

ISSN 1684-8853 (print); ISSN 2541-8610 (online)

ИНФОРМАЦИОННО- УПРАВЛЯЮЩИЕ СИСТЕМЫ

НАУЧНЫЙ ЖУРНАЛ

5(114)/2021

5(114)/2021

PEER REVIEWED JOURNAL

INFORMATSIONNO- UPRAVLIAIUSHCHIE SISTEMY (INFORMATION AND CONTROL SYSTEMS)

Founder

«Information and Control Systems», Ltd.

PublisherSaint Petersburg State University
of Aerospace Instrumentation**Editor-in-Chief**

M. Sergeev

Dr. Sc., Professor, Saint Petersburg, Russia

Deputy Editor-in-Chief

E. Krouk

Dr. Sc., Professor, Moscow, Russia

Executive secretary

O. Muravtsova

Editorial Board

S. Andreev

Dr. Sc., Tampere, Finland

V. Anisimov

Dr. Sc., Professor, Saint Petersburg, Russia

B. Bezruchko

Dr. Sc., Professor, Saratov, Russia

N. Blaunstein

Dr. Sc., Professor, Beer-Sheva, Israel

M. Buzdalov

PhD, Researcher, Saint Petersburg, Russia

C. Christodoulou

PhD, Professor, Albuquerque, New Mexico, USA

A. Dudin

Dr. Sc., Professor, Minsk, Belarus

I. Dumer

PhD., Professor, Riverside, USA

M. Favorskaya

Dr. Sc., Professor, Krasnoyarsk, Russia

L. Fortuna

PhD, Professor, Catania, Italy

A. Fradkov

Dr. Sc., Professor, Saint Petersburg, Russia

A. Hramov

Dr. Sc., Professor, Innopolis, Russia

L. Jain

PhD, Professor, Canberra, Australia

V. Khimenko

Dr. Sc., Professor, Saint Petersburg, Russia

G. Matvienko

Dr. Sc., Professor, Tomsk, Russia

A. Myllari

PhD, Professor, Grenada, West Indies

Y. Podoplyokin

Dr. Sc., Professor, Saint Petersburg, Russia

K. Samouylov

Dr. Sc., Professor, Moscow, Russia

J. Seberry

PhD, Professor, Wollongong, Australia

A. Shalyto

Dr. Sc., Professor, Saint Petersburg, Russia

A. Shepeta

Dr. Sc., Professor, Saint Petersburg, Russia

Yu. Shokin

RAS Academician, Dr. Sc., Novosibirsk, Russia

A. Smirnov

Dr. Sc., Professor, Saint Petersburg, Russia

T. Sutikno

PhD, Associate Professor, Yogyakarta, Indonesia

Z. Yuldashev

Dr. Sc., Professor, Saint Petersburg, Russia

R. Yusupov

RAS Corr. Member, Dr. Sc., Professor, Saint Petersburg, Russia

A. Zeifman

Dr. Sc., Professor, Vologda, Russia

Editor: A. Larionova**Proofreader:** T. Zvertanovskaia**Design:** M. Chernenko, Yu. Umnitsyna**Layout and composition:** Yu. Umnitsyna**Contact information**

The Editorial and Publishing Center, SUAI

67, B. Morskaia, 190000, Saint Petersburg, Russia

Website: <http://i-us.ru/en>, e-mail: i-us.spb@gmail.com

Tel.: +7 - 812 494 70 02

THEORETICAL AND APPLIED MATHEMATICS**Vostrikov A. A.** *Matrix vitrages and regular Hadamard matrices* 2**INFORMATION PROCESSING AND CONTROL****Cong Thang Pham, Thi Thu Thao Tran, Minh Tuan Pham, Thanh Cong Nguyen.** *Combined total variation of first and fractional orders for Poisson noise removal in digital images* 10**Mikhailov V. V., Kolpaschikov L. A., Sobolevskii B. A., Soloviev N. V., Yakushev G. K.** *Methodological approaches and algorithms for recognizing and counting animals in aerial photographs* 20**Timofeev A. L., Sultanov A. Kh.** *Influence of noise and sampling rate on the discrete image representation error* 33**INFORMATION AND CONTROL SYSTEMS****Demidovskij A. V., Babkin E. A.** *Adapting Neural Turing Machines for linguistic assessments aggregation in neural-symbolic decision support systems* 40**INFORMATION CHANNELS AND MEDIUM****Burkov A. A.** *Signal power and energy-per-bit optimization problems in mMTC systems* 51**INFORMATION ABOUT THE AUTHORS** 59

5(114)/2021

ИНФОРМАЦИОННО-
УПРАВЛЯЮЩИЕ
СИСТЕМЫ

РЕЦЕНЗИРУЕМОЕ ИЗДАНИЕ

Учредитель

ООО «Информационно-управляющие системы»

Издатель

Санкт-Петербургский государственный университет
аэрокосмического приборостроения

Главный редактор

М. Б. Сергеев,
д-р техн. наук, проф., Санкт-Петербург, РФ

Зам. главного редактора

Е. А. Крук,
д-р техн. наук, проф., Москва, РФ

Ответственный секретарь

О. В. Муравцова

Редакционная коллегия:

С. Д. Андреев,
д-р техн. наук, Тампере, ФинляндияВ. Г. Анисимов,
д-р техн. наук, проф., Санкт-Петербург, РФБ. П. Безручко,
д-р физ.-мат. наук, проф., Саратов, РФН. Блаунштейн,
д-р физ.-мат. наук, проф., Беэр-Шева, ИзраильМ. В. Буздалов,
канд. техн. наук, научный сотрудник, Санкт-Петербург, РФЛ. С. Джайн,
д-р наук, проф., Канберра, АвстралияА. Н. Дудин,
д-р физ.-мат. наук, проф., Минск, БеларусьИ. И. Думер,
д-р наук, проф., Риверсайд, СШАА. И. Зейфман,
д-р физ.-мат. наук, проф., Вологда, РФК. Кристофолу,
д-р наук, проф., Альбукерке, Нью-Мексико, СШАГ. Г. Матвиенко,
д-р физ.-мат. наук, проф., Томск, РФА. А. Мюллари,
д-р наук, профессор, Гренада, Вест-ИндияЮ. Ф. Подоплёкин,
д-р техн. наук, проф., Санкт-Петербург, РФК. Е. Самуйлов,
д-р техн. наук, проф., Москва, РФД. Себерри,
д-р наук, проф., Волонгонг, АвстралияА. В. Смирнов,
д-р техн. наук, проф., Санкт-Петербург, РФТ. Сутикнуо,
д-р наук, доцент, Джокьякарта, ИндонезияМ. Н. Фаворская,
д-р техн. наук, проф., Красноярск, РФЛ. Фортуна,
д-р наук, проф., Катания, ИталияА. Л. Фрадков,
д-р техн. наук, проф., Санкт-Петербург, РФВ. И. Хименко,
д-р техн. наук, проф., Санкт-Петербург, РФА. Е. Храмов,
д-р физ.-мат. наук, Иннополис, РФА. А. Шалыто,
д-р техн. наук, проф., Санкт-Петербург, РФА. П. Шепета,
д-р техн. наук, проф., Санкт-Петербург, РФЮ. И. Шокин,
акад. РАН, д-р физ.-мат. наук, проф., Новосибирск, РФЗ. М. Юлдашев,
д-р техн. наук, проф., Санкт-Петербург, РФР. М. Юсупов,
чл.-корр. РАН, д-р техн. наук, проф., Санкт-Петербург, РФ

Редактор: А. Г. Ларионова

Корректор: Т. В. Звертановская

Дизайн: М. Л. Черненко, Ю. В. Умницына

Компьютерная верстка: Ю. В. Умницына

Адрес редакции: 190000, Санкт-Петербург,

Б. Морская ул., д. 67, ГУАП, РИЦ

Тел.: (812) 494-70-02, эл. адрес: ius.spb@gmail.com,

сайт: http://i-us.ru

ТЕОРЕТИЧЕСКАЯ И ПРИКЛАДНАЯ МАТЕМАТИКА

Востриков А. А. Матричные витражи и регулярные матрицы
Адамара

2

ОБРАБОТКА ИНФОРМАЦИИ И УПРАВЛЕНИЕ

**Cong Thang Pham, Thi Thu Thao Tran, Minh Tuan Pham,
Thanh Cong Nguyen.** Combined total variation of first and fractional
orders for Poisson noise removal in digital images

10

**Михайлов В. В., Колпациков Л. А., Соболевский В. А.,
Соловьев Н. В., Якушев Г. К.** Методологические подходы
и алгоритмы распознавания и подсчета животных
на аэрофотоснимках

20

Тимофеев А. Л., Султанов А. Х. Влияние шума и частоты
дискретизации на погрешность дискретного представления
изображений

33

ИНФОРМАЦИОННО-УПРАВЛЯЮЩИЕ СИСТЕМЫ

Демидовский А. В., Бабкин Э. А. Адаптация нейронных машин
Тьюринга для задачи агрегации лингвистических оценок
в нейросимволических системах поддержки принятия решений

40

ИНФОРМАЦИОННЫЕ КАНАЛЫ И СРЕДЫ

Burkov A. A. Signal power and energy-per-bit optimization problems
in mMTC systems

51

СВЕДЕНИЯ ОБ АВТОРАХ

59

Журнал входит в БД SCOPUS и в Перечень рецензируемых научных изданий,
в которых должны быть опубликованы основные научные результаты диссертаций
на соискание ученой степени кандидата наук,
на соискание ученой степени доктора наук.

