M} K_j(p_j) (y_0 - y) (z_0 - z) d\vec{r}_0 = 0.
 \end{aligned} \tag{7}$$

Подставляя значения интегралов (7) в (6), получим формулы для компонент вектора электрической поляризации:

$$P_\alpha^{(1)}(\vec{r}) = \varepsilon_0 (\kappa_1 E_\alpha(\vec{r}) + g_1 \Delta E_\alpha(\vec{r})), \quad P_\alpha^{(2)}(\vec{r}) = \frac{1}{c} (\kappa_3 H_\alpha(\vec{r}) + g_3 \Delta H_\alpha(\vec{r})).$$

Тогда вектор электрической поляризации определяется формулой

$$\vec{P}(\vec{r}) = \varepsilon_0 (\kappa_1 \vec{E}(\vec{r}) + g_1 \Delta \vec{E}(\vec{r})) + \frac{1}{c} (\kappa_3 \vec{H}(\vec{r}) + g_3 \Delta \vec{H}(\vec{r})). \tag{8}$$

После аналогичных преобразований получим формулу для вектора магнитной поляризации

$$\vec{m}(\vec{r}) = \mu_0 (\kappa_2 \vec{H}(\vec{r}) + g_2 \Delta \vec{H}(\vec{r})) + \frac{1}{c} (\kappa_4 \vec{E}(\vec{r}) + g_4 \Delta \vec{E}(\vec{r})). \tag{9}$$

Подставляя выражения (8), (9) в уравнения (1), приходим к требуемым дифференциальным уравнениям (4).

Базисные плоские электромагнитные поля в биизотропных средах с пространственной дисперсией. Построим полную систему плоских электромагнитных полей, распространяющихся в БСПД вдоль заданного направления в пространстве, т. е. полей \vec{E}, \vec{H} , удовлетворяющих уравнениям (4). Поля аналитически выразим через базисные волновые поля [5, с. 96]:

$$\begin{aligned}
 \vec{W}^{(\mp 1)}(\vec{r}; \alpha_1, \alpha_2; k) &= \frac{i}{\lambda} (\alpha_2 \vec{e}_x - \alpha_1 \vec{e}_y) \Phi(x, y) \exp(\mp vz), \\
 \vec{W}^{(\mp 2)}(\vec{r}; \alpha_1, \alpha_2; k) &= \frac{1}{k} \left(\mp \frac{iv}{\lambda} (\alpha_1 \vec{e}_x + \alpha_2 \vec{e}_y) + \lambda \vec{e}_z \right) \Phi(x, y) \exp(\mp vz),
 \end{aligned} \tag{10}$$

где $\Phi(x, y) = \exp(i\alpha_1 x + i\alpha_2 y)$, α_1, α_2, k – произвольные комплексные постоянные, $\lambda = \sqrt{\alpha_1^2 + \alpha_2^2}$, $0 \leq \arg \lambda < \pi$, $v = \sqrt{\lambda^2 - k^2}$, $-\frac{\pi}{2} \leq \arg v < \frac{\pi}{2}$, k – волновое число, числа α_1, α_2 характеризуют направление распространения поля, $\vec{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ – волновой вектор в свободном пространстве, где $\alpha_1 = k_0 \cos \varphi_0 \sin \theta_0$, $\alpha_2 = k_0 \sin \varphi_0 \sin \theta_0$, $\alpha_3 = k_0 \cos \theta_0$, $0 \leq \theta_0 < \frac{\pi}{2}$, $0 \leq \varphi_0 < 2\pi$.

Для полей (10) выполнены формулы [5, с. 98]

$$\begin{aligned} \operatorname{rot} \vec{W}^{(\mp 1)} &= k \vec{W}^{(\mp 2)}, \quad \operatorname{rot} \vec{W}^{(\mp 2)} = k \vec{W}^{(\mp 1)}, \\ \operatorname{div} \vec{W}^{(\mp 1)} &= 0, \quad \operatorname{div} \vec{W}^{(\mp 2)} = 0, \quad \Delta \vec{W}^{(\mp j)} = -k^2 \vec{W}^{(\mp j)}, \end{aligned} \quad (11)$$

так как $\Delta \vec{W} = \operatorname{grad} \operatorname{div} \vec{W} - \operatorname{rot} \operatorname{rot} \vec{W}$.

Образум поля для биизотропной среды, комбинируя поля (10):

$$\vec{K}^{(\mp)}(\vec{r}; \alpha_1, \alpha_2; k) = \vec{W}^{(\mp 2)}(\vec{r}; \alpha_1, \alpha_2; k) - \vec{W}^{(\mp 1)}(\vec{r}; \alpha_1, \alpha_2; k). \quad (12)$$

Из соотношений (11) следуют формулы для поля (12):

$$\operatorname{div} \vec{K}^{(\mp)} = 0, \quad \operatorname{rot} \vec{K}^{(\mp)} = -k \vec{K}^{(\mp)}, \quad \Delta \vec{K}^{(\mp)} = -k^2 \vec{K}^{(\mp)}. \quad (13)$$

Построим электромагнитное поле в БСПД вида

$$\vec{E} = E_0 \vec{K}^{(\mp)}(\vec{r}; \alpha_1, \alpha_2; k), \quad \vec{H} = p E_0 \vec{K}^{(\mp)}(\vec{r}; \alpha_1, \alpha_2; k), \quad (14)$$

где k, p – постоянные, подлежащие определению; E_0 – постоянная с физической размерностью $[E_0] = \frac{\mathbf{B}}{\mathbf{M}}$, $|E_0| = 1$.

Подставим поля (14) в уравнения (4). Из первого уравнения с учетом формул (13) получим соотношение

$$ip \frac{k}{\omega} + \varepsilon_{\Pi} + p G_{\Pi} = k^2 (P_1 + p P_2). \quad (15)$$

Из второго уравнения (4) следует

$$-i \frac{k}{\omega} + p \mu_{\Pi} + Z_{\Pi} = k^2 (m_2 + p m_1). \quad (16)$$

Разделим уравнение (15) на (16) и исключим k^2 . После соответствующих преобразований выразим k через величину p :

$$k = i\omega \frac{a_2 p^2 + a_1 p + a_0}{b_2 p^2 + b_1 p + b_0}, \quad (17)$$

где

$$\begin{aligned} a_0 &= \varepsilon_{\Pi} m_2 - Z_{\Pi} P_1, \quad a_1 = m_2 G_{\Pi} + \varepsilon_{\Pi} m_1 - \mu_{\Pi} P_1 - Z_{\Pi} P_2, \\ a_2 &= m_1 G_{\Pi} - \mu_{\Pi} P_2, \quad b_0 = P_1, \quad b_1 = m_2 + P_2, \quad b_2 = m_1. \end{aligned}$$

Подставив (17) в соотношение (16), получим алгебраическое уравнение пятой степени относительно величины p :

$$(Z_{\Pi} + p\mu_{\Pi})(b_2p^2 + b_1p + b_0)^2 + (a_2p^2 + a_1p + a_0)(b_2p^2 + b_1p + b_0) + \omega^2(m_2 + pm_1)(a_2p^2 + a_1p + a_0) = 0. \quad (18)$$

Уравнение (18) имеет пять корней $p = p_s, s = 1, 2, 3, 4, 5$, в общем случае комплексных. С помощью формулы (17) определим значения волнового числа $k = k_s, s = 1, 2, 3, 4, 5$. Значения p_s, k_s определяют электромагнитные поля (14). Полученные результаты сформулируем в виде теоремы.

Теорема 2. Плоские электромагнитные поля, распространяющиеся в биизотропной среде с пространственной дисперсией, удовлетворяют уравнениям (4) и определяются формулами

$$\vec{E} = E_0 \vec{K}^{(\mp s)}(\vec{r}; \alpha_1, \alpha_2), \quad \vec{H} = p_s E_0 \vec{K}^{(\mp s)}(\vec{r}; \alpha_1, \alpha_2), \quad s = 1, 2, 3, 4, 5, \quad (19)$$

где

$$\vec{K}^{(\mp s)} = \vec{W}^{(\mp 2)}(\vec{r}; \alpha_1, \alpha_2; k_s) - \vec{W}^{(\mp 1)}(\vec{r}; \alpha_1, \alpha_2; k_s),$$

$$\vec{W}^{(\mp 1)}(\vec{r}; \alpha_1, \alpha_2; k_s) = \frac{i}{\lambda} (\alpha_2 \vec{e}_x - \alpha_1 \vec{e}_y) \Phi(x, y) \exp(\mp v_s z),$$

$$\vec{W}^{(\mp 2)}(\vec{r}; \alpha_1, \alpha_2; k_s) = \frac{1}{k_s} \left(\mp \frac{iv_s}{\lambda} (\alpha_1 \vec{e}_x + \alpha_2 \vec{e}_y) + \lambda \vec{e}_z \right) \Phi(x, y) \exp(\mp v_s z),$$

$$v_s = \sqrt{\lambda^2 - k_s^2}, \quad -\frac{\pi}{2} \leq \arg v_s < \frac{\pi}{2}.$$

Коэффициенты p_j являются корнями уравнения

$$B_5 p^5 + B_4 p^4 + B_3 p^3 + B_2 p^2 + B_1 p + B_0 = 0, \quad (20)$$

где

$$B_5 = \mu_{\Pi} b_2^2 + \omega^2 m_1 a_2^2, \quad B_4 = b_2^2 Z_{\Pi} + b_2 (a_2 + 2\mu_{\Pi} b_1) + \omega^2 a_2 (m_2 a_2 + 2m_1 a_1),$$

$$B_3 = a_1 b_2 + b_1 a_2 + \mu_{\Pi} b_1^2 + 2b_2 (b_1 Z_{\Pi} + \mu_{\Pi} b_0) + \omega^2 (m_1 a_1^2 + 2m_1 a_0 a_2 + 2m_2 a_1 a_2),$$

$$B_2 = a_0 b_2 + a_1 b_1 + a_2 b_0 + b_1^2 Z_{\Pi} + 2b_0 (b_2 Z_{\Pi} + \mu_{\Pi} b_1) + \omega^2 (m_2 a_1^2 + 2m_1 a_0 a_1 + 2m_2 a_0 a_2),$$

$$B_1 = a_0 b_1 + a_1 b_0 + \mu_{\Pi} b_0^2 + 2b_0 b_1 Z_{\Pi} + \omega^2 a_0 (m_1 a_0 + 2m_2 a_1),$$

$$B_0 = a_0 b_0 + b_0^2 Z_{\Pi} + \omega^2 m_2 a_0^2.$$

Волновые числа k_s определяются формулой

$$k_s = i\omega \frac{a_2 p_s^2 + a_1 p_s + a_0}{b_2 p_s^2 + b_1 p_s + b_0}. \quad (21)$$

Базисные сферические электромагнитные поля в биизотропных средах с пространственной дисперсией. Построим аналитически полную систему сферических электромагнитных полей в сферической системе координат $Or\theta\varphi$, распространяющихся в среде и удовлетворяющих уравнениям (4). Поля \vec{E}, \vec{H} выразим через базисные сферические волновые поля [5, с. 118]:

$$\vec{n}_{mn}(\vec{r}, k) = \frac{n(n+1)}{kr} j_n(kr) Y_n^m(\theta, \varphi) \vec{e}_r + g_n(kr) \vec{\Pi}_{mn}(\theta, \varphi), \quad (22)$$

$$\vec{m}_{mn}(\vec{r}, k) = j_n(kr) \vec{T}_{mn}(\theta, \varphi), \quad n = 1, 2, \dots, \quad m = 0, \pm 1, \dots, \pm n;$$

$$\tilde{\vec{n}}_{mn}(\vec{r}, k) = \frac{n(n+1)}{kr} h_n^{(1)}(kr) Y_n^m(\theta, \varphi) \vec{e}_r + g_n^{(1)}(kr) \vec{\Pi}_{mn}(\theta, \varphi), \quad (23)$$

$$\tilde{\vec{m}}_{mn}(\vec{r}, k) = h_n^{(1)}(kr) \vec{T}_{mn}(\theta, \varphi), \quad n=1, 2, \dots, \quad m=0, \pm 1, \dots, \pm n,$$

где k – волновое число, $Y_n^m(\theta, \varphi) = P_n^m(\cos\theta) e^{im\varphi}$, $j_n(x)$, $h_n^{(1)}(x)$ – сферические функции Бесселя, $P_n^m(x)$ – присоединенная функция Лежандра,

$$g_n(x) = \frac{1}{x} \frac{d}{dx} (x j_n(x)), \quad g_n^{(1)}(x) = \frac{1}{x} \frac{d}{dx} (x h_n^{(1)}(x)),$$

$$\vec{\Pi}_{mn}(\theta, \varphi) = \left(\frac{\partial}{\partial \theta} P_n^m(\cos\theta) \vec{e}_\theta + \frac{im}{\sin\theta} P_n^m(\cos\theta) \vec{e}_\varphi \right) e^{im\varphi},$$

$$\vec{T}_{mn}(\theta, \varphi) = \left(\frac{im}{\sin\theta} P_n^m(\cos\theta) \vec{e}_\theta - \frac{\partial}{\partial \theta} P_n^m(\cos\theta) \vec{e}_\varphi \right) e^{im\varphi}.$$

Для полей (22) выполнены соотношения

$$\operatorname{rot} \vec{n}_{mn} = k \vec{m}_{mn}, \quad \operatorname{rot} \vec{m}_{mn} = k \vec{n}_{mn}, \quad \operatorname{div} \vec{n}_{mn} = 0, \quad (24)$$

$$\operatorname{div} \vec{m}_{mn} = 0, \quad \Delta \vec{n}_{mn} = -k^2 \vec{n}_{mn}, \quad \Delta \vec{m}_{mn} = -k^2 \vec{m}_{mn}.$$

Аналогичные соотношения выполнены и для полей (23).

Построим поля в биизотропной среде, комбинируя поля (22) [5, с.121]:

$$\vec{K}_{mn}(\vec{r}, k) = \vec{n}_{mn}(\vec{r}, k) - \vec{m}_{mn}(\vec{r}, k). \quad (25)$$

Аналогично

$$\tilde{\vec{K}}_{mn}(\vec{r}, k) = \tilde{\vec{n}}_{mn}(\vec{r}, k) - \tilde{\vec{m}}_{mn}(\vec{r}, k). \quad (26)$$

Из соотношений (24) следуют формулы

$$\operatorname{div} \vec{K}_{mn} = 0, \quad \operatorname{rot} \vec{K}_{mn} = -k \vec{K}_{mn}, \quad \Delta \vec{K}_{mn} = -k^2 \vec{K}_{mn},$$

$$\operatorname{div} \tilde{\vec{K}}_{mn} = 0, \quad \operatorname{rot} \tilde{\vec{K}}_{mn} = -k \tilde{\vec{K}}_{mn}, \quad \Delta \tilde{\vec{K}}_{mn} = -k^2 \tilde{\vec{K}}_{mn}.$$

Образует электромагнитные поля в среде БСПД, используя волновые поля (25), (26):

$$\vec{E} = E_0 \vec{K}_{mn}(\vec{r}, k), \quad \vec{H} = p E_0 \tilde{\vec{K}}_{mn}(\vec{r}, k); \quad (27)$$

$$\vec{E} = E_0 \tilde{\vec{K}}_{mn}(\vec{r}, k), \quad \vec{H} = p E_0 \vec{K}_{mn}(\vec{r}, k), \quad (28)$$

где k, p – постоянные, подлежащие определению.

Подставляя поля (27) в дифференциальные уравнения (4), получим систему алгебраических уравнений (15), (16), из которых следует формула (17) и уравнение (20) для определения постоянных p и k . Аналогичные соотношения справедливы и для электромагнитных полей (28).

Теорема 3. Сферические электромагнитные поля, распространяющиеся в биизотропной среде с пространственной дисперсией, удовлетворяют уравнениям (4) и определяются формулами

$$\vec{E} = E_0 \vec{K}_{mn}^{(s)}(\vec{r}), \quad \vec{H} = p_s E_0 \tilde{\vec{K}}_{mn}^{(s)}(\vec{r}), \quad s=1, 2, 3, 4, 5,$$

$$\vec{E} = E_0 \tilde{K}_{mn}^{(s)}(\vec{r}), \quad \vec{H} = p_s E_0 \tilde{K}_{mn}^{(s)}(r), \quad n = 1, 2, \dots; \quad m = 0, \pm 1, \dots, \pm n,$$

где

$$\vec{K}_{mn}^{(s)}(\vec{r}) = \vec{n}_{mn}(\vec{r}, k_s) - \vec{m}_{mn}(\vec{r}, k_s), \quad \tilde{K}_{mn}^{(s)}(\vec{r}) = \tilde{n}_{mn}(\vec{r}, k_s) - \tilde{m}_{mn}(\vec{r}, k_s),$$

коэффициенты p_s являются корнями уравнения (20), k_s – волновые числа (21). ■

Базисные цилиндрические электромагнитные поля в биизотропных средах с пространственной дисперсией. Конструирование электромагнитных полей в цилиндрических координатах $\vec{\rho} = (\rho, \varphi, z)$, аналогичное рассмотренному в предыдущем разделе, приводит к теореме.

Теорема 4. Цилиндрические электромагнитные поля, распространяющиеся в биизотропной среде с пространственной дисперсией, удовлетворяют уравнениям (4) и определяются формулами

$$\vec{E} = E_0 \vec{K}_m^{(\pm s)}(\vec{\rho}), \quad \vec{H} = p_s E_0 \vec{K}_m^{(\pm s)}(\vec{\rho}), \quad m = 0, \pm 1, \pm 2, \dots,$$

где

$$\vec{K}_m^{(\pm s)}(\vec{\rho}) = \vec{M}_m^{(\pm 2)}(\vec{\rho}; \lambda, k_s) - \vec{M}_m^{(\pm 1)}(\vec{\rho}; \lambda, k_s), \quad s = 1, 2, 3, 4, 5,$$

$$\vec{M}_m^{(\pm 1)}(\vec{\rho}; \lambda, k_s) = \vec{V}_m^{(1)}(\lambda \rho) e^{im\varphi \pm v_s z},$$

$$\vec{M}_m^{(\pm 2)}(\vec{\rho}; \lambda, k_s) = \frac{1}{k_s} \left(\pm v_s \vec{V}_m^{(2)}(\lambda \rho) + \lambda J_m(\lambda \rho) \vec{e}_z \right) e^{im\varphi \pm v_s z},$$

$$\vec{V}_m^{(1)}(\lambda \rho) = \frac{im}{\lambda \rho} J_m(\lambda \rho) \vec{e}_\rho - J'_m(\lambda \rho) \vec{e}_\varphi,$$

$$\vec{V}_m^{(2)}(\lambda \rho) = J'_m(\lambda \rho) \vec{e}_\rho + \frac{im}{\lambda \rho} J_m(\lambda \rho) \vec{e}_\varphi,$$

$v_s = \sqrt{\lambda^2 - k_s^2}$, $-\frac{\pi}{2} \leq \arg v_s < \frac{\pi}{2}$, λ – произвольная постоянная [5, с. 131, 139], p_s – корни уравнения (20), k_s – волновые числа (21), $J_m(\cdot)$ – функция Бесселя, $J'_m(\cdot)$ – производная функции Бесселя.

Заключение. Разработан метод аналитического построения монохроматических электромагнитных полей, распространяющихся в биизотропной матрице, которая включает среды с пространственной дисперсией. Предложена математическая модель однородной биизотропной среды с пространственной дисперсией. Методика моделирования основана на преобразования интегро-дифференциальной модели к дифференциальной модели уравнений Максвелла с дифференциальными операторами второго порядка. Уравнения разрешены аналитически и построена система независимых пяти прямых и пяти обратных плоских электромагнитных волн, а также сферических и цилиндрических полей, излучаемых источниками. Поля представлены через базисные волновые поля, используемые в классической электродинамике в декартовых, сферических и цилиндрических координатах.

Работа выполнена в соответствии с заданием 1.1.09 государственной программы научных исследований «Информатика, космос и безопасность» на 2016–2020 гг.

Список использованных источников

1. Гуляев, Ю. В. Метаматериалы: фундаментальные исследования и перспективы применения / Ю. В. Гуляев, А. Н. Лагарьков, С. А. Никитов // Вестник РАН. – 2008. – Т. 78, № 5. – С. 438–457.
2. Cui, T. J. Metamaterials. Theory, design and applications / T. J. Cui, D. R. Smith, R. Lui. – Springer, 2009. – 367 p.
3. Иванов, О. В. Распространение электромагнитных волн в анизотропных и бианизотропных слоистых структурах / О. В. Иванов. – Ульяновск : УлГТУ, 2010. – 262 с.
4. Electromagnetic waves in chiral and bi-isotropic media / I. V. Lindell [et al.]. – Boston : Artech House, 1994. – 324 p.

5. Ерофеенко, В. Т. Аналитическое моделирование в электродинамике / В. Т. Ерофеенко, И. С. Козловская. – М. : Книжный дом «Либроком», 2014. – 304 с.
6. Неганов, В. А. Отражающие, волноведущие и излучающие структуры с киральными элементами / В. А. Неганов, О. В. Осипов. – М. : Радио и связь, 2006. – 280 с.
7. Ерофеенко, В. Т. Многократная фокусировка электромагнитного поля магнитного диполя плоскопараллельной двухслойной линзой из метаматериала / В. Т. Ерофеенко // Информатика. – 2016. – № 4(52). – С. 5–19.
8. Ерофеенко, В. Т. Преобразование симметричных волн круглого волновода на бианизотропно-гиротропной перегородке / В. Т. Ерофеенко, А. К. Синицын // Радиотехника и электроника. – 2016. – Т. 61, № 11. – С. 1039–1048.
9. Проникновение электромагнитных волн через композитные экраны, содержащие идеально проводящие спирали / В. Т. Ерофеенко [и др.] // Инженерно-физический журнал. – 2011. – Т. 84, № 4. – С. 740–746.
10. Ерофеенко, В. Т. Дифракция плоской электромагнитной волны на плоскостной структуре из биизотропных материалов / В. Т. Ерофеенко, С. В. Малый // Информатика. – 2012. – № 1(33). – С. 58–65.
11. Ерофеенко, В. Т. Экранирование электромагнитных полей экранами из матричных композитов, содержащих биизотропные частицы / В. Т. Ерофеенко, В. Ф. Бондаренко // Информатика. – 2014. – № 3(43). – С. 28–43.
12. Ерофеенко, В. Т. Красивая задача проникновения электромагнитных полей дипольных источников через биизотропный экран / В. Т. Ерофеенко // Вестник БГУ. Сер. 1. – 2012. – № 2. – С. 71–76.
13. Ерофеенко, В. Т. Численное исследование взаимодействия электромагнитных полей электрического и магнитного диполей с композитным экраном / В. Т. Ерофеенко, В. Ф. Бондаренко // Известия НАН Беларуси. Сер. физ.-техн. наук. – 2013. – № 4. – С. 113–120.
14. Ерофеенко, В. Т. Преобразование пучков электромагнитных волн при прохождении через экран из кирального метаматериала / В. Т. Ерофеенко, В. Ф. Бондаренко // Информатика. – 2013. – № 1(37). – С. 5–17.
15. Ерофеенко, В. Т. Взаимодействие экспоненциально затухающих осциллирующих электромагнитных полей с многослойными композитными экранами / В. Т. Ерофеенко, В. Ф. Бондаренко // Вестник БГУ. Сер. 1. – 2014. – № 1. – С. 62–67.
16. Виноградов, А. П. Электродинамика композитных материалов / А. П. Виноградов. – М. : Эдиториал УРСС, 2001. – 208 с.
17. Агранович, В. М. Кристаллооптика с учетом пространственной дисперсии и теория экситонов / В. М. Агранович, В. Л. Гинзбург. – М. : Наука, 1979. – 432 с.
18. Силин, Р. А. Обратные волны и пространственная дисперсия / Р. А. Силин, И. Р. Тимошина // Радиотехника и электроника. – 2012. – Т. 57, № 7. – С. 725–733.
19. Ерофеенко, В. Т. Моделирование распространения электромагнитных волн в средах с пространственной дисперсией / В. Т. Ерофеенко // Информатика. – 2017. – № 3(55). – С. 5–12.
20. Виноградов, А. П. К вопросу о форме материальных уравнений в электродинамике / А. П. Виноградов // Успехи физических наук. – 2002. – Т. 172, № 3. – С. 363–370.