Сдано в набор 01.09.21. Подписано в печать 21.10.21. Формат 60×84^{1/8}.

Гарнитура SchoolBookC. Печать цифровая.

Усл. печ. л. 7,3. Уч.-изд. л. 10,3. Тираж 1000 экз (1-й завод 50 экз.). Заказ № 463.

Оригинал-макет изготовлен в редакционно-издательском центре ГУАП.

190000, Санкт-Петербург, Б. Морская ул., 67.

Отпечатано с готовых диапозитивов в редакционно-издательском центре ГУАП.

190000, Санкт-Петербург, Б. Морская ул., 67.

Журнал зарегистрирован в Министерстве РФ по делам печати,
телерадиовещания и средств массовых коммуникаций.
Свидетельство о регистрации ПИ № 77-12412 от 19 апреля 2002 г.
Перерегистрирован в Роскомнадзоре.
Свидетельство о регистрации ПИ № ФС77-49181 от 30 марта 2012 г.

© Коллектив авторов, 2021

УДК 519.614

doi:10.31799/1684-8853-2021-5-2-9

Матричные витражи и регулярные матрицы Адамара

А. А. Востриков^а, канд. техн. наук, доцент, orcid.org/0000-0002-8513-3683, vostricov@mail.ru^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: кронекерово произведение матриц Адамара, когда одна матрица порядка n вставляется по месту каждого элемента другой матрицы порядка m , наследуя знак замещаемого элемента, рассматривается как основа получения этих ортогональных матриц порядка nm . Операция вставки, наследующая не только знаки, но и структурные элементы — орнаменты портретов матриц, рассматривается как операция создания витража — более общего результата. Витражи на базе типичных квазиортогональных матриц Мерсенна (M), Зейделя (S), Эйлера (E) и других, помимо наследования знака и орнамента (узора), иначе наследуют значение отличных от единицы (по амплитуде) элементов, вызывая необходимость пересмотреть и систематизировать накопленный опыт. **Цель:** описать новые алгоритмы обобщенного произведения матриц, выделяя конструкции, ведущие к регулярным матрицам Адамара высоких порядков. **Результаты:** предложен алгоритм получения матричных витражей вставкой матриц Мерсенна в матрицы Зейделя, позволяющий расширить аддитивные цепочки матриц вида $M-E-M-E-...$ и $S-E-M-E-...$, получаемые удвоениями порядков и добавлением каймы. Операция формирования матричного витража позволяет получать матрицы высоких порядков с сохранением такого важного инварианта структуры, как орнамент. Показано, что формирование матричного витража наследует логику произведения Скарпи, но не сводится к ней, поскольку ненулевое расстояние между сомножителями M и S по порядку упрощает итоговый орнамент регулярной матрицы отсутствием циклических смещений. Чередувание матриц M и S позволяет продолжить мультипликативные цепочки до известных пробелов в матрицах S . Это по-новому освещает теорию регулярных матриц Адамара как результатов произведения матриц Мерсенна и Зейделя. **Практическая значимость:** ортогональные последовательности с плавающими уровнями и алгоритмы эффективного нахождения регулярных матриц Адамара, выделенных рядом полезных свойств, имеют непосредственное практическое значение для задач помехоустойчивого кодирования, сжатия и маскирования видеоинформации.

Ключевые слова — ортогональные матрицы, регулярные матрицы, матрицы Адамара, матрицы Мерсенна, матрицы Зейделя, кронекерово произведение, алгоритм Скарпи, вставки матриц, матричные витражи.

Для цитирования: Востриков А. А. Матричные витражи и регулярные матрицы Адамара. *Информационно-управляющие системы*, 2021, № 5, с. 2–9. doi:10.31799/1684-8853-2021-5-2-9

For citation: Vostrikov A. A. Matrix vitrages and regular Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 2–9 (In Russian). doi:10.31799/1684-8853-2021-5-2-9

Введение

Квазиортогональные (ортогональные) матрицы широко используются в таких практических применениях, как обработка сигналов, сжатие, маскирование и помехоустойчивое кодирование изображений [1–4], криптография [5] и т. д. Это широко известны матрицы Адамара [6], матрицы Мерсенна [7] и др. Для поиска таких матриц известно большое количество методов и алгоритмов, отличающихся своей эффективностью при поиске матриц высоких порядков, или структурированных матриц, или матриц с целочисленными и вещественными элементами и пр. Наибольший интерес сегодня для задач обработки изображений представляет поиск квазиортогональных структурированных матриц высоких порядков [8–11], вычисления с которыми и их хранение наиболее эффективны [12].

Сегодня оборудование, осуществляющее обработку и передачу визуальной информации реального времени, оперирует разрешениями в 2, 4 и 8 тысяч пикселей внутри каждого видеокад-

ра при скорости обновления от 30 до 120 и более кадров/с. Матрицы исследуемых порядков выступают в роли операторов преобразования с различными целями, и увеличение их порядка до сотен и тысяч существенно повлияет как на скорость обработки, так и на качество восстанавливаемого изображения для потребителя. Кроме этого, ввиду малого числа уровней (количества различных значений элементов таких матриц) в процессе вычислений ресурсоемкая процедура умножения заменяется значительно более быстрой и менее энергозатратной процедурой выборки из памяти. Совокупный эффект в результате нахождения и применения новых ортогональных матриц высоких порядков способен оказать существенное влияние на эффективность перспективных систем цифровой обработки сигналов и их двумерного представления — цифровых видеоизображений.

В настоящей работе предлагаются новые алгоритмы поиска матриц на основе обобщенного произведения матриц с выделением конструкций, ведущих к регулярным матрицам Адамара высоких порядков.

Необходимые термины и определения

Определение 1. Квазиортогональная матрица A порядка n — это квадратная матрица, удовлетворяющая уравнению $A^T A = \omega(n)I$, где $\omega(n)$ — некоторая весовая функция, определяющая тип матрицы, а I — единичная матрица.

Можно использовать (с оговоркой) термин «взвешенная матрица», но это слово уже занято целочисленными ее представителями с элементами $\{0, 1, -1\}$ и линейной функцией веса $\omega(n) = n - k$, где k — целое число.

Случай $k = 0$ соответствует матрицам Адамара [6] с ее ненулевыми элементами; $k = 1$ соответствует матрицам Белевича (конференц-матрицам) [13]. Случай $k \geq 1$ относят к взвешенным матрицам $W(n, n - k)$, рассматриваемым как обобщение матриц Адамара в работах [7, 14]. Для иррациональных значений k подходит обозначение $W(n, \omega(n))$.

Иными словами, квазиортогональная матрица — это взвешенная $W(n, \omega(n))$ матрица с вещественными элементами $\{a = 1, -b\}$ в ней или образующих ее блоках. Инверсия знака при блоках не меняет состав базисных элементов. Третий элемент обычно выделен (как и у матриц Белевича) позиционированием на диагонали $\{d, a = 1, -b\}$ ($d \leq b \leq 1$) или кайме $\{a = 1, -b, s\}$ ($b \leq s \leq 1$) матрицы.

Чтобы не путать $W(n, n - k)$ с $W(n, \omega(n))$, профессор Дж. Себерри предложила называть последние *критскими* матрицами, подчеркивая нецелочисленность значений их элементов.