References

1. Gulyaev Yu. V., Lagarkov A. N., Nikitov S. A. Metamaterialy: fundamental'nye issledovaniya i perspektivy primeneniya [Metamaterials: fundamental research and prospects of application]. Vestnik RAN [Bulletin of the Russian Academy of Sciences], 2008, vol. 78, no. 5, pp. 438–457 (in Russian).
2. Cui T. J., Smith D. R., Liu R. *Metamaterials. Theory, Design and Applications*. Springer, 2009, 367 p.
3. Ivanov O. V. Rasprostranenie jelektromagnitnyh voln v anizotropnyh i bianizotropnyh sloistyh strukturah. *Propagation of Electromagnetic Waves in Anisotropic and Bianisotropic Layered Structures*. Ulyanovsk, UISTU Publ., 2010, 262 p. (in Russian).
4. Lindell I. V., Sihvola A. H., Viitanen A. J., Tretyakov S. A. *Electromagnetic Waves in Chiral and Bi-isotropic Media*. Boston, Artech House, 1994, 324 p.
5. Erofeenko V. T., Kozlovskaja I. S. Analiticheskoe modelirovanie v jelektrodinamike. *Analytical Modeling in Electrodynamics*. Moscow, Knizhnyi dom "Librocom" Publ., 2014, 304 p. (in Russian).
6. Neganov V. A., Osipov O. V. Otrazhajushhie, volnovedushhie i izluchajushhie struktury s kiral'nymi jelementami. *Reflecting, Waveguiding and Radiating Structures with Chiral Elements*. Moscow, Radio i svyaz Publ., 2006, 280 p. (in Russian).
7. Erofeenko V. T. Mnogokratnaja fokusirovka jelektromagnitnogo polja magnitnogo dipolja ploskoparallel'noj dvuhslojnoj linzoy iz metamateriala [Multiple focusing of the electromagnetic field of a magnetic dipole by a plane-parallel two-layer lens made of metamaterial]. Informatika [Informatics], 2016, no. 4(52), pp. 5–19 (in Russian).
8. Erofeenko V. T., Sinitsyn A. K. Preobrazovanie simmetrichnyh voln kruglogo volnovoda na bianizotropno-girotropnoj peregorodke [Transformation of symmetric waves of a circular waveguide on a bianisotropically gyrotropic partition]. Radiotekhnika i elektronika [Radio Engineering and Electronics], 2016, vol. 61, no. 11, pp. 1039–1048 (in Russian).
9. Erofeenko V. T., Demidchik V. I., Malyi S. V., Kornev R. V. Pronikновение jelektromagnitnyh voln cherez kompozitnye jekrany, sodержashhie ideal'no provodjashhie spirali [Penetration of electromagnetic waves]. Inzhenerno-fizicheskij zhurnal [Journal of Engineering Physics and Thermophysics], 2011, vol. 84, no. 4, pp. 740–746 (in Russian).
10. Erofeenko V. T., Malyi S. V. Difrakcija ploskoj jelektromagnitnoj volny na ploskoslojnoj strukture iz biizotropnyh materialov [Diffraction of a plane electromagnetic wave on a plane-layer structure of bi-isotropic materials]. Informatika [Informatics], 2012, no. 1(33), pp. 58–65 (in Russian).

11. Erofeenko V. T., Bondarenko V. F. Jekranirovanie jelektronnih polej jekranami iz matrichnykh kompozitov, sodержashhih biizotropnye chasticy [Screening of electromagnetic fields by screens from matrix composites containing bi-isotropic particles]. *Informatika [Informatics]*, 2014, no. 3(43), pp. 28–43 (in Russian).
12. Erofeenko V. T. Kraevaja zadacha proniknovenija jelektronnih polej dipol'nyh istochnikov cherez biizotropnyj jekran [Boundary-value problem of penetration of electromagnetic fields of dipole sources through a bi-isotropic screen]. *Vestnik BGU [Bulletin of the Belarusian State University. Ser. I]*, 2012, no. 2, pp. 71–76 (in Russian).
13. Erofeenko V. T., Bondarenko V. F. Chislennoe issledovanie vzaimodejstvija jelektronnih polej jelektricheskogo i magnitnogo dipolej s kompozitnym jekranom [Numerical investigation of the interaction of electromagnetic fields of electric and magnetic dipoles with a composite screen]. *Vesti Natsyional'noi akademii navuk Belarusi. Seryia fizika-tekhnichnykh navuk [Proceedings of the National Academy of Sciences of Belarus. Physics and Technical Series]*, 2013, no. 4, pp. 113–120 (in Russian).
14. Erofeenko V. T., Bondarenko V. F. Preobrazovanie puchkov jelektronnih voln pri prohozhdenii cherez jekran iz kiral'nogo metamateriala [Transformation of beams of electromagnetic waves when passing through a chiral metamaterial screen]. *Informatika [Informatics]*, 2013, no. 1(37), pp. 5–17 (in Russian).
15. Erofeenko V. T., Bondarenko V. F. Vzaimodejstvie jeksponencial'no zatuhajushhih oscillirujushhih jelektronnih polej s mnogoslojnymi kompozitnymi jekranami [Interaction of exponentially damped oscillating electromagnetic fields with multilayered composite screens]. *Vestnik BGU [Bulletin of the Belarusian State University. Ser. I]*, 2014, no. 1, pp. 62–67 (In Russian).
16. Vinogradov A. P. Jelektrodinamika kompozitnykh materialov. *Electrodynamics of Composite Materials*. Moscow, Editorial URSS Publ., 2001, 208 p. (in Russian).
17. Agranovich V. M., Ginzburg V. L. Kristallooptika s uchetoм prostranstvennoj dispersii i teorija jeksitonov. *Crystal Optics with Allowance for Spatial Dispersion and the Theory of Excitons*. Moscow, Nauka Publ., 1979, 432 p. (in Russian).
18. Silin R. A., Timoshina I. R. Obratnye volny i prostranstvennaja dispersija [Reverse waves and spatial dispersion]. *Radiotekhnika i elektronika [Radio Engineering and Electronics]*, 2012, vol. 57, no. 7, pp. 725–733 (in Russian).
19. Erofeenko V. T. Modelirovanie rasprostraneniya jelektronnih voln v sredah s prostranstvennoj dispersiej [Modeling the propagation of electromagnetic waves in media with spatial dispersion]. *Informatika [Informatics]*, 2017, no. 3(55), pp. 5–12 (in Russian).
20. Vinogradov A. P. K voprosu o forme material'nykh uravnenij v jelektrodinamike [On the question of the form of material equations in electrodynamics]. *Uspekhi fizicheskikh nauk [Advances in Physical Sciences]*, 2002, vol. 172, no. 3, pp. 363–370 (in Russian).

Информация об авторе

Ерофеенко Виктор Тихонович – доктор физико-математических наук, профессор, главный научный сотрудник научно-исследовательской лаборатории математических методов защиты информации, Учреждение Белорусского государственного университета «НИИ прикладных проблем математики и информатики» (пр. Независимости, 4, 220030, Минск, Республика Беларусь). E-mail: bsu_erofeenko@tut.by

Information about the author

Viktor T. Erofeenko – Dr. Sc. (Physics and Mathematics), Professor, chief research associate of the Research Laboratory of Mathematical Methods of Information Security, Research Institute for Applied Problems of Mathematics and Informatics, Belarusian State University (4, Nezavisimosti Ave., 220030, Minsk, Republic of Belarus). E-mail: bsu_erofeenko@tut.by

ISSN 1816-0301 (print)

УДК 51-7

Поступила в редакцию 16.01.2018

Received 16.01.2018

В. М. Романчук*Белорусский национальный технический университет, Минск, Беларусь***СУБЪЕКТИВНОЕ ОЦЕНИВАНИЕ ВЕРОЯТНОСТИ**

Аннотация. Предлагается метод субъективного измерения вероятности. Для этого вводится понятие величины, которая меняется равномерно от объекта к объекту. Аналогичное базовое понятие успешно используется в некоторых исследованиях для характеристики нефизических величин. С целью обоснования метода анализируется функциональная связь между объективной и субъективной вероятностями, устанавливаемая на основании эмпирических законов Фехнера и Стивенса.

Ключевые слова: субъективная вероятность, рейтинг, законы Фехнера и Стивенса, модель Раша, функция полезности, нечеткие множества, метод анализа иерархий Саати

Для цитирования. Романчук, В. М. Субъективное оценивание вероятности / В. М. Романчук // Информатика. – 2018. – Т. 15, № 2. – С. 74–82.

V. M. Romanchuk*Belarusian National Technical University, Minsk, Belarus***THE SUBJECTIVE MEASUREMENT OF PROBABILITY**

Abstract. The article offers the method of subjective measurement of probability. For this purpose the concept of quantity which changes evenly from object to object is defined. The similar basic concept is successfully used in some researches to characterize non-physical quantities. To justify the method the functional relationship between objective and subjective probabilities established on the basis of Fechner's and Stevens' empirical laws is analyzed.

Keywords: subjective probability, rating, Fechner and Stevens' laws, the Rasch model, function of usefulness, indistinct sets, method of the analysis of hierarchies Saati

For citation. Romanchuk V. M. The subjective measurement of probability. *Informatics*, 2018, vol. 15, no. 2, pp. 74–82 (in Russian).

Введение. Существуют три подхода, с помощью которых получают вероятности: можно найти значения согласно классическому определению вероятности, эмпирически получить вероятности в ходе физического эксперимента или найти значения вероятности на основании ответов эксперта [1].

Обоснованием классического подхода является априорная физическая симметрия, которая подразумевает равную вероятность. Так, если считать, что обе стороны монеты симметричны и материал монеты однороден, то вероятность выпадения каждой из сторон одинакова.

В рамках эмпирического подхода вероятности находят на основе анализа результатов случайного эксперимента. Например, можно подбрасывать монету и считать, сколько раз выпала одна из сторон. При достаточно большом числе испытаний относительная частота появления одной из сторон монеты будет приближенно равна статистической вероятности этого события. В первом и втором случаях вероятность представляет собой число, которое характеризует объективную возможность появления случайного события в данном испытании.

Субъективная вероятность – это число, которое характеризует субъективное отношение к возможности наступления некоторого события.

Постановка задачи. Пусть A_i – конечное число событий из некоторого вероятностного пространства, p_i – вероятности событий A_i , $p_i = P(A_i)$, $i = 1, 2, \dots, n$, $0 < p_i < 1$. Требуется сформулировать метод субъективного измерения вероятностей p_i .

Уточним понятие субъективного измерения. Субъективное измерение вероятностей – это не приближенная оценка объективных вероятностей, а измерение нефизической величины, которое должно рассматриваться по законам измерения нефизических величин. Измерить нефизическую величину можно только в шкале порядка [2, 3]. Недостатком такой шкалы является то, что арифметические операции в ней недопустимы. Поэтому вероятности, найденные в данной шкале, нельзя численно сравнивать. Чтобы определить субъективные вероятности в более сильной шкале, например шкале интервалов, будем находить вероятности косвенно, используя определение рейтинга и аналогию с физическим измерением.

Идею метода субъективного измерения поясним на примере. Допустим, с помощью условного прибора измерены размеры m_1, m_2, m_3 трех объектов и получены показания прибора $r_1=1, r_2=2$ и $r_3=3$ (значения r_i будем называть рейтингами). Считаем, что деления на шкале прибора расположены равномерно, но нет числовых отметок шкалы. Начало отсчета произвольно. Делениям шкалы могут соответствовать значения величины или логарифмы значений величины в зависимости от способа измерения. Способ измерения не задан. Тем не менее можно найти рейтинги, произвольно выбрать способ измерения, определить m_1, m_2, m_3 с точностью до неизвестных постоянных. Тогда для первого способа $m_1=\lambda r_1+\beta, m_2=\lambda r_2+\beta, m_3=\lambda r_3+\beta$, для второго способа $\ln(m_1)=\lambda r_1+\beta, \ln(m_2)=\lambda r_2+\beta, \ln(m_3)=\lambda r_3+\beta$ ($\lambda>0, \beta$ – неизвестные постоянные). В итоге появляется возможность сравнивать размеры объектов друг с другом. Аналогично будем предполагать, что при субъективном измерении вероятностей находятся рейтинги вероятностей с помощью эксперта. Далее на основании рейтингов (с точностью до неизвестных постоянных) получаем субъективные вероятности для двух способов измерения. Важно отметить, что в методе субъективного измерения выбор способа измерения осуществляется априори, поэтому оба способа возможны. Найденные таким образом субъективные вероятности можно сравнивать.

Классическая схема субъективного измерения. В определении классической вероятности понятие симметрии исходов является аксиоматическим и принимается на основании мнения эксперта. Например, при подбрасывании монеты эксперт может интуитивно считать, что стороны монеты достаточно симметричны и будут выпадать с одинаковой вероятностью. Классическое определение вероятности сводит вычисление вероятности к понятию событий, одинаковых по вероятности. Введем понятие величины, которая изменяется равномерно от объекта к объекту, если расположить объекты в порядке возрастания этой величины. Приведем примеры такой величины:

1. Астроном Гиппарх, наблюдая за звездами, разделил их по яркости на шесть величин, где первая величина – самый яркий объект, а шестая – наиболее тусклый [4]. Промежуточные величины он распределил равномерно между оставшимися звездами.

2. Фехнер рассматривал «едва заметные различия» субъективных ощущений [5]. Например, эксперт может сравнивать веса двух объектов. Вес второго объекта увеличивается до тех пор, пока эксперт не скажет «стал больше». Таким образом последовательно строятся объекты, которые упорядочены по возрастанию веса и едва заметно различаются. «Едва заметные различия» Фехнер считал равными потому, что эксперт просто отмечает, когда наступает едва заметное различие, и субъективно для него все различия одинаковы.

3. Терстоун использовал шкалу «равнокажущихся» интервалов для измерения исследуемых психологических и социальных характеристик [5]. Отличительной чертой шкалы Терстоуна является то, что интервалы между показателями соседних высказываний субъективно примерно одинаковы. Такое свойство шкалы достигается за счет метода ее построения.

В приведенных примерах рассматриваются последовательности объектов, величины которых изменяются равномерно (распределены равномерно, имеют «едва заметные различия», принадлежат «равнокажущимся» интервалам). Будем считать, что эксперт может указать последовательность объектов, величины которых изменяются равномерно от объекта к объекту.

Значение величины выражает размер объекта в тех или иных единицах измерения. Субъективное или объективное сравнение размеров опытным путем является единственным способом получения измерительной информации. Основных способов численного сравнения размеров всего два: сравнение разности размеров или отношения размеров величины [3]. Пусть A_1, A_2, \dots, A_n – конечное число объектов, для которых существует величина X , принимающая

значения x_i , $x_i = x(A_i)$, $i = 1, 2, \dots, n$. Будем считать, что если величина X изменяется равномерно от объекта к объекту, то разности или отношения последовательных значений величины постоянны. Для определенности будем считать, что значения величины X расположены в порядке возрастания. Это означает, что для первого способа сравнения

$$x_{i+1} - x_i = \lambda, x_i, x_{i+1} \in R, \lambda > 0,$$

для второго способа сравнения

$$x_{i+1}/x_i = \lambda, x_i, x_{i+1} \in R^+, \lambda > 1,$$

где $i = 1, 2, \dots, n - 1$, λ – неизвестная постоянная, R – множество всех действительных чисел, R^+ – множество всех положительных чисел. Следовательно, если выбран первый способ сравнения, выполняется

$$x_i - x_j = \lambda_1(r_i - r_j), x_i, x_j \in R, \lambda_1 > 0, \quad (1)$$

если выбран второй способ сравнения, выполняется

$$\ln(x_i/x_j) = \lambda_2(r_i - r_j), x_i, x_j \in R^+, \lambda_2 > 0, \quad (2)$$

где $r_i = i$, $r_j = j$, $i = 1, 2, \dots, n - 1$, $j = 1, 2, \dots, n$, λ_1, λ_2 – неизвестные постоянные.

Функцию $r_i = r(x_i)$ будем называть рейтингом величины X . В данном случае $r_i = i$, r_i – значения рейтинга. Можно сформулировать обратную задачу, найти значения величины X , если на основании субъективных наблюдений известен рейтинг и не определен способ сравнения: (1) или (2).

Пусть A_1, A_2, \dots, A_n – конечное число объектов, для которых существует величина X , принимающая значения x_i , $x_i = x(A_i)$, $i = 1, 2, \dots, n$.

Определение 1. Пусть определены значения рейтинга $r_i = r(x_i) = i$, $i = 1, 2, \dots, n$, и не определен способ сравнения. Тогда существуют значения u_1, u_2, \dots, u_n и значения v_1, v_2, \dots, v_n , такие, что

$$u_i - u_j = \lambda_1(r_i - r_j), u_i, u_j \in R, \lambda_1 > 0; \quad (3)$$

$$\ln(v_i/v_j) = \lambda_2(r_i - r_j), v_i, v_j \in R^+, \lambda_2 > 0, \quad (4)$$

где λ_1, λ_2 – неизвестные постоянные, $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$. Значения $u_i = u(x_i)$ и $v_i = v(x_i)$ будем называть значениями величины X , которые получены различными способами субъективного измерения.

Замечание 1. Последовательность u_i , $i = 1, 2, \dots, n$, является арифметической, а последовательность v_i , $i = 1, 2, \dots, n$, – геометрической.

Таким образом, определена классическая схема субъективного измерения, которую можно обобщить.

Аксиоматическая схема субъективного измерения. Дадим определение рейтинга, используя равенства (1) и (2). Пусть величина X принимает значения x_i , $i = 1, 2, \dots, n$.

Определение 2. Рейтингом величины X называется функция $r_i = r(x_i)$, $i = 1, 2, \dots, n$, такая, что имеют место (1) или (2), где $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$, λ_1, λ_2 – неизвестные постоянные, r_i – значение рейтинга.

Из определения 2 следует, что рейтинг может быть получен различными способами измерения величины X : (1) или (2). Аналогично классической схеме субъективного измерения можно сформулировать обратную задачу – найти величину X , если известны значения рейтинга и не определен способ измерения величины X .

Определение 3. Пусть величина X принимает значения x_i , $i = 1, 2, \dots, n$, причем для величины X определены значения рейтинга r_i , $i = 1, 2, \dots, n$, и не определен способ сравнения. Тогда существуют значения u_1, u_2, \dots, u_n и значения v_1, v_2, \dots, v_n , такие, что выполняются равенства (3) и (4), где $i = 1, 2, \dots, n$, $j = 1, 2, \dots, n$, λ_1, λ_2 – неизвестные постоянные.

Значения $u_i=u(x_i)$ и $v_i=v(x_i)$ будем называть значениями величины X , которые получены различными способами субъективного измерения. Решения систем (3) и (4) всегда существуют, поскольку в качестве решений можно выбрать $u_i=\lambda_1 r_i$ и $\ln(v_i)=\lambda_2 r_i$, $i=1, 2, \dots, n$.

Замечание 2. Если величина X изменяется равномерно, то ее рейтинг меняется равномерно, $r_i=i$ и последовательность u_i является арифметической, а последовательность v_i – геометрической. Этот случай соответствует классической схеме субъективного измерения.

Таким образом, определена аксиоматическая схема субъективного измерения с помощью рейтинга величины. Рейтинг величины будем находить методом парных сравнений.

Матрица парных сравнений. Пусть величина X принимает значения x_i , $x_i=x(A_i)$, $i=1, 2, \dots, n$, тогда существует $r_i=r(x_i)$ – рейтинг величины X . Определим матрицу парных сравнений D с элементами d_{ij} :

$$md_{i,j} = r_i - r_j, \quad (5)$$

где $m>0$, $i=1, 2, \dots, n$, $j=1, 2, \dots, n$.

Предполагается, что элементы $d_{i,j}$ могут быть найдены с помощью интуитивного сравнения размеров объектов A_i и A_j по величине X , $d_{i,j}=d(x_i, x_j)$. Будем считать, что для матрицы парных сравнений D выполняются свойства

$$d_{i,j} = -d_{j,i}; \quad (6)$$

$$d_{i,j} = d_{i,k} + d_{k,j}, \quad (7)$$

где d_{ij} , $d_{i,k}$, $d_{k,j}$ – элементы матрицы парных сравнений D , $i=1, 2, \dots, n$, $j=1, 2, \dots, n$, $k=1, 2, \dots, n$.

Если известна матрица парных сравнений D , которая удовлетворяет условиям (6) и (7), то существует решение системы уравнений (5). Действительно, в качестве решения можно выбрать $r_i=d_{i1}=md(x_i, x_1)$, $i=1, 2, \dots, n$.

Если известна матрица парных сравнений D , которая удовлетворяет условиям (6) и (7), то рейтинг определен с точностью до линейного преобразования. Для доказательства достаточно рассмотреть два различных решения системы линейных уравнений (5): r_{1i} , $i=1, 2, \dots, n$, и r_{2i} , $i=1, 2, \dots, n$. Решения существуют, поскольку выполняются (6) и (7), тогда $m_1 d_{i,j} = r_{1i} - r_{1j}$, $m_2 d_{i,j} = r_{2i} - r_{2j}$, m_1, m_2 – неизвестные постоянные. Следовательно, $m_1(r_{2i} - r_{2j}) = m_2(r_{1i} - r_{1j})$.