Как известно, *кронекеро* умножение $C = A \otimes B$ двух матриц A и B с элементами $\{1, -1\}$ реализуется вставкой матрицы B по месту элементов матрицы A с наследованием знака замещаемого элемента в виде

$$A \times B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & a_{n2}B & \dots & a_{nn}B \end{pmatrix}.$$

Результатом умножения, например, двух матриц Адамара порядков n и m будет матрица Адамара порядка nm . Именно произведение Кронекера использовалось для увеличения порядка ортогональных матриц первооснователями направления исследований Сильвестром и Адамаром.

Определение 2. Портрет квазиортогональной матрицы — визуальное изображение, на котором ее элементы представлены в виде клеток разных цветов, соответствующих значениям элементов.

На портрете матрицы совокупность разных по цвету клеток создает орнамент [15]. Для рассма-

триваемых в работе матриц возможны не более трех цветов клеток, соответствующих наборам значений элементов $\{1, -1\}$, $\{1, 0, -1\}$, $\{d, 1, -b\}$ или $\{1, -b, s\}$.

Первооткрывателя направления Сильвестра привлекала именно орнаментальная составляющая портрета матрицы: при обращении матрицы орнамент ее сохраняется с точностью до транспонирования. Адамар, в свою очередь, расширил количество матриц, используемых в произведениях (вставках), предложенных Сильвестром.

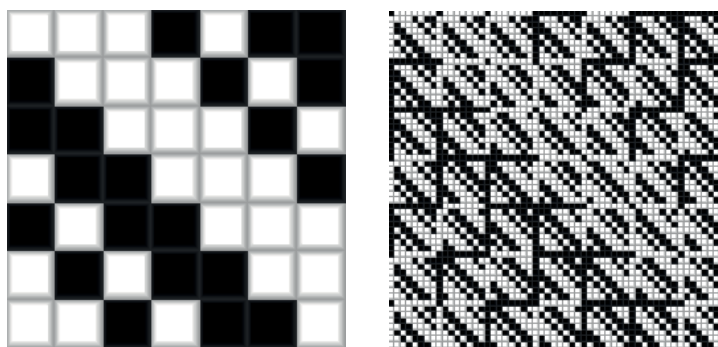
Определение 3. Витраж — это более общий результат кронекерова произведения двух матриц с косвенным наследованием числовых и структурных (орнамент матричного портрета) элементов.

У. Скарпи [16] предложил свой алгоритм реализации вставки спустя четыре года после опубликования Адамаром статьи [6], в которой тот сумел найти всего лишь пару новых матриц порядков 12 и 20, дополняющих порядки Сильвестра $n = 2^t$, где t — целое число.

Матрица Адамара после нормирования выведением в единицу элементов ее первых строки и столбца, называемых каймой, приобретает так называемую *основу* (core) отделением этой каймы. Порядки $2^t - 1$ представляют собой числа Мерсенна, причем полученную матрицу снова можно ортогонализировать введением плавающего уровня (изменением значения -1 на $-b$, где $|b| < 1$). Детерминант таких матриц выше, если число положительных элементов 1 выше числа отрицательных элементов $-b$ на единицу. Поэтому они инвертированы по знаку к основе (core), где это соотношение выдерживается с точностью до наоборот.

Получаемые в процессе этого преобразования ортогональные матрицы были названы матрицами Мерсенна [7]. Обнаружение существования таких матриц на порядках 11 и 19 позволило высказать предположение, что, как и матрицы Адамара, матрицы Мерсенна могут быть расширены на все порядки $n - 1$, где $n = 4t$ (гипотеза Н. А. Балонина). В этом виде гипотеза представляет собой альтернативную формулировку гипотезы Адамара о существовании всех выделенных им матриц на порядках, кратных 4 [17]. Но, поскольку матрицы Мерсенна — это, в том числе, и матрицы с иррациональными элементами (с учетом уровня $-b$), теория матриц с плавающим уровнем или уровнями открывает новые перспективы для доказательства упомянутых гипотез.

Произведение самим У. Скарпи было сформулировано громоздко, еще до введения понятия нормальной формы матрицы Адамара и ее основы, поэтому приведем его в нашей редакции. Формулировка даже значительно более простая по алгоритму вычисления произведения, чем предложенная Д. Джоковичем [18].



■ *Рис. 1.* Портрет матрицы Мерсенна M_7 и витраж — портрет матрицы Адамара H_{56}
 ■ *Fig. 1.* Portrait of the Mersenne matrix M_7 and stained glass — portrait of the Hadamard matrix H_{56}

Опираясь на определение матриц Мерсенна, произведение Скарпи можно переписать как обобщенное матричное произведение матриц M , где роль множителя произведения Кронекера играет кайма, которая расширяет размер вставляемой основы до размера матрицы Адамара. Такие произведения не характерны для алгоритмов Сильвестра и Адамара, поэтому Скарпи сумел получить ряд новых матриц порядка $n(n - 1)$, где $(n - 1)$ — порядки матриц Мерсенна.

Поскольку итоговый порядок кратен четырем, значение плавающего уровня $-b$ в результирующей матрице снова становится равным -1 . В результате получаем [19]

$$M \otimes M = \begin{pmatrix} \begin{pmatrix} -1 & m_{11}e^T \\ m_{11}e & M \end{pmatrix} & \begin{pmatrix} -1 & m_{12}e^T \\ m_{12}e & M \end{pmatrix} & \dots & \begin{pmatrix} -1 & m_{1(n-1)}e^T \\ m_{1(n-1)}e & M \end{pmatrix} \\ \begin{pmatrix} -1 & m_{21}e^T \\ m_{21}e & M \end{pmatrix} & \begin{pmatrix} -1 & m_{22}e^T \\ m_{22}e & TM \end{pmatrix} & \dots & \begin{pmatrix} -1 & m_{2(n-1)}e^T \\ m_{2(n-1)}e & T^{n-2}M \end{pmatrix} \\ \dots & \dots & \ddots & \dots \\ \begin{pmatrix} -1 & m_{(n-1)1}e^T \\ m_{(n-1)1}e & M \end{pmatrix} & \begin{pmatrix} -1 & m_{(n-1)2}e^T \\ m_{(n-1)2}e & T^{n-2}M \end{pmatrix} & \dots & \begin{pmatrix} -1 & m_{(n-1)(n-1)}e^T \\ m_{(n-1)(n-1)}e & T^{(n-2)(n-2)}M \end{pmatrix} \end{pmatrix},$$

где e — вектор единичных элементов каймы; T — матрица циклического смещения всех элементов, каждый блок смещается на величину произведения $(i - 1)(j - 1)$, индексы для смещений нумеруются с нуля. Символ сложного произведения не меняем.

В качестве иллюстрации на рис. 1 приведены портреты матрицы Мерсенна порядка 7 и результата матричного произведения Скарпи — витража, в котором вставляемая матрица циклически смещается на величину, определяемую произведением индексов ее положения. Знак наследуется не прямо, а в виде знака каймы.

Приведенная на рис. 1 матрица Мерсенна порядка 7 вставляется сама в себя, с добавлением каймы (первая точка ее всегда отрицательная). В итоге получаем матрицу порядка матрицы Адамара $7(7 + 1) = 56$, которая таковой и является, если считать элементы равными 1 и -1 . Из приведенного примера видно, что это произведение напоминает кронекерово и описывает достаточно просто, как вставку (витраж), обширный класс матриц, порождаемых матрицами Мерсенна.

Изобретение Скарпи предлагается расширить на произведения иных квазиортогональных матриц, например, порядков $(m - 1)$ и $(m + 1)$. В итоге становятся достижимы квадратичные порядки $m^2 - 1$, характерные для основ регулярных матриц Адамара порядка $n = m^2 = 4u^2$, где u — целое число.

Определение 4. Матрицы Адамара являются регулярными, если имеют одинаковые значения сумм элементов строки и столбцов.

Предикторы цепочек критских матриц

Наиболее интересными критскими матрицами, в аспекте данной статьи, являются матрицы Мерсенна M и матрицы Одина, совпадающие орнаментом с матрицами Зейделя S с точностью до зна-

чений их плавающих уровней. Матрицы M и S (далее матрица Одина) являются двумя предикторами цепочек критских матриц вида M (или S)– E – M – E –... [20].