Чтобы найти матрицу парных сравнений, в некоторых случаях можно интуитивно выполнить парные сравнения с помощью вербальной шкалы. Вербальную шкалу использовал Т. Саати (T. Saati) [6], чтобы находить отношения значений рейтингов нефизической величины. Приведем вариант такой шкалы (табл. 1), в которой $d_{i,j}$ – результат субъективного сравнения размеров объектов A_i и A_j по величине X .

Таблица 1

Шкала превосходства, $x_i \geq x_j$

Вербальная оценка	Числовая оценка $d_{i,j}$
Отсутствует	0
Небольшое	2
Большое	4
Очень большое	6
Максимальное	8
Промежуточные значения	1, 3, 5, 7
$x_i \geq x_j$	$d_{i,j} = -d_{j,i}$

Пример. Если превосходство x_2 над x_1 «большое», превосходство x_3 над x_1 «очень большое», то $d_{21} = 4$ и $d_{32} = 6$. Следовательно, $md_{21} = r_2 - r_1 = 4m$, $md_{32} = r_3 - r_2 = 6m$. Достаточно найти частное решение: $m=1$, $r_1=0$, $r_2=4$, $r_3=10$. Для значений, найденных способом субъективного измерения (3) и (4), получим $u_2 = u_1 + 4\lambda_1$, $u_3 = u_1 + 10\lambda_1$ и $v_2 = v_1(\lambda_2)^4$, $v_3 = v_1(\lambda_2)^{10}$; $u_1, v_1, \lambda_1 > 0$, $\lambda_2 > 0$ – неизвестные постоянные.

Аксиома субъективного измерения. Будем считать, что эксперт не в состоянии выбрать способ субъективного измерения. Тем не менее способ измерения может быть выбран исследователем. Например, исследователь предлагает эксперту оценить разность или отношение размеров. В этом случае исследователь выводит эксперта за пределы компетенции, поскольку эксперт не может делить и вычитать значения нефизических величин. Чтобы исследователь мог выбирать способ субъективного измерения и находить значения величины, сформулируем аксиому субъективного измерения.

Пусть определен рейтинг величины и не определен способ субъективного измерения.

Аксиома. Способ субъективного измерения величины выбирается априори.

Данное утверждение означает, что для объектов, для которых определен рейтинг и не определен способ измерения, можно произвольно выбрать способ измерения и найти значения величины с помощью равенств (3) или (4). Аксиома отражает особенность метода субъективного измерения. Нефизическая величина существует только в сознании людей. У нее нет измеряемых размеров и, соответственно, нельзя определить способ ее измерения. Эксперт может только найти рейтинг величины. После этого значения величины находятся, хотя и неоднозначно, на основании выбранного способа субъективного измерения. Здесь нет логического противоречия, поскольку значения величины в данном случае – это следствие математической обработки: числа, которые исследователь для удобства приписывает объектам.

В итоге получаем алгоритм метода субъективного измерения:

1. С помощью эксперта определяем элементы матрицы парных сравнений и находим рейтинг величины.

2. Априори выбираем способ измерения.

3. Находим значения величины.

Замечание 3. Способов сравнения существует много, но в теории измерений в качестве основных используют только два: разность и отношение значений [2]. Из соотношений (3) и (4) следует, что если $\lambda_1 = \lambda_2$, то имеет место равенство

$$\ln(v_i/v_j) = u_i - u_j, \quad (8)$$

где $v_i, v_j \in R^+$, $u_i, u_j \in R$, $i=1, 2, \dots, n$, $j=1, 2, \dots, n$.

Равенство (8) является частным случаем более общего равенства

$$\ln(a/b) = \ln(a) - \ln(b), \quad a, b \in R^+,$$

которое в теории групп означает, что группа R^+ по умножению изоморфна группе R по сложению, поэтому деление происходит аналогично вычитанию. Если с точки зрения теории групп изоморфные группы можно не различать, то с позиций субъективного измерения различить изоморфные группы невозможно.

В качестве обоснования аксиомы рассмотрим пример субъективного измерения вероятности. Субъективную вероятность можно интерпретировать как частично определенную вероятностную меру с неопределенным способом измерения.

Субъективная вероятность. Субъективную вероятность будем считать нефизической величиной с априори определяемым способом субъективного измерения. Будем считать, что A_1, A_2, \dots, A_n – конечное число событий из некоторого вероятностного пространства с вероятностями $p_i = P(A_i)$, $i=1, 2, \dots, n$. Пусть с помощью эксперта определен рейтинг вероятности $r_i = r(p_i)$, $i=1, 2, \dots, n$, и не определен способ измерения. Тогда допустим, что верно

$$u_i - u_j = \lambda_1(r_i - r_j), \quad u_i, u_j \in R, \quad \lambda_1 > 0, \quad (9)$$

или

$$\ln(v_i/v_j) = \lambda_2(r_i - r_j), \quad v_i, v_j \in R^+, \quad \lambda_2 > 0, \quad (10)$$

где $i=1, 2, \dots, n$, $j=1, 2, \dots, n$, $d_{i,j} = r_i - r_j$, λ_1, λ_2 – неизвестные постоянные. Значения $u_i = u(p_i)$ и $v_i = v(p_i)$ будем называть субъективными вероятностями, которые получены различными спо-

собами субъективного измерения вероятности P . Если вероятность P изменяется равномерно от события к событию, то $r_i = i$.

Из равенства (9) следует, что $u_i = \lambda_1 r_i + \beta_1$, где λ_1, β_1 – постоянные, которые можно выбрать так, что будет выполняться $0 < u_i < 1$ для всех $i=1, 2, \dots, n$. Аналогично из (10) $\ln(v_i) = \lambda_2 r_i + \beta_2$ и λ_2, β_2 можно выбрать так, что выполняется $0 < v_i < 1$ для всех $i=1, 2, \dots, n$. Следовательно, вероятность, полученная различными способами субъективного измерения, «похожа» на объективную вероятность. В качестве примера рассмотрим геометрическую вероятность.

Геометрическая вероятность события A_i – это отношение меры области $\mu(A_i)$ к мере всей области. Элементы матрицы парных сравнений $d_{i,j}$ в данном случае являются результатом субъективного сравнения мер $\mu(A_i)$ и $\mu(A_j)$, $i=1, 2, \dots, n$, $j=1, 2, \dots, n$. В частности, если площадь областей A_1, A_2, \dots, A_n изменяется равномерно, можно выбрать рейтинг $r_i = i$.

Пример. Область состоит из шести непересекающихся кругов (табл. 2, строка 2). При испытании точка случайно попадает в один из них. Требуется найти субъективную вероятность события «попасть в круг A_i ».

Таблица 2

Круги, площадь которых изменяется равномерно

r_i	1	2	3	4	5	6
A_i						

Чтобы найти вероятность, необходимо знать площади. Пусть исследователь просит эксперта оценить, во сколько раз площадь второго круга больше площади первого круга, и эксперт отвечает «в два раза» (здесь эксперт может назвать любое другое число). Далее, если площади подобраны таким образом, что с точки зрения эксперта они изменяются равномерно, то площадь третьего круга в два раза больше площади второго, площадь четвертого круга в два раза больше площади третьего и т. д. В итоге получим рейтинг площади круга, представленный в табл. 2, строке 1. В данном случае исследователь предложил сравнивать отношения площадей, а эксперт выбрал оценку «больше в два раза».

Пусть теперь исследователь просит оценить, насколько площадь второго круга больше площади первого. Это означает, что исследователь в качестве способа сравнения выбрал разность значений. По условиям эксперимента эксперт не может вычитать размеры площадей (напомним, что субъективно измеренная площадь – это число, которое именно исследователь приписывает каждому объекту на основании рейтинга). Поэтому эксперт может только ответить, например, что превосходство площади второго круга над площадью первого «большое» (значит, разность рейтингов согласно табл. 1 равна числу 4). Далее, поскольку площади кругов изменяются равномерно, эксперт ответит, что превосходство площади третьего круга над площадью второго «большое» и т. д. В этом случае получен рейтинг площади круга (табл. 2, строка 1).

Таким образом, для одних и тех же объектов из табл. 2 рейтинг совпадает, несмотря на разные способы сравнения. Это происходит потому, что эксперт не выполняет операцию деления или вычитания значений величины, а строит последовательность одинаково отличающихся по величине объектов и находит их рейтинг. Эксперт дает фактически один и тот же ответ, несмотря на изменение способа сравнения (способ сравнения в рассматриваемом примере априори указывает исследователь). Таким образом подтверждается аксиома субъективного сравнения. На основании рейтинга $r_i = i$, $i=1, 2, \dots, n$, по формулам (9) и (10) теперь можно определить субъективные вероятности.

Приведенный пример поясняет предлагаемый метод субъективного измерения вероятности. Чтобы получить более весомое подтверждение, обратимся к «настоящим» экспериментам, которые в течение более чем 100 лет проводили психологи. Итогом работы психологов стали закон Фехнера (Fechner), в котором участвуют разности значений субъективной величины, и закон Стивенса (Stevens), в котором сравниваются отношения значений субъективной величины [5].

Субъективная и объективная вероятности. Пусть A_1, A_2, \dots, A_n – конечное число событий из некоторого вероятностного пространства с вероятностями $p_i = P(A_i)$, $i=1, 2, \dots, n$, для которых с помощью эксперта определен рейтинг вероятности $r_i=r(p_i)$, $i=1, 2, \dots, n$, и не определен способ измерения.

Законы Стивенса и Фехнера связывают физические и соответствующие нефизические величины [5]. В качестве физической величины можно рассматривать объективные вероятности p_i , а в качестве нефизической величины – субъективные вероятности $u(p_i)$ или $v(p_i)$, $i=1, 2, \dots, n$.

Закон Фехнера для субъективной вероятности $u_i=u(p_i)$, $i=1, 2, \dots, n$, и объективной вероятности p_i , $i=1, 2, \dots, n$, имеет вид

$$u_i - u_j = c \ln \left(\frac{p_i}{p_j} \right), \quad (11)$$

где $i=1, 2, \dots, n$, $j=1, 2, \dots, n$, c – постоянная.

Закон Стивенса предложен для замены закона Фехнера. Исходя из закона Стивенса зависимость между субъективной вероятностью $v_i = v(p_i)$, $i=1, 2, \dots, n$, и объективной вероятностью p_i , $i=1, 2, \dots, n$, находится из выражения

$$\frac{v_i}{v_j} = \left(\frac{p_i}{p_j} \right)^\alpha, \quad (12)$$

где $i=1, 2, \dots, n$, $j=1, 2, \dots, n$, α – постоянная.

Подставив (9) в закон Фехнера (11), получим

$$\lambda_1 (r_i - r_j) = c \ln \left(\frac{p_i}{p_j} \right), \quad (13)$$

где c, λ_1 – постоянные, p_i, p_j – вероятности, $i, j=1, 2, \dots, n$. Аналогично согласно закону Стивенса (12) на основании (10)

$$\lambda_2 (r_i - r_j) = \alpha \ln \left(\frac{p_i}{p_j} \right), \quad (14)$$

$i=1, 2, \dots, n$, $j=1, 2, \dots, n$, α, λ_2 – постоянные. Видно, что законы Фехнера (11) и Стивенса (12) после перехода к рейтингам совпадают и могут быть записаны в едином виде:

$$\lambda (r_i - r_j) = \ln \left(\frac{p_i}{p_j} \right), \quad (15)$$

где p_i, p_j – объективные вероятности, r_i, r_j – рейтинги вероятностей, $\lambda > 0$, λ – постоянная, $i=1, 2, \dots, n$, $j=1, 2, \dots, n$. Совпадение законов Фехнера и Стивенса подтверждает аксиому субъективного измерения и гипотезу, что «... природа не наделила человека способностью сравнивать между собой разные свойства или их проявления в числовом формате» [3]. Таким образом, обосновывается модель субъективного измерения вероятностей.

Некоторые примеры. Для иллюстрации возможностей метода субъективного измерения вероятности рассмотрим два примера.

Пример 1. Три человека играют в игру, в которой может победить один из участников. Преимущество второго игрока перед первым «небольшое», у третьего по сравнению со вторым – «очень большое». Требуется найти вероятность победы для каждого из участников.

С помощью табл. 1 получим, что $d_{2,1}=2$ и $d_{3,2}=6$. Тогда можно выбрать $m=1$, $r_1=1$, $r_2=3$, $r_3=9$. Предположительно существует связь между объективной вероятностью и рейтингом, аналогичная закону Фехнера и Стивенса (15):

$$\ln \left(\frac{p_i}{p_j} \right) = \lambda (r_i - r_j), \quad (16)$$

где p_i, p_j – объективные значения вероятности, $i, j=1, 2, 3$; λ – постоянная, которая может быть определена по результатам статистических испытаний.

Пример 2. Используя рейтинг, можно построить математическую модель встречи двух шахматистов в одной партии. Предположим, что отношение вероятности выигрыша игроком с номером i к вероятности выигрыша игроком с номером j определяется по формуле (15) и зависит от разности рейтингов игроков r_i и r_j :

$$\frac{p_i}{p_j} = q^{r_i - r_j}, \quad (17)$$

где $q=e^\lambda$ – постоянная. Тогда вероятность p_{ij} выиграть игроку с номером i у игрока с номером j определяется по формуле

$$p_{ij} = \frac{1}{1 + q^{r_i - r_j}}. \quad (18)$$

Для проверки модели (18) можно использовать систему рейтингов Эло [7]. Рейтинги игроков r_i и r_j в системе Эло рассчитывают по специальной методике. Каждой разности рейтингов соответствует своя вероятность выигрыша (табл. 3, первая и вторая колонки).

Таблица 3

Рейтинги Эло (фрагмент)

Разность рейтингов, $r_i - r_j$	Вероятность выигрыша по системе Эло	Разность рейтингов, $r_i - r_j$	Вероятность выигрыша по модели Раша p_{ij}
0–03	0,50	0	0,50
18–25	0,53	25	0,53
47–53	0,57	50	0,57
99–106	0,64	100	0,63
146–153	0,70	150	0,70
198–206	0,76	200	0,75
560–619	0,98	600	0,96
620–734	0,99	700	0,98
735 и более	1,00	900	0,99

Если $q=1,0055$, то на основании формулы (18) получим математическую модель Раша [7], которая аппроксимирует статистические данные таблицы Эло. В этом случае для разности рейтингов r_i и r_j есть возможность рассчитать вероятность выигрыша по формуле (18). Результат такого расчета вероятностей представлен в табл. 3, третья и четвертая колонки.

Заключение. В работе рассматривается метод субъективного измерения вероятностей. Анализируются законы Фехнера и Стивенса, с помощью которых можно сравнить объективные вероятности, на основании результатов субъективного измерения. Установлено, что при переходе к рейтингу законы Фехнера и Стивенса становятся эквивалентными. Метод субъективного измерения нефизических величин может применяться в теории вероятности, в теории нечетких множеств, для нахождения функции полезности и при экспертном анализе систем.

Список использованных источников

1. Статистика для менеджеров с использованием Microsoft Excel / Д. М. Левин [и др.]. – 4-е изд. : пер. с англ. – М. : Вильямс, 2004. – 1312 с.
2. Шишкин, И. Ф. Теоретическая метрология : учебник для вузов : в 2 ч. / И. Ф. Шишкин. – СПб. : Питер, 2010. – Ч. 1 : Общая теория измерений. – 192 с.
3. Шишкин, И. Ф. Измерения нефизических величин (измерения в ноосфере) / И. Ф. Шишкин // Экономика качества [Электронный ресурс]. – 2016. – № 2(14). – Режим доступа: www.eq-journal.ru. – Дата доступа: 16.01.2018.
4. Берри, А. Краткая история астрономии / А. Берри ; пер. с англ. С. Г. Займовского. – М. : Гостехиздат, 1946. – 363 с.
5. Гусев, А. Н. Психологические измерения. Теория. Методы / А. Н. Гусев, И. С. Уточкин. – М. : Аспект Пресс, 2011. – 317 с.
6. Саати, Т. Принятие решений. Метод анализа иерархий : пер. с англ. / Т. Саати. – М. : Радио и связь, 1989. – 316 с.
7. Карнаухов, В. М. Статистическое моделирование некоторых сегментов учебного процесса / В. М. Карнаухов. – М. : РГАУ-МСХА, 2014. – 208 с.

References

1. Levine D. M., Stephan D., Krehbiel T. C., Berenson M. L. Statistika dlia menedzherov s ispol'zovaniem Microsoft Excel. *Statistics for Managers Using Microsoft Excel*. Moscow, Vil'jams Publ., 2004, 1312 p.
2. Shishkin I. F. Teoreticheskaia metrologiia. *Theoretical Metrology. Pt. 1. General Theory of Measurements*. St. Petersburg, Piter Publ, 2010, 192 p. (in Russian).
3. Shishkin I. F. Izmereniia nefizicheskikh velichin (izmereniia v noosfere) [Measurements of not physical quantities (measurements in a noosphere)]. *Jekonomika kachestva [Quality Economy]*, 2016, no. 2(14) (in Russian). Available at: www.eq-journal.ru (accessed 16.01.2018).
4. Berri A. Kratkaia istoriia astronomii. *Short History of Astronomy*. Moscow, Gostekhizdat Publ., 1946, 363 p.
5. Gusev A. N., Utochkin I. S. Psikhologicheskie izmereniia. Teorija. Metody. *Psychological Measurements. Theory. Methods*. Moscow, Aspekt Press Publ., 2011, 317 p. (in Russian).
6. Saati T. Priniatie reshenii. Metod analiza ierarkhii. *Decision Making. Method of the Analysis of Hierarchies*. Moscow, Radio i sviaz Publ., 1989, 316 p.
7. Karnaukhov V. M. Statisticheskoe modelirovanie nekotorykh segmentov uchebnogo protsesssa. *Statistical Modeling of Some Segments of Educational Process*. Moscow, RGAU-MSHA Publ., 2014, 208 p. (in Russian).

Информация об авторе

Романчак Василий Михайлович – доцент кафедры «Инженерная математика», Белорусский национальный технический университет (пр. Независимости, 65, 220013, Минск, Республика Беларусь). E-mail: Romanchak@bntu.by

Information about the author

Vasily M. Romanchak – Associate Professor "Engineering Mathematics", Belarusian National Technical University (65, Nezavisimosti Ave., 220013, Minsk, Republic of Belarus). E-mail: Romanchak@bntu.by

ISSN 1816-0301 (print)

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ*INTELLIGENT SYSTEMS*

УДК 621.397.1

Поступила в редакцию 22.02.2018

Received 22.02.2018

С. В. Кругликов¹, А. Р. Бекиш²

¹*Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск, Беларусь*
²*Военная академия Республики Беларусь, Минск, Беларусь*

**ОБОБЩЕННАЯ СТРУКТУРА ИМИТАЦИОННОЙ МОДЕЛИ
ГРУППОВОЙ ДЕЯТЕЛЬНОСТИ ОПЕРАТОРОВ АСУ**

Аннотация. Предлагается формализованная модель деятельности коллектива операторов автоматизированной системы управления, разработанная на основе применения агрегативного подхода Н. П. Бусленко в виде четырехуровневой агрегативной системы, включающей агрегаты функциональных блоков, задач, операторов, коллектива операторов. В зависимости от характера решаемых задач предлагается использовать модели трех уровней: групповой, индивидуальной деятельности и модель деятельности, учитывающую эргономические факторы. Приводятся обобщенные структурные схемы всех уровней, а также основные требования и принципы функционирования блоков, входящих в модели соответствующих уровней.

Ключевые слова: агрегат, математическая модель, имитационная модель, оператор, групповая деятельность, индивидуальная деятельность

Для цитирования. Кругликов, С. В. Обобщенная структура имитационной модели групповой деятельности операторов АСУ / С. В. Кругликов, А. Р. Бекиш // Информатика. – 2018. – Т. 15, № 2. – С. 83–97.

S. V. Kruglikov¹, A. R. Bekish²

¹*The United Institute of Informatics Problems of the National Academy
of Sciences of Belarus, Minsk, Belarus*
²*The Military Academy of the Republic of Belarus, Minsk, Belarus*

**GENERALIZED STRUCTURE OF THE SIMULATION MODEL
OF THE GROUP ACTIVITY OF ACS OPERATORS**

Abstract. A formalized model of the activity of the collective of operators of an automated control system is developed, developed on the basis of the application of Buloshenko's aggregative approach in the form of a four-level aggregative system that includes aggregates of functional blocks, tasks, operators, and the collective of operators. Depending on the nature of the tasks being solved, it is proposed to use models of three levels: group, individual activity and an activity model that takes ergonomic factors into account. Generalized structural diagrams of all levels are presented, as well as the basic requirements and principles of the functioning of the blocks included in the model of the corresponding levels.

Keywords: aggregate, mathematical model, imitation model, operator, group activity, individual activity

For citation. Kruglikov S. V., Bekish A. R. Generalized structure of the simulation model of the group activity of ACS operators. *Informatics*, 2018, vol. 15, no. 2, pp. 83–97.

Введение. Анализ применения многоуровневых автоматизированных систем управления (АСУ) свидетельствует о том, что комплексное решение задач проектирования групповой и индивидуальной деятельности операторов основной и дежурных смен АСУ возможно лишь в том случае, если в основе моделирующих систем лежит единая математическая схема моделирования, способная описывать поведение непрерывных и дискретных, детерминированных и стохастических систем [1–4]. Математическая схема представляет собой звено при переходе от содер-

жательного к формальному описанию процесса функционирования сложной системы с учетом воздействия внешней среды, т. е. имеет место цепочка «описательная модель – математическая схема – математическая модель» [5]. Введение понятия математической схемы позволяет рассматривать математику не как метод расчета, а как метод мышления, как средство формулирования понятий, что является важным при переходе от словесного описания системы к формальному представлению процесса ее функционирования в виде математической модели.

Представленный подход к формализации процесса функционирования сложной системы «человек – машина» основывается на агрегатной модели Н. П. Бусленко [6, 7], и в нем используется соответствующая терминология. При агрегатном описании сложная система разбивается на подсистемы, сохраняя при этом связи, обеспечивающие их взаимодействие. Если система оказывается сложной, то процесс расчленения продолжается до тех пор, пока не образуются подсистемы, которые в условиях рассматриваемой задачи могут считаться удобными для математического описания. Абстрактное описание модели деятельности коллектива операторов с использованием универсальной математической схемы агрегата (А-схемы) позволит обоснованно разработать обобщенную структурную схему имитационной модели.

Формализация деятельности операторов при разработке имитационной модели. Для формализации групповой деятельности операторов используем универсальный подход, сочетающий в себе комплексное применение методов декомпозиции и агрегирования [6]. Декомпозиция деятельности коллектива операторов в соответствии с принципами универсальности, иерархичности и микромоделирования позволяет выделить такие уровни модели, как типовые функциональные блоки, задачи, отдельные операторы и коллектив операторов.

В силу своей специфики каждый из уровней модели описывается разнородными математическими схемами. Для соблюдения принципов конструктивности и универсальности описания модели при ее реализации необходима унифицированная абстрактная схема, позволяющая единообразно описывать все уровни модели. В качестве такой абстрактной схемы используем агрегативный подход, развитый в работах [6, 7].

Под агрегатом понимается объект, определяемый множествами моментов времени $t \in T$, входных сигналов $x \in X$, управляющих сигналов $g \in G$, выходных сигналов $y \in Y$, состояний $z \in Z$ и конструктивных параметров агрегата $b \in B$, а также случайными операторами переходов в новое состояние H и выходов G , реализующих функции $z(t)$ и $y(t)$, структура которых описана в [7].

Оператор выходов G представляется в виде совокупности операторов G' (вырабатывает очередные моменты выдачи выходных сигналов) и G'' (формирует содержание сигналов). В пространстве состояний агрегата для каждого значения $g \in G$, $b \in B$ задается некоторое множество $Z^Y(g, \beta)$, которое определяет момент выдачи непустого выходного сигнала

$$y = G''[t, x(t), z(t), g(t), b].$$

Оператор G' вырабатывает очередной момент времени выдачи выходного сигнала, определяемый по достижении траекторией $z(t)$ подмножества $Z^Y(g, \beta)$:

$$y = G'[t, z(t), g(t), b].$$

Состояния агрегата $z(t)$ для произвольного момента времени $t > t_0$ рассчитываются по предыдущим состояниям $z(t_0)$ случайным оператором переходов

$$z(t) = H[z(t_0), t].$$

Наряду с состоянием агрегата $z(t)$ будем рассматривать также состояние $z(t_j+0)$, в которое агрегат переходит за «малый» интервал времени. Вид оператора H зависит от того, поступают или не поступают в течение рассматриваемого интервала времени входные и управляющие сигналы. Поэтому он представляется в виде совокупности операторов V' , V'' и U :

$$z(t_j+0) = V' [t, z(t_j), g(t_j), x, b],$$

$$z(t_j+0) = V'' [t, z(t_j), g, b],$$

$$z(t) = U [t, t_j, z(t_j+0), g(t_j), b].$$

Специфика построения модели операторской деятельности такова, что модель нижнего уровня входит в модель более высокого уровня. Для конкретизации агрегативного подхода применительно к решаемой задаче введем понятие «вложенного агрегата».

Будем считать, что агрегат нижнего уровня вложен в агрегат верхнего уровня, если соблюдаются следующие условия:

- агрегат верхнего уровня представляет собой множество связанных между собой агрегатов нижнего уровня;
- множество входных сигналов агрегатов нижнего уровня полностью описывается множеством входных сигналов агрегата верхнего уровня;
- множество выходных сигналов агрегата верхнего уровня полностью описывается множеством выходных сигналов агрегатов нижнего уровня.

Таким образом, модель операторской деятельности представляет собой четырехуровневую агрегативную систему (рис. 1), включающую агрегаты первого уровня (функциональные блоки, $a_{11}, \dots, a_{1k} \in A_1$), второго уровня (задачи, $a_{21}, \dots, a_{2l} \in A_2$), третьего уровня (операторы, $a_{31}, \dots, a_{3m} \in A_3$) и четвертого уровня (коллектив операторов, $a_{41} \in A_4$).

Дадим формализованное описание отдельных уровней модели в терминах агрегативного подхода.

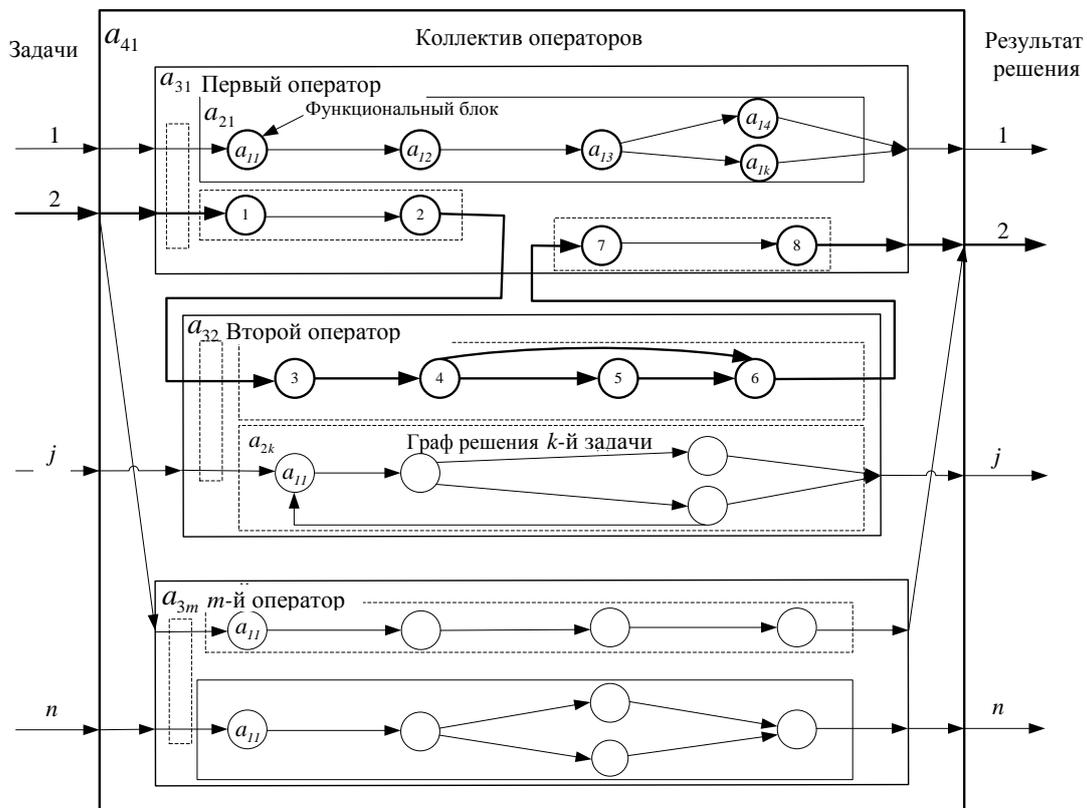


Рис. 1. Представление имитационной модели групповой деятельности операторов в виде четырехуровневой агрегативной системы

Агрегаты первого уровня (функциональные блоки). Агрегат первого уровня представляет собой вероятностный автомат, осуществляющий задержку входных сигналов и их преобразование. Множества входных и выходных сигналов принимают два значения: «0» или «1». Агрегат описывается такими параметрами, как время задержки входного сигнала t_{11} и ошибка преобразования β_{12} с законами распределения $z_{11}(t)$ и $z_{12}(\beta)$ соответственно, числовые характеристики которых определяются по эргономическим справочникам либо экспериментальным путем.

Также в параметрах агрегата может учитываться влияние двух групп факторов: характеристик оператора (уровня обученности, усталости, напряженности деятельности и т. д.) и параметров автоматизированного рабочего места (средств отображения и органов управления).

Если представить отдельные функциональные блоки вершинами некоторого графа, а очередность их следования дугами, соединяющими вершины, то получим наиболее обобщенную математическую модель алгоритма решения задачи в виде вероятностного графа. Основные этапы формирования вероятностного графа алгоритма решения задачи описаны в [8], поэтому здесь остановимся лишь на описании задачи в терминах агрегативного подхода.

Агрегаты второго уровня (задачи). Множества входных и выходных сигналов агрегата второго уровня принимают значения «0» (задача не возникла в процессе деятельности) либо «1» (задача возникла). Агрегат второго уровня полностью описывается множеством функциональных блоков и матрицей переходных вероятностей, указывающей порядок выполнения функциональных блоков.

Как и агрегат первого уровня, агрегат второго уровня реализует задержку и преобразование входного сигнала. При этом время и вероятность безошибочного преобразования входного сигнала агрегатом второго уровня a_{2l} при его прохождении через функциональные блоки a_{1k} (для случая независимых преобразований) определяются выражениями

$$t_{21}^{(a_{2l})} = \sum_{a_{1k} \in a_{2l}} t_{11}^{(a_{1k})}, \quad \beta_{22}^{(a_{2l})} = \prod_{a_{1k} \in a_{2l}} \beta_{12}^{(a_{1k})}.$$

Агрегаты третьего уровня (операторы). Для формального описания воспользуемся интерпретацией оператора одноканальной приоритетной системой массового обслуживания с отказами [9].

В случайные моменты времени t_j в систему поступают заявки, характеризующиеся случайным параметром α_j . Если система свободна, заявка немедленно принимается к обслуживанию. В противном случае она становится в очередь с допустимым временем ожидания $\gamma_j = s(\alpha_j, \beta)$, где β – параметр системы. Заявка покидает систему, если не будет принята к обслуживанию до момента $t_j + \gamma_j$. Далее к обслуживанию принимается другая заявка. Время занятости канала $\eta_j = \psi(\alpha_j, \beta)$.

Состояние агрегата описывается следующими фазовыми координатами: $z_1(t)$ – время, оставшееся до окончания обслуживания заявки, которая находится на обслуживании; $z_2(t)$ – число заявок в системе (в очереди на обслуживании).

Если $z_2(t) = 0$ (заявок нет), то $z_1(t) = 0$ для всех t от момента окончания обслуживания до момента поступления новой заявки. При $z_2(t) > 0$ заявки имеются на обслуживании в очереди. Вводятся новые координаты состояния: $z_{1+2k}(t) = \alpha'_k$ ($k = 1, 2, \dots, z_2(t) - 1$), где α'_k – параметр k -й заявки в очереди; $z_{2+2k}(t)$ – оставшееся время ожидания в очереди для k -й заявки. Заявки поступают в агрегат в момент t_j и имеют значения $x_j = \alpha_j$.

Получим операторы переходов и выходов агрегата.

Пусть в момент времени t_j поступает новая заявка. Если $z_2(t_j) > 0$, заявка поступает в очередь: $z_1(t)$ не меняется; $z_2(t)$ увеличивается на единицу; $z_{1+2k}(t)$ и $z_{2+2k}(t)$ не изменяются; возникают новые координаты $z_{1+2z_2}(t_j) = \alpha_j$ и $z_{2+2z_2}(t_j) = s(\alpha_j, \beta)$, характеризующие поступившую z_2 -ю заявку. Если в момент времени t_j канал свободен и заявок в очереди нет, то заявка принимается к обслуживанию, т. е. $z_1(t_j) = \psi(\alpha_j, \beta)$, $z_2(t_j) = 1$.

Таким образом, оператор перехода V' имеет вид

$$\left. \begin{aligned} z_1(t_j + 0) &= z_1(t_j), \\ z_2(t_j + 0) &= z_2(t_j) + 1, \\ z_{1+2k}(t_j + 0) &= z_{1+2k}(t_j) \\ z_{2+2k}(t_j + 0) &= z_{2+2k}(t_j) \end{aligned} \right\} k < z_2(t_j), \left. \begin{aligned} z_2(t_j) &> 0; \\ z_{1+2z_2}(t_j + 0) &= \alpha_j, \\ z_{2+2z_2}(t_j + 0) &= s(\alpha_j, \beta) \end{aligned} \right\}$$

$$\left. \begin{aligned} z_1(t_j + 0) &= \psi(\alpha_j, \beta), \\ z_2(t_j + 0) &= 1, \\ z_{1+2k}(t_j + 0) & \\ z_{2+2k}(t_j + 0) & \end{aligned} \right\} \left. \begin{aligned} & \text{не определяются} \\ z_2(t_j) &= 0. \end{aligned} \right\}$$

Опишем теперь множество состояний агрегата $Z^Y(g, \beta)$ и оператор выхода G . Множество $Z^Y(g, \beta)$ представим в виде объединения подмножеств Z_1^Y (определяется соотношением $z_1(t) = 0$) и Z_2^Y (определяется соотношением $z_{2+2k}(t) = 0$ хотя бы для одного k).

Под выходным сигналом будем понимать совокупность характеристик заявки, покидающей систему. Пусть $y = (y^{(1)}, y^{(2)})$, где признак $y^{(1)} = 1$, если систему покидает обслуженная заявка, или $y^{(1)} = 0$, если систему покидает необслуженная заявка, а $y^{(2)} = (\alpha_j, \beta, t^*)$ – совокупность сведений о заявке (заявка поступила в систему с характеристикой α_j , обслуживалась при значении параметра системы β , покинула систему в момент t^*). Действие оператора выхода G сводится к выбору признака $y^{(1)}$ и формированию сведений о заявке $y^{(2)}$.

Пусть обслуживание очередной заявки закончилось, т. е. в момент t_1^* состояние агрегата достигает подмножества Z_1^Y . В момент t_1^* , определяемый из соотношения $z_1(t_1^*) = 0$, агрегат формирует выходной сигнал $y = (1, \alpha_j, \beta, t_1^*)$. Пусть далее заявка покидает систему необслуженной, так как истекло время ее ожидания. В момент t_2^* , определяемый из соотношения $z_{2+2k}(t_2^*) = 0$, состояние агрегата достигает подмножества Z_2^Y , а величина $y = (0, \alpha_j, \beta, t_2^*)$.

Перейдем к оператору U . Для упрощения записи оператор удобно представить в виде совокупности операторов W' , W'' и \bar{U} .

Пусть $t = t_1^*$ (обслуживание очередной заявки окончено). Если в системе имеются заявки, к обслуживанию принимается следующая заявка в очереди. Если же в системе заявок нет, система ждет до момента поступления новой заявки и принимает ее к обслуживанию (этот случай рассмотрен выше при описании оператора V' при условии $z_2(t_j) = 0$). Определим состояния агрегата в полуинтервале от момента окончания обслуживания до момента поступления новой заявки.

Состояния агрегата $z(t_1^* + 0)$ определяются оператором W' :

$$\left. \begin{aligned} z_1(t_1^* + 0) &= \psi(z_3, \beta), \\ z_2(t_1^* + 0) &= z_2(t_1^*) - 1, \\ z_{1+2k}(t_1^* + 0) &= z_{1+2(k+1)}(t_1^*) \\ z_{2+2k}(t_1^* + 0) &= z_{2+2(k+1)}(t_1^*) \end{aligned} \right\} 1 \leq k < z_2(t_1^*) \left. \vphantom{\begin{aligned} z_1(t_1^* + 0) &= \psi(z_3, \beta), \\ z_2(t_1^* + 0) &= z_2(t_1^*) - 1, \\ z_{1+2k}(t_1^* + 0) &= z_{1+2(k+1)}(t_1^*) \\ z_{2+2k}(t_1^* + 0) &= z_{2+2(k+1)}(t_1^*) \end{aligned}} \right\} z_2(t_1^*) > 0;$$

$$\left. \begin{aligned} z_1(t_1^* + 0) &= z_1(t_1^*) = 0, \\ z_2(t_1^* + 0) &= 0, \\ z_{1+2z_2}(t_1^* + 0) & \\ z_{2+2z_2}(t_1^* + 0) & \end{aligned} \right\} \text{не определяются} \left. \vphantom{\begin{aligned} z_1(t_1^* + 0) &= z_1(t_1^*) = 0, \\ z_2(t_1^* + 0) &= 0, \\ z_{1+2z_2}(t_1^* + 0) & \\ z_{2+2z_2}(t_1^* + 0) & \end{aligned}} \right\} z_2(t_1^*) = 0.$$

В момент t_2^* , когда время ожидания l -й заявки истекает, число заявок уменьшается на единицу. Состояние агрегата $z(t_2^* + 0)$ определяется оператором W'' :

$$\begin{aligned} z_1(t_2^* + 0) &= z_2(t_2^*), \\ z_2(t_2^* + 0) &= z_2(t_2^*) - 1, \\ z_{1+2k}(t_2^* + 0) &= z_{1+2k}(t_2^*) \\ z_{2+2k}(t_2^* + 0) &= z_{2+2k}(t_2^*) \end{aligned} \left. \vphantom{\begin{aligned} z_1(t_2^* + 0) &= z_2(t_2^*), \\ z_2(t_2^* + 0) &= z_2(t_2^*) - 1, \\ z_{1+2k}(t_2^* + 0) &= z_{1+2k}(t_2^*) \\ z_{2+2k}(t_2^* + 0) &= z_{2+2k}(t_2^*) \end{aligned}} \right\} k < l,$$

$$\begin{aligned} z_{1+2k}(t_2^* + 0) &= z_{1+2(k+1)}(t_2^*) \\ z_{2+2k}(t_2^* + 0) &= z_{2+2(k+1)}(t_2^*) \end{aligned} \left. \vphantom{\begin{aligned} z_{1+2k}(t_2^* + 0) &= z_{1+2(k+1)}(t_2^*) \\ z_{2+2k}(t_2^* + 0) &= z_{2+2(k+1)}(t_2^*) \end{aligned}} \right\} l \leq k < z_2(t_2^*).$$

Оператор \bar{U} определяет состояния агрегата в полуинтервале между особыми событиями (t_n, t_{n+1}) , где t_n, t_{n+1} – соответственно моменты поступления в агрегаты входных сигналов и выдачи выходных:

$$\begin{aligned} z_1(t) &= z_1(t_n + 0) - (t - t_n), \\ z_2(t) &= z_2(t_n + 0), \\ z_{1+2k}(t) &= z_{1+2k}(t_n + 0), \\ z_{2+2k}(t) &= z_{2+2k}(t_n + 0) - (t - t_n). \end{aligned}$$

Абстрактное описание агрегата третьего уровня на этом можно считать законченным. Для конкретной реализации модели агрегата необходимо дать описание параметра α_j заявок и моментов t_j их поступления в систему, а также параметра системы β и функций

$\gamma_j = s(\alpha_j, \beta)$ и $\eta_j = \psi(\alpha_j, \beta)$, определяющих допустимое время ожидания заявки в очереди и время обслуживания заявки. Описание указанных параметров приведено в работе [9].

Агрегат четвертого уровня (коллектив операторов). Математическая модель агрегата четвертого уровня представляет собой многоканальную многофазную систему массового обслуживания. Агрегат A_4 полностью описывается множеством агрегатов третьего уровня и матрицей переходов, указывающей порядок передачи частных задач управления между операторами.

Обобщенная схема имитационной модели. Абстрактное описание модели деятельности коллектива операторов в виде четырехуровневой агрегативной системы позволяет разработать обобщенную структурную схему модели.

Существенной особенностью обобщенной модели групповой деятельности операторов является тот факт, что ее состав зависит от характера задач проектирования [10]. Выделяют модели трех уровней. Модель коллектива (первого уровня) используется при проектировании групповой деятельности; модель оператора (второго уровня) – при проектировании индивидуальной деятельности; модель оператора, учитывающая влияние эргономических факторов (третьего уровня), – при проектировании средств деятельности и задании требований к оператору.

Рассмотрим последовательно состав имитационных моделей разных уровней, выделяя при этом общие и специфические элементы модели (рис. 2).

Имитационная модель первого уровня включает следующие блоки (выделены толстыми линиями): ввода и хранения исходных данных по задачам, решаемым коллективом операторов (Б1); формирования простейшего потока задач (Б2); организации обслуживания задач операторами (Б3), включает частные модели решения задач операторами; обработки результатов моделирования (Б4); оценки адекватности модели (Б5); анализа и управления моделированием (Б6).

В Б1 вводятся и хранятся данные по задачам, решаемым коллективом операторов. Каждая задача определяется следующими данными: номером задачи $j = 1, \dots, n$; важностью (приоритетом) задачи C_j ; интенсивностью потока λ_j ; временем существования (допустимым временем решения) $t_{j\text{доп}}$, заданным законом распределения $B_j(t)$ с числовыми характеристиками $\bar{t}_{j\text{доп}}$, $\sigma_{j\text{доп}}$; временем решения j -й задачи каждым i -м оператором t_{ij} , заданным законом с числовыми характеристиками \bar{t}_{ij} , σ_{ij} .

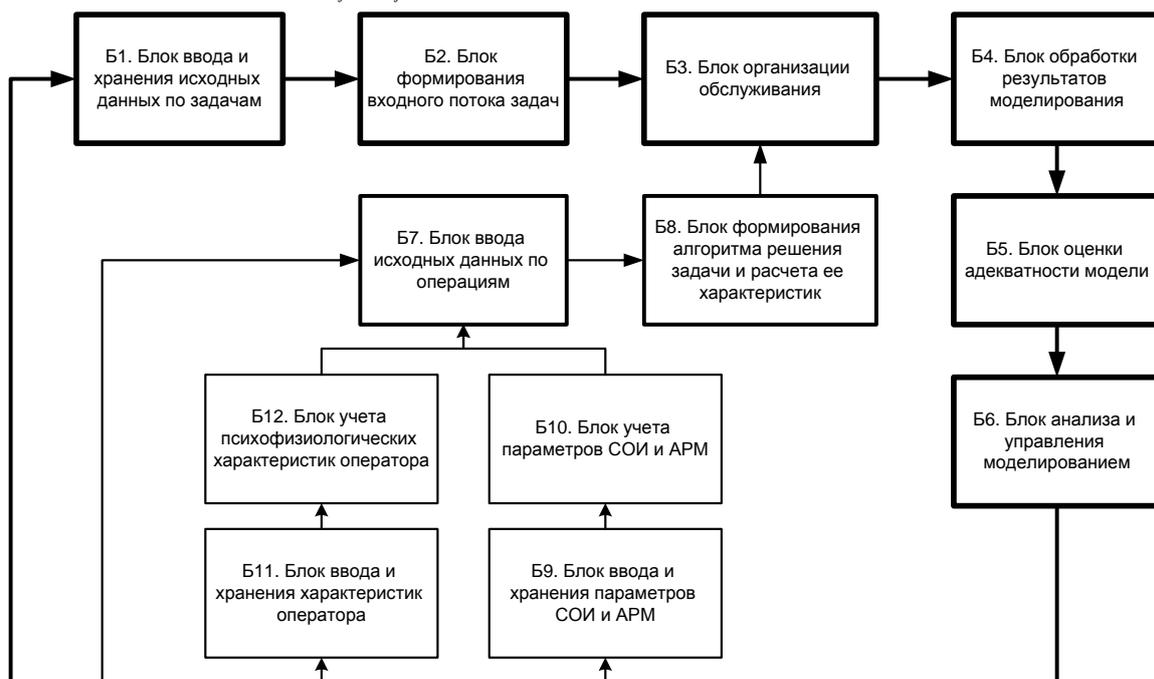


Рис. 2. Обобщенная структура имитационной модели групповой деятельности операторов

При моделировании групповой деятельности задается ориентированный граф последовательности решения каждой задачи коллективом операторов. Исходные данные по задаче удобнее представлять в виде матрицы смежности вершин графа $\|v_{ij}\|$, строками и столбцами которой являются номера операторов ($i=1, \dots, n$), а также начало (вход) и конец (выход) решения задачи. Пример графа решения задачи и соответствующая матрица $\|v_{ij}\|$ приведены на рис. 3.