Прежде всего, их можно использовать для получения простейших витражей силвестрова типа, получаемых умножением предикторов и членов цепочки на матрицу Адамара второго порядка H_2 , например $E = H_2 \otimes M$, без коррекции элементов. Для произведения $H_2 \otimes S$ витраж корректируют, исправляя элементы диагоналей матрицы Одина для получения уровня матрицы Эйлера E [19]. Здесь M — это матрица Мерсенна порядка $n = 4t - 1 = 3 \pmod{4}$, критская матрица с элементами $\{a = 1, -b\}$, $b = \frac{t}{t + \sqrt{t}}$, $k = (2t - 1) \times (1 - b^2)$, $\omega(n) = ((n + 1) + (n - 1)b^2)/2$, $t > 0$ — целое число. Рабочее предположение (гипотеза [17]) состоит в том, что эти матрицы существуют для любого выделенного для них порядка.

Для порядков значений простых целых чисел это циклические матрицы, для степеней целых чисел — блочные. Как видно из цепочки матриц E и M , большинство из них можно найти через бициклические матрицы Эйлера E . Будучи предиктором матриц Мерсенна, т. е. основой без специфической каймы, матрицы Эйлера тоже существуют на всех выделенных для них порядках $n = 4t - 2 = 2 \pmod{4}$. Весовую функцию и элементы матрицы Эйлера E несложно определить с помощью связи $E = H_2 \otimes M$, она сложена из двух циклических матриц Мерсенна. Аналогичен путь ее построения из циклических матриц Одина структуры Зейделя, с некоторыми поправками входящий в алгоритм [19, 21].

Матрица S порядка $n = 4t - 3 = 1 \pmod{4}$ — это критская матрица с элементами $\{d, a = 1, -b\}$, $b = 1 - 2d$, $d = \frac{1}{1 + \sqrt{n}}$, $k = 1 - d^2 + 2(t - 1)(1 - b^2)$, $\omega(n) = d^2 + ((n - 1) + (n - 1)b^2)/2$, где $t > 1$ — целое число. Эти матрицы существуют для любого выделенного для них порядка, если их порядок — сумма двух квадратов. Первые проблемные для матриц Одина порядки: 21, 33, 57 и т. п. До первой сотни существуют, но неизвестны (не найдены) матрицы Одина составных простых порядков 65 и 85.

Оригинальность бицикла E и его место в цепочках критских матриц состоит в том, что как связной элемент бесконечных цепочек итераций M (или S)– E – M – E –... он существует независимо от форм, которые принимают первоначальные матрицы Мерсенна и Одина. По сути, на старте это вполне себе независимая матрица, которую мы можем предопределять через предикторы M или S (служащих основой) только ради экономии сил и времени на ее отыскание.

Регулярные матрицы Адамара $H = [S \otimes M]$ или $H = [M \otimes S]$

Каждая из описанных нами цепочек в пределе стремится к матрице Адамара, но на бесконечно большом порядке, так как пара $\{a = 1, -b\}$ стремится к $\{a = 1, -1\}$. Иными словами, функция уровней $b(n)$ элементов матриц Мерсенна — монотонная, стремящаяся к единице. Особенность критских цепочек состоит в том, что формула Скарпи более универсальная и позволяет умножать не только матрицы Мерсенна (сами на себя), но и пары близко расположенных матриц $H = [S \otimes M]$ или $H = [M \otimes S]$. Используемые здесь квадратные скобки обозначают добавляемую к произведению (основе) кайму из единиц.

Формула не требует смещения, характерного для произведения Скарпи, диагональные элементы сомножителей замещены некоторыми компенсаторами, в остальном это обычное кронекерово произведение. Тем самым бесполезные, на первый взгляд, уровневые матрицы Одина выделены тем, что в сочетании с матрицами Мерсенна дают регулярные матрицы Адамара. Напомним, что в форме Буша (помимо тривиальных случаев, связанных с цепочкой матриц Силвестра) за всю историю поисков найдено всего три регулярные матрицы Адамара первых порядков 36, 100 и 324. Матрица порядка 196 не найдена.

Витраж как основа построения регулярных матриц Адамара

Структура витража, опирающаяся на кронекерово произведение матриц, относится к простым структурам. Она порождает квазиортогональную матрицу с абсолютными значениями элементов, равными произведениям сомножителей. Эта версия произведения нам менее интересна, хотя она и имеет право на существование как источник ортогональных матриц. Основой регулярных матриц Адамара являются нормализованные произведения $H = [S^{(1)} \otimes M^{(1)}]$ или $H = [M^{(1)} \otimes S^{(1)}]$ двух смежных (отличающихся на 2) простых порядков.

Сложный витраж отличается от результата кронекерова произведения общей для всей матрицы каймой и матрицей компенсатором, размещенным вдоль оси симметрии (диагонали).

В качестве блоков на оси симметрии (диагональных блоков) используется матрица из единиц J , если первый сомножитель больше по порядку второго. В противном случае — матрица $J - 2I$, где I — единичная матрица (знак центральных элементов квадрата компенсируется), причем диагональ вставляемой матрицы варьи-

руется в зависимости от знака замещаемого блоком элемента.

В остальном это произведение не отличается от кронекерова, чем и привлекательно. Вставки поднимают абсолютные значения элементов произведения квазиортогональных матриц до единицы. Это компенсационный принцип как у алгоритма Скарпи с его окаймлением блоков. Если порядки матриц Мерсенна составные, то используется сомножитель $M^{(2)}$ порядка $(2m + 1)$, построенный из матриц простых порядков m .

В любом таком случае мы можем указать положение компенсатора на диагонали: у матриц Белевича это нули; у произведения роль нуля играют матричные блоки без орнамента, входящие на ось симметрии или антисимметрии. Компенсатор помогает дополнять не равные по числу положительных и отрицательных элементов в строках (столбцах) матрицы сосредоточениями однородных элементов, обеспечивающих необходимый нулевой баланс.

Для освоения этой техники приведем примеры. Отметим, что циклическая форма матриц Мерсенна существует и для порядков, равных произведениям близких пар целых чисел, отличающихся на 2. Например, для порядков $3 \times 5 = 15$, $5 \times 7 = 35$, $7 \times 9 = 63$ и т. п., но они не отличаются явно выраженной осью симметрии. Такие ущербные матрицы наделены общей для всех матриц Мерсенна способностью порождать матрицу Адамара добавлением к этой основе каймы из отрицательных элементов — вид обобщенного умножения на матрицу Адамара первого порядка.

Регулярные матрицы Адамара

Регулярные матрицы Адамара существуют на порядках $4u^2 = (m + 1)(m - 1) + 1$, где u — целое число, т. е. 4, 16, 36, 100, 144, 196, 324, 196, 324, 676, 900, 1444, 1764, 2116, 2500, 2916 и т. п. Для вычисления промежуточных матриц порядков $21 \times 23 + 1 = 484$ и $33 \times 35 + 1 = 1156$ нет соответствующих матриц Одина.

Рассмотрим первые три циклические матрицы для получения витражей. Это матрицы Мерсенна M_3 , Одина S_5 и Мерсенна M_7 . Будем далее обозначать $M = (M_n)$, $S = (S_n)$ округленные до целых значений элементов матрицы с нулями на диагонали. В таком случае эти матрицы перестанут быть ортогональными, и их портреты совпадут с матрицами Q квадратичных вычетов с выделенными диагональными (осевыми) элементами, представленными на рис. 2.

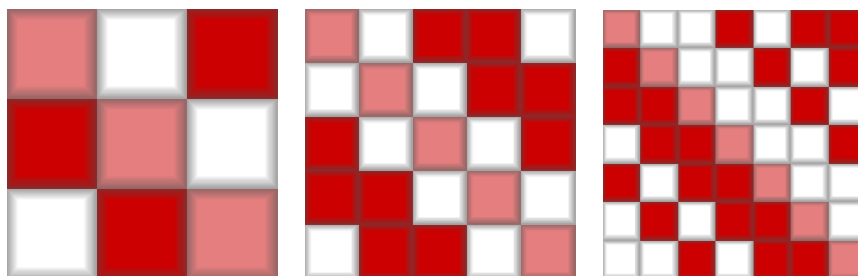
Для четных порядков 4 (тривиальный), 16, 36, 100, 144, (не $196 = 13 \times 15 + 1$), 324, (не 676), 900, (не 1444), 1764, (не 2116), 2500, (не 2916) возможны простые сомножители.

Алгоритм поиска. Для произведения $H = [M \otimes S]$ вычисляем $A = S - I$, $B = -S - I$, а для $H = [S \otimes M]$, наоборот: $A = M - I$, $B = -M - I$. Эти матрицы замещают положительный и отрицательный элементы первого сомножителя, диагональные элементы на оси симметрии замещаются $J - 2I$. Итоговую матрицу наращиваем каймой.