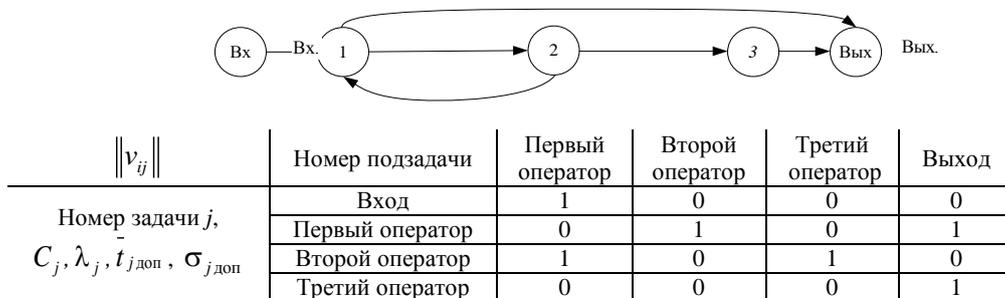


Рис. 3. Граф решения задачи коллективом операторов и его матрица смежности

В Б1 также хранятся сведения об алгоритме решения j -й задачи v_j , используемые моделями второго и третьего уровней. Алгоритм решения задачи v_j задается в виде вероятностного графа, определенного матрицей смежности вершин ω_{lk}^j , строками и столбцами которой являются операции алгоритма, а элементы P_{lk}^j характеризуют вероятность перехода после выполнения l -й операции к k -й при решении j -й задачи.

Пример графа алгоритма решения задачи и соответствующая матрица смежности вершин показаны на рис. 4. Кроме того, для каждой задачи имеется таблица идентификации номера операции с типом, содержащая также числовые характеристики выполнения операции.

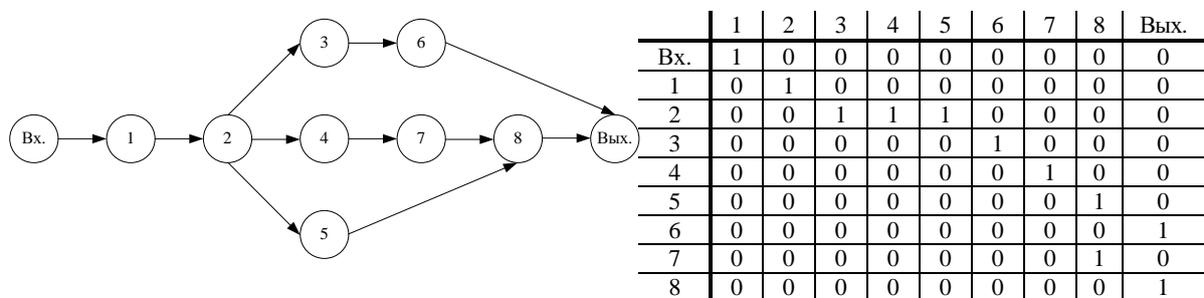


Рис. 4. Граф алгоритма решения задачи оператором и его матрица смежности

В Б2 формируется пуассоновский (простейший) поток входных задач, в котором вероятность поступления k задач в промежутке времени в простейшем случае задается формулой Пуассона:

$$P_k(\Delta t) = \frac{(\lambda \Delta t)^k}{k!} e^{-\lambda \Delta t}.$$

В простейшем потоке промежутки времени между поступлениями задач распределяются по показательному закону, поэтому процедура моделирования простейшего потока заключается в следующем. Весь интервал моделирования разбивается на определенное число шагов, в пределах которого считается, что поток задач распределен по закону Пуассона. Время поступления последующей задачи рассчитывается как сумма времени поступления предыдущей задачи и отрезка времени, распределенного по показательному закону:

$$t_i = t_{i-1} + \xi_i,$$

где ξ_i – случайный промежуток времени, распределенный по показательному закону.

Блок организации взаимодействия операторов и решения задач Б3 специфический и включается только в модель первого уровня. Этот блок обеспечивает включение моделей операторской деятельности в соответствии с графом решения задачи.

Моделируя групповую деятельность, необходимо учитывать последовательное или параллельное соединение операторов. При последовательном соединении включение моделей индивидуальной деятельности производится однозначно в соответствии с графом решения задачи. Параллельное соединение может быть задано разнообразной логикой включения моделей: случайно, по мере освобождения операторов; по приоритету; по минимуму длины очереди или времени ожидания и т. д. Могут быть учтены и характеристики взаимодействия операторов (время взаимодействия, вероятность безошибочного взаимодействия).

В Б3 включаются частные модели, моделирующие процесс решения задач операторами. Они представляют собой, по существу, приоритетные однолинейные системы массового обслуживания с ограниченным временем ожидания в очереди (пребывания в системе). Более подробно процесс организации обслуживания задач будет описан при рассмотрении модели второго уровня.

В блоке обработки результатов моделирования Б4 рассчитываются показатели качества решения задач и загрузки операторов. Так, при проведении каждого r -го ($r = \overline{1, z}$) эксперимента на модели вычисляются:

- вероятность своевременного решения j -й задачи в r -м эксперименте

$$P_{jr}^{св} = \frac{k_{jr}}{n_{jr}},$$

где k_{jr}, n_{jr} – соответственно количество возникших и своевременно решенных задач j -го типа в r -м эксперименте;

- вероятность своевременного и безошибочного решения j -й задачи в r -м эксперименте

$$P_{jr} = \frac{m_{jr}}{n_{jr}},$$

где m_{jr} – количество своевременно и безошибочно решенных задач j -го типа в r -м эксперименте;

- загрузка оператора решением задач j -го типа в r -м эксперименте

$$\rho_{jr} = \frac{\gamma_{jr}}{\Delta t},$$

где $\gamma_{jr} = \sum_j t_{jr}$ – время решения оператором всех задач j -го типа в r -м эксперименте.

Кроме того, Б4 вычисляет средние значения параметров:

- вероятности своевременного решения задач j -го типа

$$P_j^{св} = \frac{1}{z} \sum_{r=1}^z P_{jr}^{св};$$

- вероятности своевременного и безошибочного решения задач j -го типа

$$P_j = \frac{1}{z} \sum_{r=1}^z P_{jr};$$

– загрузки оператора решением задач j -го типа

$$\rho_j = \frac{1}{z} \sum_{r=1}^z \rho_{jr} ;$$

– вероятности своевременного и безошибочного решения всех задач оператором

$$P = \frac{1}{n} \sum_{j=1}^n P_j ;$$

– загрузки оператора

$$\rho = \sum_{j=1}^n \rho_j .$$

Б5 предназначен для оценки соответствия результатов моделирования деятельности оператора результатам реальной деятельности. Оценка адекватности модели производится по таким выходным показателям, как качество решения задач, качество работы операторов и параметры загрузки операторов. Порядок оценки адекватности модели и основные расчетные соотношения приведены в [11].

Б6 обеспечивает расчет и построение зависимостей входных показателей модели от характеристик задач, алгоритмов их решения, параметров средств отображения информации (СОИ) и автоматизированных рабочих мест (АРМ), характеристик оператора. Другой важной функцией блока является поиск оптимальных значений указанных факторов по полученным результатам моделирования, т. е. управление моделированием.

Для определения оптимальных значений параметров информационной модели целесообразно использовать параллельный метод стохастической оптимизации [12, 13]. Выбор данного метода определяется прежде всего его максимальным быстродействием, связанным с тем, что в качестве среднего направления шага оптимизации используется не направление статистического градиента (оно может изменяться в широких пределах), а непосредственное направление на экстремум в среднем по области определения.

Имитационная модель второго уровня отличается от модели первого уровня наличием блоков ввода и хранения исходных данных по операциям (Б7), формирования алгоритма решения задачи и расчета его характеристик (Б8) (на рис. 2 выделены контуром средней толщины), а также организации обслуживания с детализацией дисциплины обслуживания заявок оператором (Б3).

В Б7 вводятся и хранятся данные по типовым операциям (функциональным блокам), полученные либо из банка эргономических данных, либо экспериментальным путем.

По каждой операции имеются следующие данные:

- тип операции (идентификатор);
- временные характеристики выполнения операции, заданные законом распределения $\beta(\tau)$ с математическим ожиданием m_τ и среднеквадратическим отклонением σ_τ ;
- точностные характеристики выполнения операции, заданные законом распределения ошибки $\varphi(x)$ с математическим ожиданием m_x и среднеквадратическим отклонением σ_x ;
- вероятность безошибочного выполнения операции $P_{оп}$.

Б8 формирует вероятностный алгоритм решения задачи, заданный матрицей смежности операций, и рассчитывает время и вероятность правильного и своевременного решения задачи, используя данные блока Б7 по операциям. Основная идея формирования вероятностного алгоритма заключается в следующем.

В первой строке матрицы, соответствующей входу в алгоритм, находится ненулевой элемент: $P_{вхi} \neq 0$. Здесь возможны два случая: $P_{вхi} = 1$ для данного i и $0 < P_{вхi} < 1$ для нескольких i , при этом $\sum_i P_{вхi} = 1$.

При $P_{\text{вх}i} = 1$ по номеру операции i с помощью таблицы идентификации, хранящейся в Б1, находится тип операции. По типу операции реализуется обращение к Б7 и определяются время и правильность выполнения операции. Эти данные фиксируются в массиве задачи.

При $0 < P_{\text{вх}i} < 1$ для выбора очередной операции алгоритма используется датчик случайных чисел, параметрами которого являются величины $P_{\text{вх}i}$. Аналогично по номеру i анализируется следующая i -я строка матрицы смежности. Признаком завершения алгоритма является переход к строке, соответствующей операции из алгоритма решения задачи.

Блок Б3 предназначен для организации решения задачи оператором. Моделирующий алгоритм реализует как беспriorитетное обслуживание заявок, так и различные виды приоритетного обслуживания: относительный, абсолютный и динамический приоритеты.

При беспriorитетном обслуживании момент начала обслуживания очередной заявки соответствует моменту ее поступления, если система не занята обслуживанием или этот момент задерживается на время ожидания $\Delta t_{\text{ож}} = t_{\text{осв}} - t$, пока система освободится от обслуживания предшествующих заявок.

При относительном приоритете момент начала обслуживания поступившей заявки совпадает с моментом освобождения системы после обслуживания заявок более высокого приоритета. Если заявки обслуживаются по абсолютному приоритету, то момент начала обслуживания всегда соответствует моменту поступления заявки высшего приоритета.

Модель позволяет реализовать динамические приоритеты, особенностью которых является изменение порядка обслуживания в зависимости от условий (загрузки оператора).

Процесс обслуживания моделируется следующим образом. Считается, что все n_{jr} задач j -го типа в r -м моделировании обслужены, т. е. содержание счетчика K_{jr} количества одновременно обслуженных задач равно n_{jr} . После этого путем анализа времени поступления задачи, времени обслуживания и допустимого времени ожидания (или пребывания) устанавливается факт обслуживания задачи или ее потери (частичного обслуживания). Факт несвоевременного обслуживания регистрируется как вычитание единицы со счетчика K_{jr} . В результате содержимое счетчика K_{jr} будет соответствовать числу своевременно обслуженных задач j -го типа в r -м моделировании.

Для определения величины m_{jr} своевременно и правильно обслуженных задач содержимое счетчика K_{jr} обрабатывается таким образом, что всякий раз из K_{jr} вычитается единица, если в блоке Б3 зафиксирован факт неправильного выполнения алгоритма.

Существенным отличием *имитационной модели третьего уровня* (все блоки на рис. 2) от модели второго уровня является то, что характеристики операций рассчитываются и вводятся в блок Б7 по данным, характеризующим параметры СОИ и АРМ и психофизиологические особенности оператора. В зависимости от быстродействия и объема памяти ЭВМ, на которой реализуется модель, эти расчеты могут производиться либо перед моделированием деятельности оператора, либо непосредственно в процессе моделирования.

Указанные расчеты осуществляются с помощью блоков:

- ввода и хранения параметров СОИ в АРМ (Б9);
- учета параметров СОИ и АРМ (Б10);
- ввода и хранения характеристик оператора (Б11);
- учета психофизиологических характеристик оператора (Б12).

В Б9 вводятся и хранятся значения информационных, системных и светотехнических параметров СОИ, значения параметров пульта ввода АРМ (количество органов управления d , неупорядоченность их расположения L).

В Б11 вводятся и хранятся данные, позволяющие учесть:

– уровень подготовки оператора (время обучения; коэффициенты кривых обучения оператора; временные, вероятностные и точностные характеристики работы оператора в начале и конце обучения);

– процессы адаптации и утомления оператора во время работы (период адаптации оператора, время начала утомляемости, коэффициенты кривых адаптации и утомления);

– влияние напряженности (загрузки оператора) на качество деятельности.

Б10 обеспечивает расчет временных и вероятностных характеристик типовых операций в зависимости от параметров СОИ и пульта управления АРМ.

В Б10 на основе обработки статистического материала формируется функциональная зависимость вида

$$y = F(x_1, \dots, x_n),$$

где y – количественные характеристики выполнения типовой операции ($m_\tau, \sigma_\tau, m_x, \sigma_x, P_{\text{оп}}$), x_1, \dots, x_n – параметры системы отображения.

Для получения указанной зависимости могут использоваться корреляционно-регрессионный анализ или метод многорядовой селекции с применением комбинаторного перебора (метод группового учета аргументов, МГУА). Более предпочтительным является МГУА, так как он обладает рядом достоинств [13]:

– позволяет получать зависимость на основе минимального объема априорной информации;

– позволяет получать модели практически любой сложности;

– имеет высокую точность;

– структура модели в этом методе, в отличие от регрессионного анализа, заранее не фиксируется, а выбирается из множества вариантов по заданным критериям.

Общая идея МГУА заключается в том, что на основе статистических исходных данных, введенных в ЭВМ, и критерия выбора вычисляется единственная модель оптимальной сложности при помощи перебора большого числа моделей.

В частном случае, когда пульт управления характеризуется неупорядоченностью L и количеством органов управления d [10], имеет место выражение

$$m_\tau = 0,5d + 1,7(1,5 \frac{d-1}{25} - 1) \left[1 + \frac{4}{\pi} \arctg(2L-1) \right].$$

Б12 обеспечивает учет влияния особенностей деятельности оператора и его характеристик (уровня обученности, адаптации, загрузки и др.) на качество выполнения операций.

Уровень обученности оператора может быть учтен в модели с помощью либо таблиц, либо формульной зависимости. В таблице указываются значения величин $m_\tau, \sigma_\tau, m_x, \sigma_x, P_{\text{оп}}$ для разных (обычно трех) уровней обученности: удовлетворительного, хорошего и отличного.

При использовании показательной модели обучаемости зависимость величины m_τ от уровня имеет вид [10]

$$m_\tau = m_{\tau k} - (m_{\tau k} - m_{\tau 0}) e^{-\frac{t_1}{t_0}},$$

где $m_{\tau 0}, m_{\tau k}$ – время выполнения операции неподготовленным и отлично подготовленным операторами соответственно; t_0 – коэффициент, характеризующий скорость обучения; t_1 – время обучения.

Адаптация и утомление оператора могут привести к возрастанию времени выполнения операций, ухудшению точности и уменьшению вероятности безошибочной работы. Как показано в [8], эти факторы могут быть учтены с помощью кривых, приведенных на рис. 5.

Формульная зависимость среднего времени выполнения операции может быть представлена, например, следующим образом:

$$m_{\tau} = \begin{cases} \tau_0 + k_1 \operatorname{tg} \alpha (t_a - t) & \text{при } 0 \leq t \leq t_a, \\ \tau_0 & \text{при } t_a \leq t \leq t_{\text{yr}}, \\ \tau_0 + k_2 \operatorname{tg} \beta (t - t_{\text{yr}}) & \text{при } t \geq t_{\text{yr}}, \end{cases}$$

где τ_0 – среднее время выполнения операции без учета адаптации и утомляемости; t_a – период адаптации; t_{yr} – время начала утомляемости; t – текущее время работы; k_1 , k_2 , α , β – время обучения.

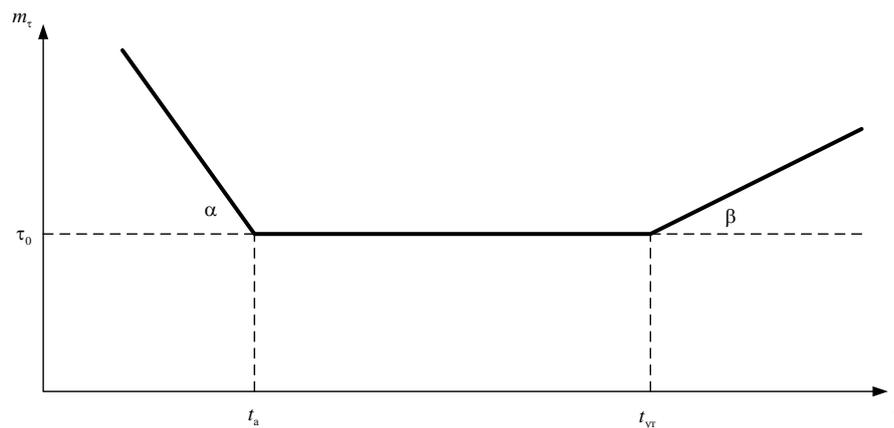


Рис. 5. Зависимость времени выполнения операций с учетом адаптации и утомляемости оператора

Напряженность в работе оператора связана с возникновением дефицита времени при обработке большого количества информации. Главным образом напряженность в работе отрицательно сказывается на безошибочности и точности работы оператора. Экспериментально полученная зависимость вероятности безошибочного выполнения операции от уровня загрузки оператора имеет вид

$$P_{\text{оп}} = P_0 - (P_0 - P_{\rho}) \left(1 - e^{-\frac{\rho - P_0}{C_0}}\right), \quad \rho \geq \rho_0,$$

где P_0 , P_{ρ} – значения вероятности безошибочного выполнения операции при нормальной загрузке и произвольной загрузке ($\rho \geq \rho_0$) соответственно; C_0 – коэффициент, характеризующий способность оператора работать в напряженных условиях.

В блоке 12 имеется возможность учесть и такой важный фактор, как «свертывание» алгоритма решения задачи оператором при дефиците времени. При «свертывании» алгоритма уменьшается время решения задачи, но возрастает вероятность ошибки [10]. Остальные блоки модели третьего уровня функционируют аналогично моделям первого и второго уровней.

На этом разработку обобщенной структурной схемы имитационной модели групповой и индивидуальной деятельности операторов в зависимости от характера решаемых задач можно считать завершенной.

Необходимо отметить, что разделение модели на три уровня является довольно условным, так как в зависимости от требуемой степени детализации при проектировании групповой деятельности может возникнуть необходимость учета деталей организации деятельности отдельного оператора. В свою очередь, организация индивидуальной деятельности оператора может сильно зависеть от особенностей организации групповой деятельности и результатов труда других операторов.

Заключение. На основе агрегативного подхода разработана формализованная модель деятельности коллектива операторов в виде четырехуровневой агрегативной системы, включающей функциональные блоки (агрегаты первого уровня), модели решения задач (агрегаты вто-

рого уровня), модели операторов (агрегаты третьего уровня), модель коллектива операторов (агрегат четвертого уровня). Агрегаты первого уровня представляют собой вероятностные автоматы, реализующие задержку и преобразование входных сигналов, второго уровня – вероятностный граф, третьего уровня – одноканальную систему массового обслуживания с отказами, агрегат четвертого уровня – многоканальную систему массового обслуживания.

В зависимости от характера задач проектирования деятельности предлагается использовать модели первого уровня (модель групповой деятельности коллектива операторов), второго (модель оператора) и третьего (модель оператора, учитывающую влияние эргономических факторов). Разработаны обобщенные структурные схемы моделей всех уровней. Определены основные требования и принципы функционирования блоков, входящих в модели соответствующего уровня.

Список использованных источников

1. Бекиш, А. Р. Подход к моделированию индивидуальной и групповой операторской деятельности с использованием CASE-технологий / А. Р. Бекиш, С. В. Кругликов // Сб. науч. ст. Воен. акад. Респ. Беларусь. – 2012. – № 22. – С. 30–37.
2. Бекиш, А. Р. Вариант построения имитационной модели групповой деятельности операторов автоматизированных систем управления тактического звена / А. Р. Бекиш, С. В. Кругликов // Вестн. Воен. акад. Респ. Беларусь. – 2015. – № 2(47). – С. 79–87.
3. Бекиш, А. Р. Научно-методический аппарат проектирования и оценки организации боевой работы в системах управления военного назначения / А. Р. Бекиш, С. В. Кругликов // Вестн. Воен. акад. Респ. Беларусь. – 2015. – № 3(48). – С. 101–111.
4. Кругликов, С. В. Адаптивные информационно-управляющие системы специального назначения: теория и практика синтеза : моногр. / С. В. Кругликов. – Минск : ВА РБ, 2014. – 233 с.
5. Санников, С. П. Моделирование систем / С. П. Санников. – Екатеринбург : Изд-во УГЛТУ, 2012. – 35 с.
6. Бусленко, Н. П. Лекции по теории сложных систем / Н. П. Бусленко, В. В. Калашников, И. Н. Коваленко. – М. : Сов. радио, 1973. – 440 с.
7. Бусленко, Н. П. Моделирование сложных систем / Н. П. Бусленко. – М. : Наука, 1978. – 400 с.
8. Суходольский, Г. В. Структурно-алгоритмический анализ и синтез деятельности / Г. В. Суходольский. – Л. : Изд-во ЛГУ, 1976. – 120 с.
9. Ронжин, О. В. Информационные методы исследования эргатических систем / О. В. Ронжин. – М. : Энергия, 1976. – 208 с.
10. Зинченко, В. П. Введение в эргономику / В. П. Зинченко. – М. : Сов. радио, 1974. – 349 с.
11. Бекиш, А. Р. Методика оценки адекватности имитационной модели групповой деятельности операторов / А. Р. Бекиш // Сб. науч. ст. Воен. акад. Респ. Беларусь. – 2016. – № 31. – С. 68–78.
12. Егоров, Б. М. Параллельный метод стохастической оптимизации многопараметрических объектов / Б. М. Егоров [и др.] // Вычисл. и прикл. матем. – Киев : Изд-во КГУ, 1979. – 191 с.
13. Ивахненко, А. Г. Принятие решения на основе самоорганизации / А. Г. Ивахненко, Ю. П. Зайченко, В. Д. Дмитров. – М. : Сов. радио, 1976. – 280 с.

References

1. Bekish A. R., Kruglikov S. V. Podkhod k modelirovaniyu individual'noy i gruppovoy operatorskoy deyatelnosti s ispol'zovaniem CASE-tehnologiy [The approach to modeling of individual and group operator activity with the use of CASE-technologies]. *Sbornik nauchnykh statey Voennoy akademii Respubliki Belarus' [A Collection of Scientific Articles of the Military Academy of the Republic of Belarus]*, 2012, no. 22, pp. 30–37 (in Russian).
2. Bekish A. R., Kruglikov S. V. Variant postroyeniya imitatsionnoy modeli gruppovoy deyatelnosti operatorov avtomatizirovannykh sistem upravleniya takticheskogo zvena [A variant of constructing an imitation model of group activity of operators of automated tactical control systems]. *Vestnik Voennoy akademii Respubliki Belarus' [Vestnik of the Military Academy of the Republic of Belarus]*, 2015, no. 2(47), pp. 79–87 (in Russian).
3. Bekish A. R., Kruglikov S. V. Nauchno-metodicheskiy apparat proyektirovaniya i otsenki organizatsii boyevoy raboty v sistemakh upravleniya voennogo naznacheniya [Scientific and methodical apparatus for designing and evaluating the organization of combat operations in military control systems]. *Vestnik Voennoy akademii Respubliki Belarus' [Vestnik of the Military Academy of the Republic of Belarus]*, 2015, no. 3(48), pp. 101–111 (in Russian).
4. Kruglikov S. V. Adaptivnyye informatsionno-upravlyayushchiye sistemy spetsial'nogo naznacheniya: teoriya i praktika sinteza. *Adaptive Information-Control Systems for Special Purposes: Theory and Practice of Synthesis*. Minsk, VA RB Publ., 2014, 233 p. (in Russian).
5. Sannikov, S. P. Modelirovaniye sistem. *Modeling of Systems*. Ekaterinburg, UGLTU Publ., 2012, 35 p. (in Russian).
6. Buslenko N. P., Kalashnikov V. V., Kovalenko I. N. Leksii po teorii slozhnykh sistem. *Lectures on the Theory of Complex Systems*. Moscow, Sov. radio Publ., 1973, 440 p. (in Russian).

7. Buslenko N. P. Modelirovaniye slozhnykh sistem. *Simulation of Complex Systems*. Moscow, Nauka Publ., 1978, 400 p.
8. Sukhodol'skiy G. V. Strukturno-algoritmicheskiy analiz i sintez deyatelnosti. *Structural-Algorithmic Analysis and Synthesis of Activity*. Leningrad, LGU Publ., 1976, 120 p. (in Russian).
9. Ronzhin O. V. Informatsionnyye metody issledovaniya ergaticheskikh sistem. *Information Methods for Studying Ergatic Systems*. Moscow, Energiya Publ., 1976, 208 p. (in Russian).
10. Zinchenko V. P. Vvedeniye v ergonomiku. *Introduction to Ergonomics*. Moscow, Sov. radio Publ., 1974, 349 p. (in Russian).
11. Bekish A. R. Metodika otsenki adekvatnosti imitatsionnoy modeli gruppovoy deyatelnosti operatorov [The method of assessing the adequacy of the simulation model of the operators group activity]. *Sbornik nauchnykh statey Voennoy akademii Respubliki Belarus'* [A Collection of Scientific Articles of the Military Academy of the Republic of Belarus], 2016, no. 31, pp. 68–78 (in Russian).
12. Yegorov, B. M. Parallel'nyy metod stokhasticheskoy optimizatsii mnogoparametricheskikh ob'yektov [Parallel method of stochastic optimization of multiparameter objects]. *Vychislitel'naya i prikladnaya matematika [Computational and Applied Mathematics]*, Kiev, Izd-vo KGU, 1979, 191 p. (in Russian).
13. Ivakhnenko A. G., Zaychenko Yu. P., Dmitrov V. D. Prinyatiye resheniya na osnove samoorganizatsii. *Decision Making on the Basis of Self-Organization*. Moscow, Sov. radio Publ., 1976, 280 p. (in Russian).

Информация об авторах

Кругликов Сергей Владимирович – доктор военных наук, кандидат технических наук, доцент, заместитель генерального директора по научной и инновационной работе, Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь).

E-mail: kruglikov_s@newman.bas-net.by

Бекиш Александр Романович – научный сотрудник научно-исследовательской части, Военная академия Республики Беларусь (пр. Независимости, 220, 220057, Минск, Республика Беларусь). E-mail: beker1@rambler.ru

Information about the authors

Sergey V. Kruglikov – Doctor of Military Sciences, Ph. D. (Engineering), Associate Professor, Deputy Director General for Scientific and Innovative Work, The United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Sursanova Str., 220012, Minsk, Republic of Belarus). E-mail: kruglikov_s@newman.bas-net.by

Alexander R. Bekish – Research Fellow of the Research Section, The Military Academy of the Republic of Belarus (220, Independence Ave., 220057, Minsk, Republic of Belarus). E-mail: beker1@rambler.ru

ISSN 1816-0301 (print)

УДК 004.9, 004.94, 004.56

Поступила в редакцию 22.01.2018

Received 22.01.2018

Н. И. Червяков¹, А. А. Коляда², Н. А. Коляда², В. А. Кучуков¹, С. Ю. Протасеня²¹Северо-Кавказский федеральный университет, Ставрополь, Россия²Научно-исследовательское учреждение «Институт прикладных физических проблем имени А. Н. Севченко» Белорусского государственного университета, Минск, Беларусь**НЕЙРОННЫЕ СЕТИ КОНЕЧНОГО КОЛЬЦА НА ОСНОВЕ РЕДУКЦИОННОЙ СХЕМЫ ПОЗИЦИОННО-МОДУЛЯРНОГО КОДОВОГО ПРЕОБРАЗОВАНИЯ**

Аннотация. Рассматривается проблема создания нейросетевых модулярных вычислительных структур для высокопроизводительных выражений в области защиты информации. Главное внимание уделяется редукционной технологии позиционно-модулярного преобразования масштабируемых целых чисел, которая служит основой для построения так называемых нейронных сетей конечного кольца (НСКК). Для повышения скорости сходимости используемой редукционной схемы понижения разрядности элементов формируемой последовательности вычетов предложен эффективный табличный метод. Разработанный подход позволяет свести к теоретическому минимуму количество итераций редукционного процесса. Это достигается за счет применения гибкого адаптивного механизма проверки принадлежности поитерационных вычетов к специальному диапазону, допускающему табличную декомпозицию его элементов на пары остатков по модулям модулярной системы счисления. На базе модифицированного редукционного метода синтезированы быстрый алгоритм и параллельная структура НСКК с обратной связью, обеспечивающая реализацию редукционной схемы за время $(S(\lceil \log_2 b \rceil + 1) + 2)t_{\text{сл}}$, где S – число итераций, b – разрядность входного числа, $t_{\text{сл}}$ – длительность операции сложения двух вычетов.

Ключевые слова: нейронная сеть, нейронная сеть конечного кольца, нейронная сеть с обратной связью, синоптические веса, модулярная система счисления, модулярная арифметика, редукционная схема понижения разрядности чисел, табличный метод

Для цитирования. Нейронные сети конечного кольца на основе редукционной схемы позиционно-модулярного кодового преобразования / Н. И. Червяков [и др.] // Информатика. – 2018. – Т. 15, № 2. – С. 98–110.

N. I. Chervyakov¹, A. A. Kolyada², N. A. Kolyada², V. A. Kuchukov¹, S. U. Protasenia²¹North-Caucasus Federal University, Stavropol, Russia²Scientific Research Institution "Institute of Applied Physical Problems named after A. N. Sevchenko" of the Belarusian State University, Minsk, Belarus**NEURAL NETWORKS OF THE FINAL RING BASED ON THE REDUCTION SCHEME OF THE POSITION-MODULAR-CODE TRANSFORMATION**

Abstract. The article studies the problem of creating a neural network of modular computing structures for high-performance expressions in the field of information security. The main attention is paid to the reduction technology of position-modular transformation of scalable integers, which serves as the basis for constructing the so-called neural networks of the finite ring (NNFR). To increase the speed of convergence of the reduction scheme used to reduce the number of elements of the generated sequence of residues, an effective tabular method is proposed. The developed approach makes it possible to reduce the number of iterations of the reduction process to a theoretical minimum. This is achieved through flexible adaptive mechanism check botheration deductions to a special range, allowing a tabular decomposition of its elements into pairs of residues in modules of the modular number system. On the basis of a modified reduction method there was synthesized a fast algorithm and a parallel structure of the NNFR with feedback, which ensures the implementation of the reduction scheme in a time order $(S(\lceil \log_2 b \rceil + 1) + 2)t_{\text{sum}}$, were S – the number of iterations, b – the bit width of the input number, t_{sum} – the duration of the addition operation of two deductions.

Keywords: neural network, neural network end rings, a neural network with feedback, the synaptic weight, modular number system, modular arithmetic, reducing the scheme of reduction of bit numbers, the table method

For citation. Chervyakov N. I., Kolyada A. A., Kolyada N. A., Kuchukov V. A., Protasenia S. U. Neural networks of the final ring based on the reduction scheme of the position-modular-code transformation. *Informatics*, 2018, vol. 15, no. 2, pp. 98–110 (in Russian).

Введение. Как известно, важнейшей неотъемлемой составляющей математического и компьютерного обеспечения современных систем защиты информации являются вычислительные технологии на диапазонах больших чисел (ДБЧ). Особое место среди таких технологий занимают модулярные вычислительные технологии (МВТ) [1–6]. Обладая естественным кодовым параллелизмом, модулярные вычислительные структуры имеют ряд существенных преимуществ над позиционными структурами, и наиболее ярко эти преимущества проявляются при оперировании в ДБЧ. Важным фактором, способствующим неуклонному повышению уровня востребованности технологий требуемого класса на основе модулярной арифметики (МА), является их идеальная приспособленность к нейросетевым реализациям [1, 3, 7–11].

Активно развиваемое в настоящее время новое направление фундаментальных и прикладных исследований в криптографии, нацеленное на как можно более полную реализацию оптимально согласованных свойств параллелизма нейронных сетей (НС) и арифметики модулярных систем счисления (МСС), предоставляет принципиально новые возможности для построения высокопроизводительных криптосистем различного функционального назначения. Многообещающие перспективы совместного применения модулярной и нейросетевой вычислительных технологий обусловлены тем, что при согласованном числе синапсов НС, используемых в процессе взаимодействия ее нейронов, и мощности модулярного базиса сеть становится естественным представлением применяемой МСС. На отмеченное обстоятельство указывают, в частности, следующие признаки:

- семантическое сходство позиционных форм модулярных чисел [1, 10, 12, 13] с расчетными соотношениями формального нейрона [3, 10];
- существование адекватного отображения алгоритмов арифметических операций в МСС на многослойные НС;
- простота реализации основных операций нейросетевого логического базиса в модулярном коде;
- равнозначность модулярного кодирования информации ассоциативной нейронной памяти.

Основополагающая идея осуществляемых разработок по созданию методологического, алгоритмического и программно-аппаратного обеспечения нейросетевой МВТ состоит в переводе вычислений из ДБЧ в компьютерные диапазоны целых чисел (ЦЧ) стандартной разрядности. Ключевую роль в процессе решения сформулированной задачи выполняют НС на конечных кольцах вычетов по рабочему базису модулей [1, 3, 14, 15]. Операционную основу НСКК составляют главным образом операции приведения ЦЧ к остаткам по используемым модулям. Как структурно, так и на операционном уровне НСКК максимально должны быть согласованы с естественным кодовым параллелизмом МА. В полной мере данному условию удовлетворяет рассматриваемый в данной статье редукционный метод позиционно-модулярного кодового преобразования.

Редукционный метод позиционно-модулярного преобразования масштабируемых целых чисел. Введем обозначения:

Z – множество ЦЧ;

$\lfloor a \rfloor$ и $\lceil a \rceil$ – наибольшее и наименьшее ЦЧ соответственно, не большие и не меньшие естественной величины a ;

$Z_m = \{0, 1, \dots, m-1\}$, $Z_m^- = \{-\lfloor m/2 \rfloor, -\lfloor m/2 \rfloor + 1, \dots, \lceil m/2 \rceil - 1\}$ – множества наименьших неотрицательных и абсолютно наименьших вычетов по натуральному модулю m ;

$|a|_m$ и $|a|_m^-$ – элементы множеств Z_m и Z_m^- , сравнимые с a (в общем случае рациональным числом) по модулю m ;

$M = \{m_1, m_2, \dots, m_k\}$ – набор модулей базовой МСС (модулярный базис), где k – мощность базиса;

$X = (|X|_{m_1}, |X|_{m_2}, \dots, |X|_{m_k})$ – представление ЦЧ X в МСС с базисом M .

Традиционно в качестве операционной основы НСКК используется преобразование ЦЧ X из двоичной системы счисления в модулярную [3, 14, 15]. Принимая, однако, во внимание то обстоятельство, что при построении различных конфигураций МА важную роль выполняют преобразования масштабируемых вычетов в остатке по модулям МСС, в настоящей статье рас-

сма­три­ва­ет­ся рас­ши­рен­ный клас­с НСКК, ко­то­рые осу­ществ­ля­ют тре­буе­мые опе­ра­ции – пре­об­ра­зо­ва­ния ви­да $X \rightarrow |CX|_m$, где X – не­от­ри­ца­тель­ное ЦЧ, за­дан­ное сво­им дво­ич­ным ко­дом $(x_{b-1} x_{b-2} \dots x_0)_2$ раз­ряд­но­стью b бит ($x_j \in \{0,1\}$ ($j = \overline{0, b-1}$)); C – це­ло­чис­лен­ная кон­стан­та (мас­шта­би­ру­ю­щий мно­жи­тель); $m \in \mathbf{M}$. По кри­те­рию про­сто­ты ней­ро­се­те­вой ре­али­за­ции наи­бо­лее при­ем­ле­мым ме­то­дом вы­пол­не­ния пре­об­ра­зо­ва­ния $X \rightarrow |CX|_m$ яв­ля­ет­ся ме­то­д мо­ду­ляр­ной ре­дук­ции по ре­кур­сив­ной схе­ме по­сле­до­ва­тель­но­го сни­же­ния раз­ряд­но­сти эле­мен­тов фор­ми­ру­е­мой по­сле­до­ва­тель­но­сти вы­че­тов [1, 3, 15].

По­ло­жим

$$X^{(0)} = (x_{b_0-1}^{(0)} x_{b_0-2}^{(0)} \dots x_0^{(0)})_2 = \sum_{j=0}^{b_0-1} 2^j x_j^{(0)} \quad (b_0 = b, x_j^{(0)} = x_j) \quad (1)$$

и пусть

$$W_j(C) = |C \cdot 2^j|_{m_0}^- = \begin{cases} |C \cdot 2^j|_m, & \text{если } |C \cdot 2^j|_m < \left\lfloor \frac{m}{2} \right\rfloor, \\ |C \cdot 2^j|_m - m, & \text{если } |C \cdot 2^j|_m \geq \left\lfloor \frac{m}{2} \right\rfloor, \end{cases} \quad (2)$$

$$(j = \overline{0, b-1}).$$

При $C=1$ да­лее упо­тре­б­ля­ет­ся обо­зна­че­ние $W_j = W_j(1)$.

При­ме­ня­е­мая ре­дук­ци­он­ная схе­ма опи­сы­ва­ет­ся опе­ра­ци­он­ной по­сле­до­ва­тель­но­стью

$$\left\{ \begin{aligned} X^{(1)} &= \sum_{j=0}^{b_0-1} W_j(C) x_j^{(0)} = (x_{b_1-1}^{(1)} x_{b_1-2}^{(1)} \dots x_0^{(1)})_2 - 2^{b_1} x_{b_1-1}^{(1)} = \sum_{j=0}^{b_1-2} 2^j x_j^{(1)} - 2^{b_1-1} x_{b_1-1}^{(1)}, \\ X^{(s)} &= \sum_{j=0}^{b_{s-1}-2} W_j x_j^{(s-1)} - W_{b_{s-1}-1} x_{b_{s-1}-1}^{(s-1)} = (x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_0^{(s)})_2 - 2^{b_s} x_{b_s-1}^{(s)} = \\ &= \sum_{j=0}^{b_s-2} 2^j x_j^{(s)} - 2^{b_s-1} x_{b_s-1}^{(s)} \quad (s = \overline{2, S}); \end{aligned} \right\},$$

$$\chi = |X^{(S)}|_m$$

где b_1 и b_s – дли­ны до­пол­ни­тель­ных дво­ич­ных ко­дов $(x_{b_1-1}^{(1)} x_{b_1-2}^{(1)} \dots x_0^{(1)})_2$ и $(x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_0^{(s)})_2$ со­от­вет­ствен­но ЦЧ $X^{(1)}$ и $X^{(s)}$, ко­то­рые, как сле­ду­ет из (2), в прин­ци­пе мо­гут бы­ть и от­ри­ца­тель­ны­ми; S – ко­ли­че­ство и­те­ра­ций схе­мы.

Осно­во­по­ла­га­ю­щая идея ре­дук­ци­он­но­го ме­то­да, ре­али­зу­е­мая в рам­ках вы­чис­ли­тель­ной схе­мы (3) в це­лях при­ве­де­ния ЦЧ CX к остат­ку по мо­ду­лю m , со­сто­ит в за­ме­не коэф­фи­ци­ен­тов $C2^j$ пра­вой ча­сти ра­вен­ства $CX = CX^{(0)} = \sum_{j=0}^{b_0-1} (C2^j) x_j^{(0)}$ на вы­че­ты $W_j(C)$, опре­де­ля­е­мые по пра­ви­лу (2), а коэф­фи­ци­ен­тов 2^j , $2^{b_{s-1}-1}$ вы­ра­же­ния $X^{(s)} = \sum_{j=0}^{b_{s-1}-2} 2^j x_j^{(s-1)} - 2^{b_{s-1}-1} x_{b_{s-1}-1}^{(s-1)}$ – на вы­че­ты W_j , $W_{b_{s-1}-1}$ при $s = \overline{2, S}$. Вви­ду (2) все по­лу­ча­е­мые по­сле ука­зан­ных за­мен ЦЧ $X^{(s)}$ ($s = \overline{1, S}$) ра­вно­ос­та­точ­ны по мо­ду­лю m . Они яв­ля­ют­ся эле­мен­та­ми од­но­го и то­го же клас­са \overline{X} вы­че­тов по дан­но­му мо­ду­лю: $\mathbf{R}_m(CX) = \{R \in \mathbf{Z} | R \equiv CX \pmod{m}\}$.

Спра­вед­ли­во сле­ду­ю­щее ут­вер­жде­ние.

Те­о­ре­ма. Пусть мо­ду­ль m яв­ля­ет­ся про­стым чис­лом ($m > 2$) и име­ет раз­ряд­но­сть $b_{\text{mod}} = \lceil \log_2 m \rceil$ бит. Тогда для дли­ны b_s до­пол­ни­тель­но­го дво­ич­но­го ко­да $(x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_0^{(s)})_2$ вы-

чета $X^{(s)}$, определяемого по редуccionной схеме (3) для входного ЦЧ (1) с использованием синтаксических весов (2), верна оценка

$$b_s < b_{mod} + \log_2 b_{s-1}, \quad (4)$$

где b_{s-1} – разрядность ЦЧ $X^{(s-1)}$, $s = \overline{1, S}$.

При этом количество S итераций редуccionной схемы (3) удовлетворяет неравенству

$$\begin{aligned} S &\leq \min \{s \mid b_s - b_{mod} < \\ &< \log_2 (b_{mod} + \log_2 (b_{mod} + \log_2 (\dots \log_2 (b_{mod} + \log_2 b_0) \dots))) \} \leq \\ &\leq \Delta_{min}; s \geq 1 \} (b_0 = b), \end{aligned} \quad (5)$$

где Δ_{min} – установленный порог.