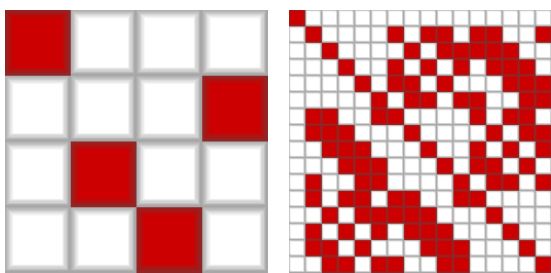
Тривиальный случай — это порядок $1 \times 3 + 1 = 4$, когда регулярная матрица Адамара получается расширением округленной матрицы Мерсенна M_3 с помощью каймы из положительных элементов, кроме первого отрицательного элемента. На случай порядка $3 \times 5 + 1 = 16$ матрица M_3 образует витраж с матрицей S_5 , ведущий к матрице H_{16} (рис. 3).

Для порядка $5 \times 7 + 1 = 36$ матрица S_5 образует витраж с матрицей M_7 , ведущий к матрице H_{36} . Произведем выравнивание сумм элементов строк и столбцов матриц Адамара инверсией второй половины блоков ее ядра. Регулярные матрицы H_{16} и H_{36} (рис. 4) сходны. Для синтеза матриц порядков $7 \times 9 + 1 = 64$ и $9 \times 11 + 1 = 100$ используем блочную матрицу $S_9 = M_3 \otimes M_3$, соответствующую кратной степени.

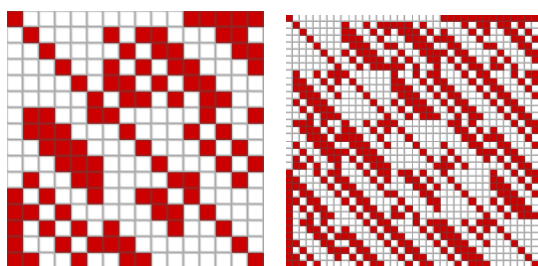
Первый по-настоящему не рабочий для такого алгоритма порядок это уже известный порядок 196 ($13 \times 15 + 1$). Заметим, что построение матрицы S_{13} проблем не вызывает. В разложении непростого числа $15 = 2 \times 7 + 1$ фигурирует простой порядок уже знакомой матрицы $M^{(1)} = M_7$, с помощью которой можно составить кососимметри-



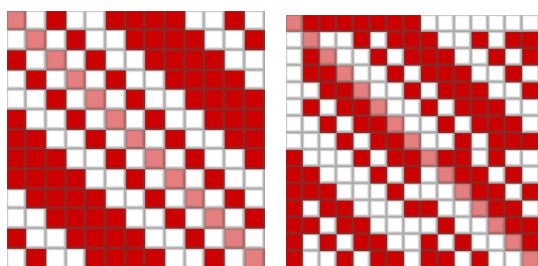
■ **Рис. 2.** Портреты трех циклических матриц квадратичных вычетов порядков 3, 5, 7
 ■ **Fig. 2.** Portraits of three cyclic matrices of quadratic residues of orders 3, 5, 7



■ *Рис. 3.* Портреты матриц Адамара H_4 и H_{16}
 ■ *Fig. 3.* Portraits of the Hadamard matrices H_4 и H_{16}



■ *Рис. 4.* Портреты регулярных матриц Адамара H_{16} и H_{36}
 ■ *Fig. 4.* Portraits of the regular Hadamard matrices H_{16} и H_{36}



■ *Рис. 5.* Портреты матриц S_{13} и $M^{(2)} = M_{15}$
 ■ *Fig. 5.* Portraits of the matrices S_{13} и $M^{(2)} = M_{15}$

ческую матрицу $M^{(2)} = M_{15}$ с симметрией второго порядка (рис. 5). Тем самым путь к построению регулярных матриц Адамара порядка 196 открыт.

Заключение

Скарпи трактовал свое произведение как череду операций с матрицей Адамара. Предложенный в работе алгоритм построения регулярных матриц Адамара через витражи является новым, поскольку в качестве сомножителей произведения не используются матрицы Адамара — впервые для этого вводятся критские матрицы с плавающими уровнями и манипуляции с ними, в том числе повышение уровня по мере роста порядка матрицы.

Кроме матриц Адамара, результатом обобщенного кронекерова произведения являются и матрицы Мерсенна более высоких порядков. В обобщенном кронекеровом произведении вполне можно обойтись без процедуры общего окаймления, трактуя его именно как путь получения матриц Мерсенна.

Кроме аддитивных цепочек матриц вида $M-E-M-E-\dots$ и $S-E-M-E-\dots$, получаемых удвоениями порядков и добавлением каймы, возникают мультипликативные цепочки генерации критских матриц.

Изложенный материал является основой для получения регулярной матрицы порядка 196 и далее матрицы порядка 676. Поскольку порядок матрицы S_{25} является квадратом простого числа, то для нее легко строится матрица квадратичных вычетов и регулярная матрица порядка 1444. Построение матрицы порядка 2116 зависит от блочной матрицы S_{45} , являющейся ядром особой конференц-матрицы S_{46} . И, наконец, матрицу S_{25} можно использовать при построении матрицы порядка 2916.

В практическом смысле результаты проведенных исследований существенно расширяют возможности применения цифровых систем обработки как одномерных сигналов, так и изображений высокого разрешения в целях их сжатия и защиты от несанкционированного доступа.

Благодарность

Автор выражает благодарность профессорам Н. А. Балонину и М. Б. Сергееву за конструктивные замечания и помощь в подготовке статьи.

Финансовая поддержка

Статья подготовлена при финансовой поддержке Министерства науки и высшего образования Российской Федерации, соглашение № FSRF-2020-0004.

Литература

1. Wang R. *Introduction to Orthogonal Transforms with Applications in Data Processing and Analysis*. Cambridge University Press, 2010. 504 p.
2. Mironovsky L. A., Slaev V. A. *Strip-Method for Image and Signal Transformation*. De Gruyter, 2011. 175 p.
3. Vostrikov A., Sergeev M. Expansion of the quasi-orthogonal basis to mask images. *Smart Innovation*,

- Systems and Technologies*, 2015, vol. 40. pp. 161–168. doi:10.1007/978-3-319-19830-9_15
4. Ahmed N., Rao R. *Orthogonal Transforms for Digital Signal Processing*. Berlin-Heidelberg-New York: Springer-Verlag, 1975. 263 p.
 5. Koukouvinos C., Simos D., Varbanov Z. Hadamard matrices, designs and their secret-sharing schemes. *Algebraic Informatics*, 2011, pp. 216–229. doi:10.1007/978-3-642-21493-6_14
 6. Hadamard J. Résolution d'une Question Relative aux Déterminants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246.
 7. Сергеев А. М. О взаимосвязи одного вида квазиортогональных матриц, построенных на порядках последовательностей $4k$ и $4k-1$. *Известия ЛЭТИ*, 2017, № 7, с. 12–17.
 8. Di Matteo O., Djokovic D. Z., Kotsireas I. S. Symmetric Hadamard matrices of order 116 and 172 exist. *Special Matrices*, 2015, vol. 3.1, pp. 227–234.
 9. Kharaghani H. and Tayfeh-Rezaie B. A Hadamard matrix of order 428. *J. Combin. Designs*, 2005, vol. 13, pp. 435–440.
 10. Horadam K. J. Hadamard matrices and their applications: Progress 2007–2010. *Cryptography and Communications*, 2010, no. 2, iss. 2, pp. 129–154.
 11. Seberry J., Yamada M. *Hadamard Matrices: Constructions using Number Theory and Linear Algebra*. John Wiley & Sons, 2020. 352 p.
 12. Vostrikov A., Sergeev M., Balonin N., Sergeev A. Use of symmetric Hadamard and Mersenne matrices in digital image processing. *Procedia Computer Science*, 2018, pp. 1054–1061. doi:10.1016/j.procs.2018.08.042
 13. Belevitch V. Theorem of $2n$ -terminal networks with application to conference telephony. *Electrical Communication*, 1950, no. 26, pp. 231–244.
 14. Balonin Yu. N., Sergeev A. M. Two-circulant Hadamard matrices, weighing matrices, and Ryser's conjecture. *Информационно-управляющие системы*, 2018, № 3, с. 2–9. doi:10.15217/issn1684-8853.2018.3.2
 15. Sergeev A., Sergeev M., Vostrikov A., Kurtyanik D. Portraits of orthogonal matrices as a base for discrete textile ornament patterns. *Smart Innovation, Systems and Technologies*, 2019, vol. 143, pp. 135–143. doi:10.1007/978-981-13-8303-8_12
 16. Scarpis U. Sui determinanti di valore massimo, *Rendiconti della R. Istituto Lombardo di Scienze e Lettere*, 1898, vol. 31, pp. 1441–1446.
 17. Сергеев А. М. Обобщенные матрицы Мерсенна и гипотеза Балонина. *Автоматика и вычислительная техника*, 2014, № 4, с. 35–43.
 18. Djokovic D. Generalization of Scarpis' theorem on Hadamard matrices. *Linear and Multilinear Algebra*, 2017, vol. 65, iss. 10, pp. 1985–1987.
 19. Балонин Н. А., Сергеев А. М., Сеницына О. И. Алгоритмы конечных полей и групп поиска ортогональных последовательностей. *Информационно-управляющие системы*, 2021, № 4, с. 2–17. doi:10.31799/1684-8853-2021-6-2-17
 20. Balonin N. A., Vostrikov A. A., Sergeev M. B. On two predictors of calculable chains of quasi-orthogonal matrices. *Automatic Control and Computer Sciences*, 2015, vol. 49, no. 3, pp. 153–158. doi:10.3103/S0146411615030025
 21. Балонин Н. А., Сергеев М. Б., Себерри Дж., Сеницына О. И. Окружности на решетках и матрицы максимума детерминанта. *Информационно-управляющие системы*, 2020, № 6, с. 2–11. doi:10.31799/1684-8853-2020-6-2-11

UDC 519.614

doi:10.31799/1684-8853-2021-5-2-9

Matrix vitrages and regular Hadamard matrices

A. A. Vostrikov^a, PhD, Tech., Associate Professor, orcid.org/0000-0002-8513-3683, vostricov@mail.ru

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: The Kronecker product of Hadamard matrices when a matrix of order n replaces each element in another matrix of order m , inheriting the sign of the replaced element, is a basis for obtaining orthogonal matrices of order nm . The matrix insertion operation when not only signs but also structural elements (ornamental patterns of matrix portraits) are inherited provides a more general result called a “vitrage”. Vitrages based on typical quasi-orthogonal Mersenne (M), Seidel (S) or Euler (E) matrices, in addition to inheriting the sign and pattern, inherit the value of elements other than unity (in amplitude) in a different way, causing the need to revise and systematize the accumulated experience. **Purpose:** To describe new algorithms for generalized product of matrices, highlighting the constructions that produce regular high-order Hadamard matrices. **Results:** We have proposed an algorithm for obtaining matrix vitrages by inserting Mersenne matrices into Seidel matrices, which makes it possible to expand the additive chains of matrices of the form M–E–M–E–... and S–E–M–E–..., obtained by doubling the orders and adding an edge. The operation of forming a matrix vitrage allows you to obtain matrices of high orders, keeping the ornamental pattern as an important invariant of the structure. We have shown that the formation of a matrix vitrage inherits the logic of the Scarpi product, but is cannot be reduced to it, since a nonzero distance in order between the multiplicands M and S simplifies the final regular matrix ornamental pattern due to the absence of cyclic displacements. The alternation of M and S matrices allows you to extend the multiplicative chains up to the known gaps in the S matrices. This sheds a new light on the theory of a regular Hadamard matrix as a product of Mersenne and Seidel matrices. **Practical relevance:** Orthogonal sequences with floating levels and efficient algorithms for finding regular Hadamard matrices with certain useful properties are of direct practical importance for the problems of noise-proof coding, compression and masking of video data.

Keywords — orthogonal matrices, regular matrices, Hadamard matrices, Mersenne matrices, Seidel matrices, Kronecker product, Scarpi algorithm, matrix inserts, matrix vitrages.

For citation: Vostrikov A. A. Matrix vitrages and regular Hadamard matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 2–9 (In Russian). doi:10.31799/1684-8853-2021-5-2-9

References

1. Wang R. *Introduction to Orthogonal Transforms with Applications in Data Processing and Analysis*. Cambridge University Press, 2010. 504 p.
2. Mironovskij L. A., Slaev V. A. *Strip-Method for Image and Signal Transformation*. De Gruyter, 2011. 175 p.
3. Vostrikov A., Sergeev M. Expansion of the quasi-orthogonal basis to mask images. *Smart Innovation, Systems and Technologies*, 2015, vol. 40, pp. 161–168. doi:10.1007/978-3-319-19830-9_15
4. Ahmed N., Rao R. *Orthogonal Transforms for Digital Signal Processing*. Berlin-Heidelberg-New York: Springer-Verlag, 1975. 263 p.
5. Koukouvinos C., Simos D., Varbanov Z. Hadamard matrices, designs and their secret-sharing schemes. *Algebraic Informatics*, 2011, pp. 216–229. doi:10.1007/978-3-642-21493-6_14
6. Hadamard J. Résolution d'une Question Relative aux Déterminants. *Bulletin des Sciences Mathématiques*, 1893, vol. 17, pp. 240–246 (In French).
7. Sergeev A. M. On interconnection of one type of quasi-orthogonal matrices constructed by orders of $4k$ and $4k-1$ sequences. *Izvestia LETI*, 2017, no. 7, pp. 12–17 (In Russian).
8. Di Matteo O., Djokovic D. Z., Kotsireas I. S. Symmetric Hadamard matrices of order 116 and 172 exist. *Special Matrices*, 2015, vol. 3.1, pp. 227–234.
9. Kharaghani H. and Tayfeh-Rezaie B. A Hadamard matrix of order 428. *J. Combin. Designs*, 2005, vol. 13, pp. 435–440.
10. Horadam K. J. Hadamard matrices and their applications: Progress 2007–2010. *Cryptography and Communications*, 2010, no. 2, iss. 2, pp. 129–154.
11. Seberry J., Yamada M. *Hadamard Matrices: Constructions using Number Theory and Linear Algebra*. John Wiley & Sons, 2020. 352 p.
12. Vostrikov A., Sergeev M., Balonin N., Sergeev A. Use of symmetric Hadamard and Mersenne matrices in digital image processing. *Procedia Computer Science*, 2018, pp. 1054–1061. doi:10.1016/j.procs.2018.08.042
13. Belevitch V. Theorem of $2n$ -terminal networks with application to conference telephony. *Electrical Communication*, 1950, no. 26, pp. 231–244.
14. Balonin Yu. N., Sergeev A. M. Two-circulant Hadamard matrices, weighing matrices, and Ryser's conjecture. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 2–9. doi:10.15217/issn1684-8853.2018.3.2
15. Sergeev A., Sergeev M., Vostrikov A., Kurtyanik D. Portraits of orthogonal matrices as a base for discrete textile ornament patterns. *Smart Innovation, Systems and Technologies*, 2019, vol. 143, pp. 135–143. doi:10.1007/978-981-13-8303-8_12
16. Scarpis U. Sui determinanti di valore massimo, Rendiconti della R. Istituto Lombardo di Scienze e Lettere. 1898, vol. 31, pp. 1441–1446 (In Italian).
17. Sergeev A. M. Generalized Mersenne matrices and Balonin's hypothesis. *Automatic Control and Computer Sciences*, 2014, no. 4, pp. 35–43 (In Russian).
18. Djokovic D. Generalization of Scarpis' theorem on Hadamard matrices. *Linear and Multilinear Algebra*, 2017, vol. 65, iss. 10, pp. 1985–1987.
19. Balonin Yu. N., Sergeev A. M., Sinicyna O. I. Algorithms for finite fields and search groups for orthogonal sequences. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 4, pp. 2–17 (In Russian). doi:10.31799/1684-8853-2021-6-2-17
20. Balonin N. A., Vostrikov A. A., Sergeev M. B. On two predictors of calculable chains of quasi-orthogonal matrices. *Automatic Control and Computer Sciences*, 2015, vol. 49, no. 3, pp. 153–158. doi:10.3103/S0146411615030025
21. Balonin N. A., Sergeev M. B., Seberry J., Sinitsyna O. I. Circles on lattices, and maximum determinant matrices. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2020, no. 6, pp. 2–11 (In Russian). doi:10.31799/1684-8853-2020-6-2-11

УВАЖАЕМЫЕ АВТОРЫ!