Доказательство. Как известно, множество всех степеней числа 2 по модулю m в мультипликативной группе кольца \mathbf{Z}_m образует так называемую циклическую подгруппу, порождаемую элементом 2. Порядок N этой подгруппы (число ее элементов) служит делителем функции Эйлера $\varphi(m) = m-1$. Сказанное относится и к совокупности абсолютно наименьших остатков от деления указанных степеней на m , т. е. к определяемому по формуле (2) набору вычетов:

$$\{W_j \in \mathbf{Z}_m^- \mid W_j = |2^j|_m^-; j = \overline{0, N-1}\}. \quad (6)$$

При $b_{mod} < b_{s-1}$ последовательности весовых коэффициентов

$$\{W_0, W_1, \dots, W_{N-1}, W_N, \dots, W_{b_{s-1}-1}\} \quad (s = \overline{2, S}), \quad (7)$$

используемые в (3), имеют циклическую структуру. Сегменты длины N в (7) с начальными элементами W_{iN} ($i = \overline{0, \lfloor b_{s-1}/N \rfloor - 1}$) совпадают с последовательностью (6). В случае, когда b_{s-1} не делится нацело на N , последний $\lfloor b_{s-1}/N \rfloor$ -й сегмент в (7) оказывается неполным. Пусть N_+ и N_- – количество соответственно положительных и отрицательных вычетов в множестве (6). Тогда с учетом вышесказанного при $s = \overline{2, S}$ максимально возможное значение числа $X^{(s)}$ (см. (3)) сверху можно оценить следующим образом:

$$\begin{aligned} \max\{X^{(s)}\} &< \frac{b_{s-1}}{N} \sum_{j=0}^{N_+-1} \left(\frac{m-1}{2} - 1 \right) = \frac{b_{s-1}}{N} \left(\frac{m-1}{2} + \frac{m-1}{2} - N_+ + 1 \right) \frac{N_+}{2} = \\ &= \frac{b_{s-1}}{N} (m - N_+) \frac{N_+}{2}. \end{aligned}$$

Аналогично для минимального значения ЦЧ $X^{(s)}$ верна оценка

$$\begin{aligned} \min\{X^{(s)}\} &> \frac{b_{s-1}}{N} \sum_{j=0}^{N_- - 1} \left(-\frac{m-1}{2} + j \right) = \\ &= \frac{b_{s-1}}{N} \left(-\frac{m-1}{2} - \frac{m-1}{2} + N_- - 1 \right) \frac{N_-}{2} = \\ &= -\frac{b_{s-1}}{N} (m - N_-) \frac{N_-}{2}. \end{aligned}$$

Следовательно,

$$\begin{aligned}
 \max\{X^{(s)}\} - \min\{X^{(s)}\} &< \frac{b_{s-1}}{N} \left(\frac{N_+}{2} (m - N_+) + \frac{N_-}{2} (m - N_-) \right) = \\
 &= \frac{b_{s-1}}{N} \left(\frac{m}{2} (N_+ + N_-) - \frac{N_+^2 + N_-^2}{2} \right) = \\
 &= \frac{b_{s-1}}{2N} (mN - (N_+^2 + N_-^2 + 2N_+N_- - 2N_+N_-)) = \\
 &= \frac{1}{2} b_{s-1} \left(m - \frac{(N_+ + N_-)^2}{N} + \frac{2N_+N_-}{N} \right) = \\
 &= \frac{1}{2} b_{s-1} \left(m - N + 2N \left(\left(\frac{N_+}{N} \cdot \frac{N_-}{N} \right)^{\frac{1}{2}} \right)^2 \right) \leq \\
 &\leq \frac{1}{2} b_{s-1} \left(m - N + 2N \left(\frac{N_+}{N} + \frac{N_-}{N} \right) \cdot \frac{1}{4} \right) = \\
 &= \frac{1}{2} b_{s-1} \left(m - N + \frac{1}{2} N \right) = \frac{1}{2} b_{s-1} \left(m - \frac{1}{2} N \right).
 \end{aligned}$$

Так как N является делителем функции Эйлера $\varphi(m) = m-1$, оно представимо в виде $N = (m-1)/d$, где d – делитель ЦЧ $\varphi(m) = m-1$ ($d \neq N$). С учетом отмеченного обстоятельства из оценки для $\max\{X^{(s)}\} - \min\{X^{(s)}\}$ получаем

$$\max\{X^{(s)}\} - \min\{X^{(s)}\} < \frac{1}{2} b_{s-1} \left(m - \frac{m-1}{2d} \right) = \frac{1}{2} b_{s-1} m \left(1 - \frac{1}{2d} + \frac{1}{2dm} \right).$$

Отсюда следует

$$\begin{aligned}
 &\log_2(\max\{X^{(s)}\} - \min\{X^{(s)}\} + 1) + 1 < \\
 &< \log_2(b_{s-1}) + b_{\text{mod}} + \log_2 \left(1 - \frac{1}{2d} + \frac{1}{2dm} \right).
 \end{aligned}$$

Таким образом, ввиду $\frac{1}{2} < 1 - \frac{1}{(2d)} + \frac{1}{(2dm)} < 1$ для разрядности ЦЧ $X^{(s)}$ справедлива оценка

$$b_s = \lceil \log_2(\max\{X^{(s)}\} - \min\{X^{(s)}\} + 1) \rceil < b_{\text{mod}} + \log_2(b_{s-1}).$$

Что касается числа $X^{(1)}$ (см. (3)), то для оценки его разрядности b_1 также применим рассмотренный выше подход. Таким образом, при любом целочисленном C последовательность $\{W_0(C), W_1(C), \dots, W_{N-1}(C), W_N(C), W_{N+1}(C), \dots, W_{2N-1}(C), \dots, W_{b_0-1}(C)\}$ абсолютно наименьших остатков по модулю m , определяемых по правилу (2), благодаря выполнению равенств $W_{iN+j}(C) = W_j(C)$ ($i = 1, \lfloor \frac{b_0}{N} \rfloor - 1; j = \overline{0, N-1}$) и последовательности (7) обладает циклической структурой (с периодом N). Математические выкладки, приведенные выше для $X^{(s)}$ ($s = \overline{2, S}$), имеют вид $b_1 < b_{\text{mod}} + \log_2 b_0 = b_{\text{mod}} + \log_2 b$.

Существование искомого числа S итераций редуцирующей схемы, удовлетворяющего условию (5), вытекает из оценки (4), записанной в развернутом виде (по рекуррентному правилу), и выбора порога Δ_{min} . ■

Замечание 1. Фигурирующий в (5) порог Δ_{min} подбирается экспериментально так, чтобы последовательность b_0, b_1, \dots, b_S была строго убывающей. Выполнение данного условия обеспечивается тем, что при S , удовлетворяющем (5), величины $\log_2 b_0, \log_2 b_1, \dots, \log_2 b_S$ образуют убывающую последовательность.

Замечание 2. В соответствии с (5) и замечанием 1 в качестве признака завершения редукционного процесса (3), естественно, следует принять выполнение неравенства

$$\Delta_S = b_S - b_{mod} \leq \Delta_{min}. \quad (8)$$

При использовании в (8) $\Delta_{min} = 0$ искомое значение выходной величины схемы (3) формируется по правилу

$$\chi = \begin{cases} X^{(S)} + m, & \text{если } X^{(S)} < 0, \\ X^{(S)}, & \text{если } 0 \leq X^{(S)} < m, \\ X^{(S)} - m, & \text{если } m \leq X^{(S)}. \end{cases} \quad (9)$$

Замечание 3. Расчетные соотношения схемы (3) целиком согласуются с принципами нейросетевой вычислительной технологии. При суммировании синаптических весов (2) с последующим вычислением активационной функции $\chi = |CX|_m = |X^{(S)}|_m$, реализуемой, например, в виде (9), набор необходимых весовых коэффициентов (2) рассчитывается предварительно и хранится в памяти.

Табличный метод ускоренной реализации редукционной схемы. С увеличением значения переменной s скорость приближения b_s к разрядности b_{mod} модуля m снижается (см. (4), (5)). Поэтому в целях уменьшения количества S итераций вычислительной схемы (3) до приемлемого уровня в (8) следует использовать выбираемый надлежащим образом порог $\Delta_{min} > 0$. При этом необходимо учитывать то обстоятельство, что с ростом Δ_{min} расчетное соотношение для активационной функции нейронной сети в сравнении с (9) становится более сложным. Выбор наиболее приемлемых конфигураций блока суммирования синаптических весов, активационной функции $f(F^{(s)})$ базового нейрона ($s = \overline{1, S}$), способа ее реализации и организации цикла по переменной x является ключевым аспектом задачи оптимизации объема производимых вычислений и соответствующих реализационных затрат для формирования последовательности вычетов $X^{(1)}, X^{(2)}, \dots, X^{(S)}$, а также генерирования и анализа по правилу типа (8) на каждой итерации признака окончания рекурсивного процесса (3). Представленный подход к построению НСКК предусматривает применение табличного метода ускоренного выполнения рекурсивной редукционной схемы (3). Предлагаемый метод обеспечивает уменьшение числа S итераций схемы и снижение сложности активационной функции до уровня сложности расчетного соотношения (9).

Разобьем b_s – битовый дополнительный двоичный код $(x_{b_{s-1}}^{(s)} x_{b_{s-2}}^{(s)} \dots x_0^{(s)})_2$ ЦЧ $X^{(s)}$ – на три части: младшую $(x_{b_{mod}-2}^{(s)} x_{b_{mod}-3}^{(s)} \dots x_0^{(s)})_2$, среднюю $(x_{b_{mod}+\Delta_{min}-1}^{(s)} x_{b_{mod}+\Delta_{min}-2}^{(s)} \dots x_{b_{mod}-1}^{(s)})_2$ и старшую $(x_{s-1}^{(s)} x_{s-2}^{(s)} \dots x_{b_{mod}+\Delta_{min}}^{(s)})_2$, которые имеют соответственно разрядности $b_{mod}-1$, $\Delta_{min} + 1$ и $\Delta' = \Delta - \Delta_{min}$ ($\Delta = b_s - b_{mod}$). Основополагающая идея табличного метода ускоренного выполнения редукционного процесса (3) состоит в увеличении порога Δ_{min} до максимально допустимого (с точки зрения размера используемой таблицы) значения, обеспечивающего минимизацию количества S итераций реализуемой схемы. На искомую заключительную итерацию редукционного процесса указывает нулевое значение управляющего сигнала

$$\Gamma = \left| \left(\bigvee_{j=b_s-\Delta'-1}^{b_s-1} x_j^{(s)} \right) + \left(\bigwedge_{j=b_s-\Delta'-1}^{b_s-1} x_j^{(s)} \right) \right|_2. \quad (10)$$

Булево выражение (10) фактически представляет собой реализацию проверки условия (8) из выражения (2). Значение $\Gamma=0$ сигнала (10) указывает на то, что все разряды старшей Δ' битовой части $(x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_{b_s-\Delta'-1}^{(s)})_2$ дополнительного двоичного кода $(x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_0^{(s)})_2$ вычета $X^{(s)}$ совпадают со старшим разрядом в средней части $(x_{b_{mod}+\Delta_{min}-1}^{(s)} x_{b_{mod}+\Delta_{min}-2}^{(s)} \dots x_{b_{mod}-1}^{(s)})_2$ данного кода:

$$x_j^{(s)} = x_{b_s-\Delta'}^{(s)} = x_{b_{mod}+\Delta_{min}-1}^{(s)} \quad (j = \overline{b_s - \Delta' - 1, b_s - 1}). \quad (11)$$

При выполнении условия (11) ЦЧ $X^{(s)}$ является элементом диапазона $\mathbb{Z}_{2^{b_{mod}+\Delta_{min}-1}}^- = \{-2^{b_{mod}+\Delta_{min}-1}, -2^{b_{mod}+\Delta_{min}-1} + 1, \dots, 2^{b_{mod}+\Delta_{min}-1} - 1\}$. При этом $X^{(s)}$ полностью определяется дополнительным двоичным кодом $(x_{b_{mod}+\Delta_{min}-1}^{(s)}, x_{b_{mod}+\Delta_{min}-2}^{(s)}, \dots, x_0^{(s)})_2$ и, следовательно,

$$X^{(s)} = \sum_{j=0}^{b_{mod}+\Delta_{min}-2} 2^j x_j^{(s)} - 2^{b_{mod}+\Delta_{min}-1} x_{b_{mod}+\Delta_{min}-1}^{(s)}. \quad (12)$$

По достижении итерации, на которой фиксируется $\Gamma=0$, текущее значение s присваивается переменной S . В этом случае из (12) для $X^{(S)}$ получаем соотношение

$$|X^{(S)}|_m = |X_0^{(S)} + X_1^{(S)}|_m, \quad (13)$$

где

$$X_0^{(S)} = \sum_{j=0}^{b_{mod}-2} 2^j x_j^{(S)} = \sum_{j=0}^{b_{mod}-2} W_j x_j^{(S)}; \quad (14)$$

$$\begin{aligned} X_1^{(S)} &= \left| \sum_{j=b_{mod}-1}^{b_{mod}+\Delta_{min}-2} 2^j x_j^{(S)} - 2^{b_{mod}+\Delta_{min}-1} x_{b_{mod}+\Delta_{min}-1}^{(S)} \right|_m = \\ &= \left| \sum_{j=b_{mod}-1}^{b_{mod}+\Delta_{min}-2} W_j x_j^{(S)} - W_{b_{mod}+\Delta_{min}-1} x_{b_{mod}+\Delta_{min}-1}^{(S)} \right|_m. \end{aligned} \quad (15)$$

Значения b_{mod} – битового вычета $X_1^{(S)}$ по модулю m – рассчитываются согласно (15) предварительно и записываются в табличную память (таблицу $TRes_MP$) по правилу

$$TRes_MP [(x_{b_{mod}+\Delta_{min}-1}^{(s)} x_{b_{mod}+\Delta_{min}-2}^{(s)} \dots x_{b_{mod}-1}^{(s)})_2] = X_1^{(s)}. \quad (16)$$

Емкость таблицы $TRes_MP$ составляет $2^{\Delta_{min}+1}$ слов с разрядностью b_{mod} бит.

В случае, когда управляющий сигнал Γ , генерируемый по формуле (10), принимает значение $\Gamma=1$, реализация редуцированной схемы (3) продолжается при очередном значении s .

Приведение ЦЧ $X^{(s)}$ к результирующему остатку по модулю m описанным табличным методом на основе (13)–(16) позволяет значительно расширить диапазон анализируемых вычетов при генерировании признака завершения редуцированного процесса. Именно это обстоятельство и обеспечивает существенное уменьшение числа S итераций вычислительной схемы (3).

Редуциционный алгоритм ускоренного позиционно-модулярного преобразования масштабируемых целых чисел. На базе представленного метода последовательного уменьшения разрядности ЦЧ по редуциционной схеме (3) рекурсивного типа синтезирован алгоритм

ускоренного позиционно-модулярного кодового преобразования, ориентированный на нейросетевую реализацию.

Параметры алгоритма:

– попарно простые модули m_1, m_2, \dots, m_k , имеющие соответственно разрядности $b_mod_1, b_mod_2, \dots, b_mod_k$ бит ($b_mod_i = \lceil \log_2 m_i \rceil$ ($i = \overline{1, k}$); $k \geq 1$);

– порог $\Delta_{min} > 0$ для решающего правила завершения редуccionного процесса.

Входные данные:

– двоичный код $(x_{b-1} x_{b-2} \dots x_0)_2$ исходного ЦЧ X (b – длина кода);

– целочисленные коэффициенты C_i произведений $C_i X$, подлежащих приведению к остаткам по модулям m_i ($i = \overline{1, k}$).

Выходные данные: модулярный код $(\chi_1, \chi_2, \dots, \chi_k)$ ($\chi_i = |C_i X|_{m_i}$ ($i = \overline{1, k}$)) по заданному базису модулей – $\mathbf{M} = \{m_1, m_2, \dots, m_k\}$.

Предварительно получаемые данные:

1. Рассчитанные согласно правилу (2) наборы весов

$$\mathbf{W}_i(C) = \{W_{j,i}(C) | W_{j,i}(C) = |C_i 2^j|_{m_i}^-; j = \overline{0, b-1}\}; \quad (17)$$

$$\mathbf{W}_i = \{W_{j,i} | W_{j,i} = |2^j|_{m_i}^-; j = \overline{0, b_1-1}; b_1 < b\} (i = \overline{1, k}). \quad (18)$$

2. Таблицы $TRes_MP_1, TRes_MP_2, \dots, TRes_MP_k$ остатков от ЦЧ вида

$$A_i(a_0, a_1, \dots, a_{\Delta_{min}}) = \sum_{j=0}^{\Delta_{min}-1} W_{j+b_mod_i-1} a_j - W_{\Delta_{min}+b_mod_i-1} a_{\Delta_{min}}$$

по модулям m_i (см. (15), (16)), генерируемые согласно правилу

$$TRes_MP_i[(a_{\Delta_{min}} a_{\Delta_{min}-1} \dots a_0)_2] = |A_i(a_0, a_1, \dots, a_{\Delta_{min}})|_{m_i} (a_j \in \{0, 1\} (j = \overline{0, \Delta_{min}}); i = \overline{1, k}). \quad (19)$$

В случае, когда базис \mathbf{M} содержит только один модуль ($\mathbf{M} = \{m\}$), индекс i в (17), (18) опускается.

Алгоритм позиционно-модулярного кодового преобразования по высокоскоростной редуccionной схеме понижения разрядности вычетов по модулю состоит из следующих шагов:

– ПМ_РС.1. Положить $b_0 = b$, $X^{(0)} = (x_{b_0-1}^{(0)} x_{b_0-2}^{(0)} \dots x_0^{(0)})_2 = (x_{b-1} x_{b-2} \dots x_0)_2 = X$, $i=1$.

– ПМ_РС.2. Номеру итерации редуccionного процесса присвоить начальное значение $s=1$.

– ПМ_РС.3. Следуя (3), с учетом (17) вычислить $X^{(1)} = \sum_{j=0}^{b_0-1} W_{j,i}(C_i) x_j^{(0)}$, сформировав двоичный код $(x_{b_1-1}^{(1)} x_{b_1-2}^{(1)} \dots x_0^{(1)})_2$ длиной b_1 бит ЦЧ $X^{(1)}$.

– ПМ_РС.4. По старшей $(b_s - b_mod_i - \Delta_{min} + 1)$ -битовой части $(x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_{b_mod_i+\Delta_{min}-1}^{(s)})_2$ двоичного кода $(x_{b_s-1}^{(s)} x_{b_s-2}^{(s)} \dots x_0^{(s)})_2$ ЦЧ $x^{(s)}$ получить булеву величину

$$\Gamma = \left| \left(\bigvee_{j=b_mod_i+\Delta_{min}-1}^{b_s-1} x_j^{(s)} \right) + \left(\bigwedge_{j=b_mod_i+\Delta_{min}-1}^{b_s-1} x_j^{(s)} \right) \right|_2.$$

– ПМ_РС.5. Если $\Gamma=1$, то инкрементировать s ($s:=s+1$), найти

$$X^{(s)} = \sum_{j=0}^{b_{s-1}-2} W_{j,i} x_j^{(s-1)} - W_{b_{s-1}-1,i} x_{b_{s-1}-1}^{(s-1)},$$

получив код $(x_{b_{s-1}}^{(s)} x_{b_{s-2}}^{(s)} \dots x_0^{(s)})_2$ длиной b_s бит числа $X^{(s)}$, и перейти к ПМ_РС.4.

ПМ_РС.6. Ввиду $\Gamma=0$ в соответствии с (13) – (16), (18), (19) переменной S присвоить значение $S=s$ и для фиксации искомого значения i -й цифры формируемого модулярного кода выполнить следующие действия:

ПМ_РС.6А. Рассчитать вычет

$$X_0^{(s)} = \sum_{j=0}^{b_{\text{mod } i}-2} 2^j x_j^{(s)} = \sum_{j=0}^{b_{\text{mod } i}-2} W_{j,i} x_j^{(s)}.$$

ПМ_РС.6Б. Из таблицы $TRes_MP_i$ извлечь вычет

$$\begin{aligned} X_1^{(s)} &= TRes_MP_i[(x_{b_{\text{mod } i}+\Delta_{\text{min}}-1}^{(s)} x_{b_{\text{mod } i}+\Delta_{\text{min}}-2}^{(s)} \dots x_{b_{\text{mod } i}-1}^{(s)})_2] = \\ &= \left| \sum_{j=b_{\text{mod } i}-1}^{b_{\text{mod } i}+\Delta_{\text{min}}-2} 2^j x_j^{(s)} - 2^{b_{\text{mod } i}+\Delta_{\text{min}}-1} x_{b_{\text{mod } i}+\Delta_{\text{min}}-1}^{(s)} \right|_{m_i} = \\ &= \left| \sum_{j=b_{\text{mod } i}-1}^{b_{\text{mod } i}+\Delta_{\text{min}}-2} W_{j,i} x_j^{(s)} - W_{b_{\text{mod } i}+\Delta_{\text{min}}-1,i} x_{b_{\text{mod } i}+\Delta_{\text{min}}-1}^{(s)} \right|_{m_i}. \end{aligned}$$

ПМ_РС.6В. Найти сумму $\chi_i = X_0^{(s)} + X_1^{(s)}$.

ПМ_РС.6Г. При $\chi_i \geq m_i$ положить $\chi_i := \chi_i - m_i$.

ПМ_РС.7. Если $i \neq k$, то инкрементировать переменную i ($i:=i+1$) и перейти к ПМ_РС.2.

ПМ_РС.8. Завершить работу алгоритма.

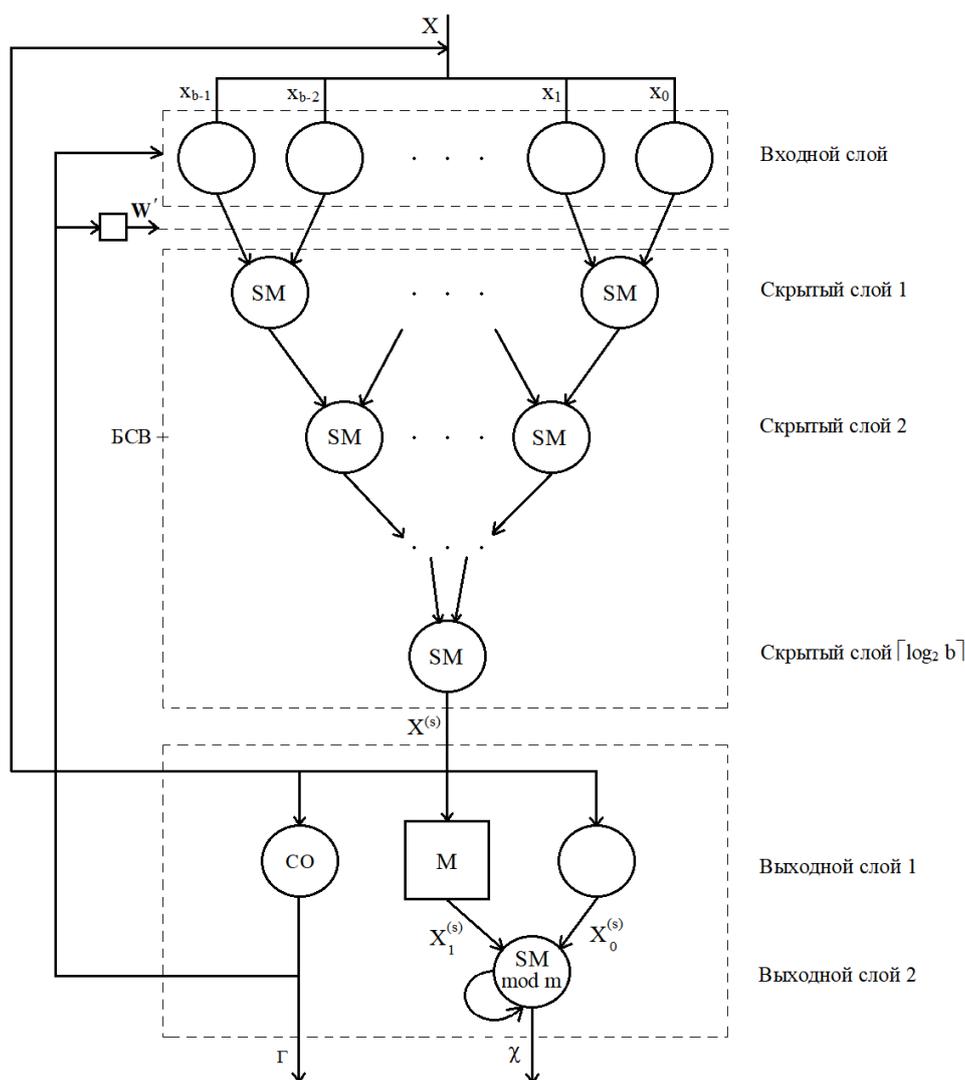
Нейросетевая реализация редукционного алгоритма позиционно-модулярного преобразования масштабируемых чисел. Синтезированная процедура ПМ_РС.1–ПМ_РС.8 приведения масштабируемых ЦЧ к остаткам по модулям $m \in \mathbf{M} = \{m_1, m_2, \dots, m_k\}$ ($k \geq 1$) может быть реализована как программно, так и аппаратным способом с применением нейросетевой вычислительной технологии [1, 3, 14, 15].