Научные базы данных, включая SCOPUS и Web of Science, обрабатывают данные автоматически. С одной стороны, это ускоряет процесс обработки данных, с другой — различия в транслитерации ФИО, неточные данные о месте работы, области научного знания и т. д. приводят к тому, что в базах оказывается несколько авторских страниц для одного и того же человека. В результате для всех по отдельности считаются индексы цитирования, что снижает рейтинг ученого.

Для идентификации авторов в сетях Thomson Reuters проводит регистрацию с присвоением уникального индекса (ID) для каждого из авторов научных публикаций.

Процедура получения ID бесплатна и очень проста, есть возможность провести регистрацию на 12-ти языках, включая русский (чтобы выбрать язык, кликните на зеленое поле сверху справа на стартовой странице): <https://orcid.org>

UDC 004.93

doi:10.31799/1684-8853-2021-5-10-19

Combined total variation of first and fractional orders for Poisson noise removal in digital images

Pham Cong Thang^a, PhD, Lecturer, orcid.org/0000-0002-6428-102X, pcthang@dut.udn.vn

Tran Thi Thu Thao^b, MSc., Lecturer, orcid.org/0000-0001-7705-2405

Pham Minh Tuan^a, PhD, Lecturer, orcid.org/0000-0001-9843-9676

Nguyen Thanh Cong^a, MSc., Specialist, orcid.org/0000-0002-8060-0238

^aThe University of Danang – University of Science and Technology, 54, Nguyen Luong Bang St., 550000, Danang, Vietnam

^bThe University of Danang – University of Economics, 71, Ngu Hanh Son St., 550000, Danang, Vietnam

Introduction: Many methods have been proposed to handle the image restoration problem with Poisson noise. A popular approach to Poissonian image reconstruction is the one based on total variation. This method can provide significantly sharp edges and visually fine images, but it results in piecewise-constant regions in the resulting images. **Purpose:** Developing an adaptive total variation-based model for the reconstruction of images contaminated by Poisson noise, and an algorithm for solving the optimization problem. **Results:** We proposed an effective way to restore images degraded by Poisson noise. Using the Bayesian framework, we proposed an adaptive model based on a combination of first-order total variation and fractional order total variation. The first-order total variation model is efficient for suppressing the noise and preserving the keen edges simultaneously. However, the first-order total variation method usually causes artifact problems in the obtained results. To avoid this drawback, we can use high-order total variation models, one of which is the fractional-order total variation-based model for image restoration. In the fractional-order total variation model, the derivatives have an order greater than or equal to one. It leads to the convenience of computation with a compact discrete form. However, methods based on the fractional-order total variation may cause image blurring. Thus, the proposed model incorporates the advantages of two total variation regularization models, having a significant effect on the edge-preserving image restoration. In order to solve the considered optimization problem, the Split Bregman method is used. Experimental results are provided, demonstrating the effectiveness of the proposed method. **Practical relevance:** The proposed method allows you to restore Poissonian images preserving their edges. The presented numerical simulation demonstrates the competitive performance of the model proposed for image reconstruction. **Discussion:** From the experimental results, we can see that the proposed algorithm is effective in suppressing noise and preserving the image edges. However, the weighted parameters in the proposed model were not automatically selected at each iteration of the proposed algorithm. This requires additional research.

Keywords – total variation, image restoration, Poisson noise, minimization method.

For citation: Pham C. T., Tran T. T., Pham M. T., Nguyen T. C. Combined total variation of first and fractional orders for Poisson noise removal in digital images. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 10–19. doi:10.31799/1684-8853-2021-5-10-19

Introduction

Poisson noise is known as photon noise or shot noise [1]. Poisson noise removal is an important task in various applications such as electronic microscopy [2, 3], tomography [4, 5], X-ray [6, 7], etc. In electronic microscopy imaging, the number of electrons collected to create an image pixel follows the Poisson distribution [8]. It means that parts of an image with higher intensity have a higher noise level than parts with low intensity. A low signal-to-noise ratio leads to a high level of noise, which affects detected edges and accuracy in measurement [9, 10]. With the electron tomography technique, data acquisition is the detection of scattered electron particles. The emission and subsequent detection of electrons follow a Poisson process [11]. These processes degrade data quality due to Poisson noise [12, 13]. Meanwhile, X-ray is a popular medical imaging technique. X-rays are produced using photons, which have very small wavelengths and high

penetration ability [14]. X-ray images are produced using a digital receptor. The photon scattering follows Poisson noise distribution which leads to degradation of the X-ray image quality and to the difficulty of diagnosis [15].

For many years, many variational methods have been proposed to handle the restoration problem with Poisson noise [16–23]. One of variational models for Poissonian image reconstruction is the one based on the total variation (TV) norm as regularization term [24]:

$$\min_{z \in \Omega} (\|z\|_{TV} + \beta \langle \mathbf{1}, Kz - f \log Kz \rangle), \quad (1)$$

where β is a positive regularization parameter; z must be positive in Ω ; K is a blurring operator, the operator $\|z\|_{TV}$ stands for the TV of u . Authors in [25] replaced the TV regularization term in (1)

by $\|z\|_{TV} + \frac{\lambda}{2} \|z\|_2^2$ and proposed bounded varia-

tion-based Poissonian images restoration model as follows (TV-model):

$$z^* = \arg \min_{z \in \Omega} \left(\|z\|_{TV} + \frac{\lambda}{2} \|z\|_2^2 + \beta \langle 1, Kz - f \log Kz \rangle \right). \quad (2)$$

The model (2) performs very well for preserving edges while removing noise. However, it often causes undesired artifact effects in smooth regions. To overcome these effects, some high-order models have been introduced for restoring blurred images corrupted by Poisson noise. The authors in [26] replaced the term $\|z\|_{TV}$ in (2) with higher-order $\|z\|_{HTV}$ and proposed following model (HTV-model):

$$z^* = \arg \min_{z \in \Omega} \left(\|z\|_{HTV} + \frac{\lambda}{2} \|z\|_2^2 + \beta \langle 1, z - f \log z \rangle \right). \quad (3)$$

Recently, fractional-order derivatives are widely applied in image processing [27–31]. The works have reflected the good performance of the fractional-order derivative in image denoising with edge-preserving. Following [29], the fractional-order TV model for Poissonian image denoising and deblurring is as follows (FTV-model):

$$z^* = \arg \min_u (\|z\|_{TV_\alpha} + \beta \langle 1, Kz - f \log Kz \rangle), \quad (4)$$

where β is positive parameter; $\|z\|_{TV_\alpha}$ stands for fractional-order TV defined in Eq. (10).

Motivated by the above studies, we introduce an adaptive TV based optimization problem as follows:

$$z^* = \arg \min_{z \in S(\Omega)} E(z), \quad (5)$$

$$E(z) = \mu_1 \|z\|_{TV} + \mu_2 \|z\|_{TV_\alpha} + \frac{\lambda}{2} \|z\|_2^2 + \beta \langle 1, Kz - f \log Kz \rangle,$$

where μ_1 , μ_2 and β are positive parameters, $\mu_1, \mu_2 \in (0, 1)$.

Motivated by the previous works, we propose a Poisson noise removal model that can substantially reduce artifact effects while preserving edges in the restored images. The proposed model is designed by combining advantages of the first order TV and fractional order TV. We extend the split Bregman method for solving the optimization problem. Furthermore, we provide experimental results to demonstrate the efficiency of our algorithm for the considered problem, in comparison with state-of-the-art methods.

Preliminaries

We recall the principle behind Eq. (5). We aim at reconstructing the original image u with the known noisy image f . Our strategy is to find the image u which maximizes the conditional probability $P(u|f)$. Bayes's rule gives

$$P(z|f) = \frac{P(f|z)P(z)}{P(f)}. \quad (6)$$

The probability density function of the observed image f corrupted by Poisson noise is:

$$P(f|z) = \frac{z^f \exp(-z)}{f!}.$$

Suppose that f has size $M \times N$, and let $\mathcal{I} = \{1, \dots, M\} \times \{1, \dots, N\}$ denote the domain of f . For $i \in \mathcal{I}$, we write f_i the pixel of f at position i (and similarly u_i the pixel of u at position i) [32]. Then:

$$P(f|z) = \prod_{i \in \mathcal{I}} \frac{(z_i)^{f_i} e^{-z_i}}{f_i!}.$$

Maximizing $P(z|f)$ is equivalent to minimizing $-\log(P(z|f))$, so let us compute the quantity $-\log(P(f|z))$:

$$\sum_{i \in \mathcal{I}} z_i - f_i \log(z_i) + \log(f_i!). \quad (7)$$

Since f is constant, we can ignore the term $\log(f_i!)$. Now we assume that $P(z)$ follows a choice of the prior:

$$P(z) = \exp\left(-\frac{1}{\tau} \phi(z)\right), \quad (8)$$

where τ is a normalization factor being positive and constant.