На рисунке изображена структура параллельной НСКК с обратной связью. Данная НС содержит входной слой, нейроны которого образуют b -разрядный регистр для фиксации двоичного кода $(x_{b-1} x_{b-2} \dots x_0)_2$ ЦЧ X , $\lceil \log_2 b \rceil$ скрытых слоев, в совокупности составляющих блок суммирования вычетов – взвешенных компонент наборов вида (17), (18):

$$\mathbf{W}(C) = \{W_j(C) | W_j(C) = |C - 2^j|_m^- ; j = \overline{0, b-1}\}; \quad (20)$$

$$\mathbf{W} = \{W_j | W_j = |2^j|_m^- ; j = \overline{0, b_1-1} ; b_1 < b\}, \quad (21)$$

а также два выходных слоя, осуществляющих проверку принадлежности получаемых блоком суммирования вычетов (БСВ) чисел $X^{(s)} = (x_{b_{s-1}}^{(s)} x_{b_{s-2}}^{(s)} \dots x_0^{(s)})_2$ ($s = \overline{1, S}$) к диапазону $\mathbf{Z}_{b_{\text{mod } i}+\Delta_{\text{min}}}^-$ и приведение ЦЧ $X^{(s)}$ к остатку по модулю m .



Параллельная НСКК с обратной связью для позиционно-модулярного преобразования по рекурсивной редукционной схеме

В скрытых слоях используются сумматоры SM , которые выполняют операции сложения пар вычетов, формируемых в соответствующих предшествующих слоях. Если в l -й слой БСВ ($l = \overline{1, \lceil \log_2 b \rceil}$) поступает нечетное число вычетов, то вычет, не вошедший в пару, подвергается задержке на время сложения пар вычетов в данном слое. Обладая параллельной $\lceil \log_2 b \rceil$ -каскадной «пирамидальной» архитектурой, БСВ формирует дополнительный двоичный код ЦЧ конвейерного типа. Это обеспечивает получение на выходе БСВ двоичного кода ЦЧ $X^{(s)}$ (см. (3)) за время $\lceil \log_2 b \rceil t_{\text{сл}}$, где $t_{\text{сл}}$ – длительность операции сложения двух вычетов.

Изображенная на рисунке НСКК работает следующим образом. Двоичный код $(x_{b-1} x_{b-2} \dots x_0)_2$ ЦЧ X , подлежащего преобразованию $X \rightarrow |CX|_m$, фиксируется в нейронах входного слоя, откуда вместе с набором (20) синаптических весов поступает затем в БСВ. Реализуя рекурсивную $\lceil \log_2 b \rceil$ -каскадную процедуру суммирования аддитивных компонент выражения для $X^{(1)}$ (см. (3)) в режиме максимального распараллеливания вычислительного процесса на уровне двухместных операций сложения в скрытых слоях с первого по $\lceil \log_2 b \rceil$ -й, БСВ получает дополнительный двоичный код $(x_{b_1-1}^{(1)} x_{b_1-2}^{(1)} \dots x_0^{(1)})_2$ числа $X^{(1)}$. Старшая $(\Delta' + 1)$ -битовая часть $(x_{b_1-1}^{(1)} x_{b_1-2}^{(1)} \dots x_{b_1-\Delta'-1}^{(1)})_2$ ($\Delta' = b_1 - b_{\text{mod}} - \Delta_{\text{min}}$) сформированного кода подается в управляющий элемент CO первого выходного слоя, который в соответствии с булевым выражением (10), рассматриваемым при $s=1$, генерирует сигнал Δ . Единичное значение сигнала Δ указы-

вает на то, что ЦЧ, полученное на выходе БСВ, находится за пределами диапазона $Z_{b_mod+\Delta_{min}}^-$. В этом случае $\Delta = 1$ инициирует включение в НСКК обратной связи, что обеспечивает переход к очередной итерации редуционного процесса (3). Двоичный код с выхода БСВ передается во входной слой и согласно правилу

$$W' = \begin{cases} W(C), & \text{если } \Gamma = 0, \\ W, & \text{если } \Gamma = 1, \end{cases}$$

набор $W(C)$ синаптических весов замещается набором W (см. (21)). В результате на выходе БСВ сформируется двоичный код очередного элемента последовательности вычетов $X^{(1)}, X^{(2)}, \dots, X^{(S)}$.

Функционирование НСКК в рекурсивном режиме завершается с появлением на выходе БСВ в рамках заключительной S -й итерации двоичного кода $(x_{b_s-1}^{(S)} x_{b_s-2}^{(S)} \dots x_0^{(S)})_2$ ЦЧ $X^{(S)}$ из диапазона $Z_{2b_mod+\Delta_{min}}^-$. В данном случае сигнал Γ , генерируемый управляющим элементом CO , принимает значение $\Gamma=0$, что приводит к блокировке обратной связи. При этом извлекаемый из табличной памяти M первого выходного слоя по адресу $(x_{b_mod+\Delta_{min}-1}^{(S)} x_{b_mod+\Delta_{min}-2}^{(S)} \dots x_{b_mod-1}^{(S)})_2$ вычет $X_1^{(S)}$ вместе с сохраняемым в регистре вычетом $X_0^{(S)}$ (см. (14)–(16)) подаются на входы сумматора $SMmod\ m$ второго выходного слоя, который в соответствии с (13) получает искомым остаток: $\chi = |CX|_m = |X^{(S)}|_m$.

Отметим, что значение параметра Δ_{min} можно регулировать, поэтому описанная конфигурация НСКК с обратной связью позволяет гибко устанавливать оптимальный баланс между количеством итераций применяемой редуционной схемы и объемом используемой табличной памяти.

Заключение. В статье представлены наиболее важные результаты разработки по созданию базовых НСКК для высокопроизводительных МА-приложений в области защиты информации.

Для построения НС на кольцах вычетов по модулям МСС как основы нейросетевого обеспечения криптографических МА-приложений использована редуционная технология позиционно-модулярного преобразования масштабируемых чисел. Развиваемые подходы к решению поставленной задачи обеспечивают оптимальные условия для согласования и реализации фундаментальных свойств параллелизма НС и МА. При этом для повышения скорости сходимости базовой рекурсивной вычислительной схемы применен табличный метод сокращения количества необходимых итераций.

Синтезирован новый эффективный алгоритм ускоренного преобразования масштабируемых чисел из двоичной системы в модулярную систему счисления по редуционной схеме понижения разрядности элементов формируемой последовательности вычетов. Благодаря применению гибкого адаптивного механизма проверки принадлежности поитерационных вычетов к специально выбираемому диапазону, допускающему табличную декомпозицию его элементов на пары компонент – остатков по модулям заданного базиса, реализуемый метод позволяет достичь значительного повышения скорости осуществляемого преобразования.

Разработана параллельная структура НСКК с обратной связью, выполняющей позиционно-модулярное преобразование масштабируемого целого числа по редуционной схеме за минимизированное количество S итераций – суммарное время $(S(\lceil \log_2 b \rceil + 1) + 2)t_{сл}$, где b – разрядность входного числа, $t_{сл}$ – длительность операции сложения двух вычетов.

Список использованных источников

1. Модулярная арифметика и ее приложения в инфокоммуникационных технологиях / Н. И. Червяков [и др.]. – М. : Физматлит, 2017. – 400 с.
2. Ananda Mohan, P. V. Residue Number Systems: Theory and Applications / P. V. Ananda Mohan. – Basel : Birkhauser, Mathematics, 2016. – 351 p.
3. Применение искусственных нейронных сетей и системы остаточных классов в криптографии / Н. И. Червяков [и др.]. – М. : Физматлит, 2012. – 280 с.
4. Инютин, С. А. Основы модулярной алгоритмики / С. А. Инютин. – Ханты-Мансийск : Полиграфист, 2009. – 347 с.

5. Omandi, A. *Residue Number Systems: Theory and Implementation* / A. Omandi, B. Premkumar. – Singapore : Imperial College Press, 2007. – 311 p.
6. Оцоков, Ш. А. Способ организации высокоточных вычислений в модулярной арифметике / Ш. А. Оцоков // Первая Междунар. конф. «Параллельная компьютерная алгебра и ее приложения в новых инфокоммуникационных системах». – Ставрополь, 20–24 окт., 2014. – Ставрополь : ИИЦ «Фабула», 2014. – С. 270–277.
7. Fast modular network implementation support vector machines / G.-B. Gulang [et al.] // *IEEE Trans. Neural Networks*. – 2005. – Vol. 16, no. 6. – P. 1651–1663.
8. Тихонов, Э. Е. Программно-аппаратная реализация нейронных сетей : моногр. / Э. Е. Тихонов, А. А. Евдокимов. – Невинномысск : НИЭУП, 2013. – 116 с.
9. Sanches, D. Optimization of modular granular neural networks using arhierarchical genetic algorithm based on the database comlexcity applied to human recognition / D. Sanches, P. Melin, O. Castillo // *Informations Sciences. Tjuana Institute of Technology*. – 2015. – Vol. 309. – P. 73–101.
10. Кондрашев, А. В. Нейронная сеть для преобразования чисел, представленных в позиционном коде, в систему остаточных классов [Электронный ресурс] / А. В. Кондрашев, Д. В. Горденко, Д. Н. Павлюк // Исследования в области естественных наук. – 2015. – № 1. – Режим доступа: <http://science.snauka.ru/2015/01/8925>. – Дата доступа: 22.01.2018.
11. Разработка нового нейросетевого метода вычисления модульного умножения в системе остаточных классов / М. Г. Бабенко [и др.] // *Нейрокомпьютеры: разработка и применение*. – 2016. – № 10. – С. 41–48.
12. Коляда, А. А. Обобщенная интегрально-характеристическая база модулярных систем счисления / А. А. Коляда // *Информационные технологии*. – 2017. – Т. 23, № 9. – С. 641–649.
13. Чернявский, А. Ф. Преобразование кода модулярной системы счисления в обобщенный позиционный код / А. Ф. Чернявский, А. А. Коляда // Доклады Нац. акад. наук Беларуси. – 2017. – Т. 61, № 4. – С. 26–30.
14. Червяков, Н. И. Нейронные сети конечного кольца для реализации пороговых схем разделения секрета / Н. И. Червяков, А. А. Евдокимов // *Нейрокомпьютеры: разработка, применение*. – 2007. – № 2–3. – С. 45–50.
15. Червяков, Н. И. Нейронная сеть конечного кольца прямого распространения для операций на эллиптических кривых / Н. И. Червяков, А. Б. Спельников, А. Ф. Мезенцева // *Нейрокомпьютеры: разработка, применение*. – 2008. – № 1–2. – С. 28–34.

References

1. Chervjakov N. I., Koljada A. A., Ljahov P. A., Babenko M. G., Lavrinenko I. N., Lavrinenko A. V. *Moduljarnaja arifmetika i ee prilozhenija v infokommunikacionnyh tehnologijah. Modular Arithmetic and its Applications in Infocommunication Technologies*. Moscow, Fizmatlit Publ., 2017, 400 p. (in Russian).
2. Ananda Mohan P. V. *Residue Number Systems: Theory and Applications*. Basel, Birghauser, Mathematics, 2016, 351 p.
3. Chervjakov N. I., Evdokimov A. A., Galushkin A. I., Lavrinenko I. N., Lavrinenko A. V. *Primenenie iskusstvennyh nejronnyh setej i sistemy ostatocnyh klassov v kriptografii. The Use of Artificial Neural Networks and the Residual Class System in Cryptography*. Moscow, Fizmatlit Publ., 2012, 280 p. (in Russian).
4. Injutin S. A. *Osnovy moduljarnoj algoritmiki. Fundamentals of Modular Algorithms*. Khanty-Mansiysk, Poligrafist Publ., 2009, 347 p. (in Russian).
5. Omandi A., Premkumar B. *Residue Number Systems: Theory and Implementation*. Singapore, Imperial College Press, 2007, 311 p.
6. Ocovok Sh. A. *Sposob organizatsii vysokotochnykh vychislenij v moduljarnoj arifmetike [The way to organize high-precision calculations in modular arithmetic]*. Pervaja Mezhdunar. konf. «Parallelnaja komp'juternaja algebra i ee prilozhenija v novyh infokommunikacionnyh sistemah» [First International Conf. "Parallel Computer Algebra and Its Applications in New Infocommunication Systems"]. Stavropol, 20–24 Okt., 2014. Stavropol, Fabula Publ., 2014, pp. 270–277 (in Russian).
7. Gulang G.-B., Mao K.-Z., Siew C. K., Huang D.-S. Fast modular network implementation support vector machines. *IEEE Trans. Neural Networks*, 2005, vol. 16, no. 6, pp.1651–1663.
8. Tihonov Je. E., Evdokimov A. A. *Programmno-apparatnaja realizacija nejronnyh setej : monografija. Software and Hardware Implementation of Neural Networks*. Nevinnomyssk, NIJeUP Publ., 2013, 116 p. (in Russian).
9. Sanches D., Melin P., Castillo O. Optimization of modular granular neural networks using arhierarchical genetic algorithm based on the database comlexcity applied to human recognition. *Informations Sciences. Tjuana Institute of Technology*, 2015, vol. 309, pp. 73–101.
10. Kondrashjov A. V., Gordenko D. V., Pavljuk D. N. *Nejronnaya set' dlya preobrazovaniya chisel, predstavlennykh v pozitsionnom kode, v sistemu ostatocnyh klassov [Neural network for converting the numbers represented in the positional code to the residual class system]*. Issledovaniya v oblasti estestvennyh nauk [Research in the Field of Natural Sciences], 2015, no. 1 (in Russian). Available at: <http://science.snauka.ru/2015/01/8925> (accessed 22.01.2018).
11. Babenko M. G., Chernyh A. N., Kuchukov V. A., Derjabin M. A., Kuchukova N. N. *Razrabotka novogo nejrosетеvogo metoda vychisleniya modul'nogo umnozheniya v sisteme ostatocnyh klassov [Development of a new neural network method for calculating modular multiplication in a system of residual classes]*. *Nejrokompjutyery: razrabotka i primenenie [Neurocomputers: Development and Application]*, 2016, no. 10, pp. 41–48 (in Russian).
12. Koljada A. A. *Obobshhennaya integral'no-kharakteristicheskaya baza moduljarnykh sistem schisleniya [Generalized integral-characteristic base of modular number systems]*. *Informacionnye tehnologii [Information Technology]*, 2017, vol. 23, no. 9, pp. 641–649 (in Russian).

13. Chernjavskij A. F., Koljada A. A. Preobrazovanie koda modulyarnoj sistemy schisleniya v obobshhennyj pozitsionnyj kod [Conversion of a modular number system code to a generalized positional code]. *Doklady Natsional'noi akademii nauk Belarusi [Doklady of the National Academy of Sciences of Belarus]*, 2017, vol. 61, no. 4, pp. 26–30 (in Russian).

14. Chervjakov N. I., Evdokimov A. A. Nejronnye seti konechnogo kol'tsa dlya realizatsii porogovykh skhem razdeleniya sekreta [Neural networks of the finite ring for the implementation of threshold separation schemes for secretion]. *Nejrokompjutery: razrabotka, primenenie [Neurocomputers: Development and Application]*, 2007, no. 2–3, pp. 45–50 (in Russian).

15. Chervjakov N. I., Spel'nikov A. B., Mezenceva A. F. Nejronnaya set' konechnogo kol'tsa pryamogo rasprostraneniya dlya operatsij na ehllipticheskikh krivykh [Neural network of a finite ring of direct propagation for operations on elliptic curves]. *Nejrokompjutery: razrabotka, primenenie [Neurocomputers: Development and Application]*, 2008, no. 1–2, pp. 28–34 (in Russian).

Информация об авторах

Червяков Николай Иванович – доктор технических наук, профессор, Северо-Кавказский федеральный университет (пр. Кулакова, 2, 355029, Ставрополь, Российская Федерация). E-mail: Chervyakov@yandex.ru

Коляда Андрей Алексеевич – доктор физико-математических наук, доцент, главный научный сотрудник лаборатории специализированных вычислительных систем, Научно-исследовательское учреждение «Институт прикладных физических проблем им. А. Н. Севченко» Белорусского государственного университета (ул. Курчатова, 7, 220045, Минск, Республика Беларусь). E-mail: razan@tut.by

Коляда Назар Андреевич – научный сотрудник лаборатории специализированных вычислительных систем, Научно-исследовательское учреждение «Институт прикладных физических проблем им. А. Н. Севченко» Белорусского государственного университета (ул. Курчатова, 7, 220045, Минск, Республика Беларусь). E-mail: razan@tut.by

Кучуков Виктор Андреевич – специалист-патентовед отдела научно-технической информации, наукометрии и экспортного контроля, профессор, Северо-Кавказский федеральный университет (пр. Кулакова, 2, 355029, Ставрополь, Российская Федерация). E-mail: patentncfu@yandex.ru

Протасеня Стелла Юрьевна – младший научный сотрудник лаборатории специализированных вычислительных систем, Научно-исследовательское учреждение «Институт прикладных физических проблем им. А. Н. Севченко» Белорусского государственного университета (ул. Курчатова, 7, 220064, Минск, Республика Беларусь). E-mail: Estellita@mail.ru

Information about the authors

Nikolai I. Chervyakov – D. Sc. (Engineering), Professor, North-Caucasus Federal University (2, Kulakova Ave., 355029, Stavropol, Russian Federation). E-mail: Chervyakov@yandex.ru

Andrey A. Kolyada – D. Sc. (Physics and Mathematics), Chief Researcher of the Laboratory of Specialized Computational Systems, Scientific Research Institution "Institute of Applied Physical Problems named after A. N. Sevchenko" of the Belarusian State University (7, Kurchatov Str., 220045, Minsk, Republic of Belarus). E-mail: razan@tut.by

Nazar A. Kolyada – Researcher, Laboratory of Specialized Computing Systems, Scientific Research Institution "Institute of Applied Physical Problems named after A. N. Sevchenko" of the Belarusian State University (7, Kurchatov Str., 220045, Minsk, Republic of Belarus). E-mail: razan@tut.by

Victor A. Kuchukov – Patent Holder of the Department of Scientific and Technical Information, Sciencemetry and Export Control, Professor, North-Caucasus Federal University (2, Kulakova Ave., 355029, Stavropol, Russian Federation).

E-mail: patentncfu@yandex.ru

Stella Y. Protasenia – Junior Scientific Employee, Laboratory of Specialized Computational Systems, Scientific Research Institution "Institute of Applied Physical Problems named after A. N. Sevchenko" of the Belarusian State University (7, Kurchatov Str., 220064, Minsk, Republic of Belarus). E-mail: Estellita@mail.ru

Правила для авторов

Редакция журнала «Информатика» просит авторов руководствоваться приведенными ниже правилами:

1. Статьи принимаются в редакцию через электронную систему подачи по адресу <http://inf.grid.by> в формате файлов текстовых редакторов Microsoft Word. Основной текст статьи не должен превышать 17 стр., включая рисунки, таблицы и достаточное количество наиболее актуальных ссылок; обзорной статьи – 10 стр., включая все основные ссылки. Текст набирается с переносами, шрифт Times New Roman 11 пт, интервал между строками одинарный, абзацный отступ 1 см, поля по 2,5 см со всех сторон.

Изложенный в статье материал должен быть четко структурированным: введение, цели и задачи, методы, результаты, заключение (выводы).

2. Статьи о результатах работ, проведенных в научных учреждениях, должны иметь разрешение на публикацию (сопроводительное письмо за подписью руководителя или выписку из заседания ученого совета, отдела или кафедры, акт экспертизы).

3. Статья в обязательном порядке должна иметь следующую структуру: индекс по универсальной десятичной классификации (УДК); инициалы и фамилии всех авторов, название статьи, полное название учреждений, где работают авторы, с указанием города, страны, аннотацию (150–250 слов), подрисуночные надписи, названия таблиц и ключевые слова (7–10) на русском и английском языках, адрес электронной почты каждого автора и номер телефона.

4. Аннотация (авторское резюме) должна кратко представлять результаты работы и быть информативной, содержательной. Приветствуется структура аннотации, повторяющая структуру статьи и включающая введение, цели и задачи, методы, результаты, заключение.

5. Формулы, рисунки, таблицы в статье нумеруются в соответствии с порядком их упоминания в тексте. Ссылки на рисунки и таблицы в тексте обязательны. Рисунки должны быть выполнены с хорошим разрешением в масштабе, позволяющем четко различать надписи и обозначения. Подрисуночные подписи с расшифровкой всех позиций, представленных на рисунке, набираются шрифтом гарнитуры основного текста размером 9 пт. Цветные иллюстрации печатаются только в том случае, когда это необходимо для понимания излагаемого материала.

6. Набор формул выполняется в формульном редакторе Microsoft Equation или Math Type. Прямым шрифтом набираются: греческие и русские буквы; математические символы (\sin , \lg , ∞); символы химических элементов (C, Cl, CHCl₃); цифры (римские и арабские); векторы; индексы (верхние и нижние), являющиеся сокращениями слов. Курсивом набираются латинские буквы, символы физических величин (в том числе и в индексе).

7. Сокращения в тексте статьи (за исключением единиц измерения) могут быть использованы только после упоминания полного термина. Единицы измерения физических величин следует приводить в Международной системе единиц (СИ).

8. Цитируемые в статье фамилии авторов теорем, теорий, законов и т. д. следует приводить в скобках на языке оригинала после русского написания. Например, теорема Эйлера (Euler).

9. Список использованной литературы оформляется в соответствии с требованиями Высшей аттестационной комиссии Республики Беларусь (ГОСТ 7.5–2008). Номер литературной ссылки в тексте дается порядковым номером в квадратных скобках. Ссылаться на неопубликованные работы не допускается.

10. Отдельно приводится список цитированных источников в *романском* (латинском) алфавите со следующей структурой: авторы (транслитерация), название статьи в транслитерированном варианте [перевод названия статьи на английский язык в квадратных скобках], название русскоязычного источника (транслитерация) [перевод названия источника на английский язык – парафраз (для журналов можно не делать)], выходные данные с обозначениями на английском языке.

11. Поступившие в редакцию статьи направляются на рецензирование специалистам. Основным критерием целесообразности публикации является новизна и информативность статьи. Если по рекомендациям рецензента статья возвращается автору на доработку, а переработанная рукопись вновь рассматривается редколлегией, датой поступления считается день получения редакцией ее окончательного варианта. Статьи не по профилю журнала возвращаются авторам после заключения редколлегии.

12. Статьи, направляемые на доработку, должны быть возвращены в исправленном виде с ответами на все замечания.

13. Редакция журнала предоставляет возможность первоочередного опубликования статей, представленных лицами, которые осуществляют послевузовское обучение (аспирантура, докторантура, соискательство) в год завершения обучения.

14. Авторы несут ответственность за направление в редакцию статей, уже опубликованных ранее или принятых к публикации другими изданиями.

15. Редакция оставляет за собой право на редакционные изменения, не искажающие основное содержание статьи. Окончательное решение о публикации принимается редакционной коллегией.

Журнал «Информатика» включен Высшей аттестационной комиссией Республики Беларусь в список научных изданий для опубликования результатов диссертационных исследований.

Индексы

00827

для индивидуальных
подписчиков

008272

для предприятий и
организаций