In this work, we assume that

$$\phi(z) = \mu_1 \|z\|_{TV} + \mu_2 \|z\|_{TV_\alpha} + \frac{\alpha}{2} \|z\|_2^2.$$

The assumption on $P(u)$ means that each pixel depends (weakly) on the neighbouring pixels only, so we do not lose much by assuming independence.

We now have all the ingredients to maximize $P(z|f)$. By Eq. (6), this amounts to minimize the expression $-\log(P(f|z)) - \log(P(z))$, so we can plug in Equations (7) and (8) to get:

$$z^* = \arg \min_z \sum_{i \in \mathcal{I}} \left(\frac{1}{\tau} \phi(z_i) + (z_i - f_i \log(z_i)) \right)$$

and we can view this expression as a discrete approximation of the functional $E(\cdot)$ defined as

$$E(z) = \left(\mu_1 \|z\|_{TV} + \mu_2 \|z\|_{TV_\alpha} + \frac{\alpha}{2} \|z\|_2^2 + \beta \langle 1, z - f \log z \rangle \right),$$

where $\beta = \tau$ is positive and constant parameter.

In case of the blur effect, we can generalize the model (5) for restoring a blurred image corrupted by Poisson noise as follows:

$$E(z) = \left(\mu_1 \|z\|_{TV} + \mu_2 \|z\|_{TV_\alpha} + \frac{\alpha}{2} \|z\|_2^2 + \beta \langle 1, Kz - f \log Kz \rangle \right),$$

where β is positive and constant parameter.

The discrete gradients of an image u for the pixel location (i, j) in z ($i = 1..M; j = 1..N$) are defined like [33–35]:

$$\|z\|_{TV} = \sqrt{(\nabla_1 z)^2 + (\nabla_2 z)^2}, \quad (9)$$

$$\nabla_1 z_{i,j} = z_{i+1,j} - z_{i,j}, \quad \nabla_2 z_{i,j} = z_{i,j+1} - z_{i,j},$$

$$\nabla_{11} z_{i,j} = z_{i+1,j} - 2z_{i,j} + z_{i-1,j},$$

$$\nabla_{22} z_{i,j} = z_{i,j+1} - 2z_{i,j} + z_{i,j-1},$$

$$\nabla_{12} z_{i,j} = \nabla_{21} z_{i,j} = z_{i,j} - z_{i,j-1} - z_{i-1,j} + z_{i-1,j-1},$$

$$\|z\|_{HTV} = \sqrt{(\nabla_{11} z)^2 + (\nabla_{12} z)^2 + (\nabla_{21} z)^2 + (\nabla_{22} z)^2}.$$

Due to the convenience in numerical implementation, from Greunwald — Letnikov (GL) fractional-order derivative, the discrete fractional-order TV $\|z\|_{TV_\alpha}$ of $z \in \Omega$ is defined as follows [27, 31]:

$$\|z\|_{TV_\alpha} = \sqrt{(\nabla_1^\alpha z)^2 + (\nabla_2^\alpha z)^2}, \quad (10)$$

where the discrete gradients $\nabla_1^\alpha z$ and $\nabla_2^\alpha z$ are defined as follows:

$$(\nabla_1^\alpha z)_{i,j} = \sum_{k=0}^{L-1} C_k^\alpha z_{i-k,j}, \quad (\nabla_2^\alpha z)_{i,j} = \sum_{k=0}^{L-1} C_k^\alpha z_{i,j-k}.$$

Parameter L is the number of neighboring pixels that are used to compute the fractional-order derivative at each pixel; the coefficients $\{C_k^\alpha\}$ are defined as follows [27]:

$$C_k^\alpha = \frac{\Gamma(\alpha+1)}{\Gamma(k+1)\Gamma(\alpha+1-k)}, \quad k = 1..L-1.$$

Computational method

In this section, we derive the numerical method for the problem (5) in detail. There are many meth-

ods which can be employed to obtain the solution of the optimization problem. In this article, we decide to employ the split Bregman method for solving the optimization problem.

The split Bregman method performs break the minimization problem down into easy subproblems [36–38]. Subproblems can be directly solved with tools like fast Fourier transform (FFT), shrinkage operator that makes the optimization algorithm rather fast. We have a scalar γ and two convex functionals $\Psi(\cdot)$ and $G(\cdot)$; and that we need to solve the following constrained optimization problem:

$$\arg \min_{z,w} (\|w\| + \gamma G(z)), \quad \text{s.t. } w = \Psi(z). \quad (11)$$

We convert (11) into an unconstrained problem:

$$\arg \min_{z,w} \left(\|w\| + \gamma G(z) + \frac{\rho}{2} \|d - \Psi(z) - b\|_2^2 \right), \quad (12)$$

where w — splitting variable; ρ is positive constant parameters; b — variable of the Bregman iterations.

The solution to problem (12) can be approximated by the Split Bregman method:

$$z^{(k+1)} = \arg \min_z \left(\gamma G(z) + \frac{\rho}{2} \|w^{(k)} - \Psi(z) - b^{(k)}\|_2^2 \right);$$

$$w^{(k+1)} = \arg \min_w \left(\|w\|_1 + \frac{\rho}{2} \|w - \Psi(z^{(k+1)}) - b^{(k)}\|_2^2 \right),$$

$$b^{(k+1)} = b^{(k)} + \Psi(z^{(k+1)}) - w^{(k+1)}.$$

We return to the problem (5). By introducing three auxiliary variables u, p and q , Eq. (5) is equivalent to the constrained optimization problem:

$$\min_{p,q,z,u} \left(\mu_1 \|p\| + \mu_2 \|q\| + \frac{\theta}{2} \|z\|_2^2 + \beta \langle 1, u - f \log u \rangle \right),$$

$$\text{s.t. } p = \nabla z, \quad q = \nabla^\alpha z, \quad u = Kz, \quad (13)$$

where ∇z and $\nabla^\alpha z$ are defined in (9) and (10) respectively.

We convert the problem (5) to the unconstrained one as follows:

$$\min_{p,q,z,u} \left(\mu_1 \|p\| + \mu_2 \|q\| + \frac{\theta}{2} \|z\|_2^2 + \beta \langle 1, u - f \log u \rangle + \frac{\rho_1}{2} \|p - \nabla z\|_2^2 + \frac{\rho_2}{2} \|q - \nabla^\alpha z\|_2^2 + \frac{\rho_3}{2} \|u - Kz\|_2^2 \right), \quad (14)$$

where μ_1, μ_2, β and γ are positive parameters; ρ_1, ρ_2 are Lagrangian multipliers.

The extended split Bregman iterative method employed for solving the problem (5) can be described as follows:

$$\begin{cases} z^{(k+1)} = \arg \min_z \left(\begin{aligned} & \frac{\theta}{2} \|z\|_2^2 + \frac{\rho_1}{2} \|p^{(k)} - \nabla z - b_p^{(k)}\|_2^2 + \\ & + \frac{\rho_2}{2} \|q^{(k)} - \nabla^\alpha z - b_q^{(k)}\|_2^2 + \\ & + \frac{\rho_3}{2} \|u^{(k)} - Kz - b_z^{(k)}\|_2^2 \end{aligned} \right) \\ u^{(k+1)} = \arg \min_u \left(\begin{aligned} & \beta \langle \mathbf{1}, u - f \log u \rangle + \\ & + \frac{\rho_3}{2} \|u - Kz^{(k+1)} - b_z^{(k)}\|_2^2 \end{aligned} \right) \\ p^{(k+1)} = \arg \min_p \left(\mu_1 \|p\| + \frac{\rho_1}{2} \|p - \nabla z^{(k+1)} - b_p^{(k)}\|_2^2 \right) \\ q^{(k+1)} = \arg \min_q \left(\mu_2 \|q\| + \frac{\rho_2}{2} \|q - \nabla^\alpha z^{(k+1)} - b_q^{(k)}\|_2^2 \right) \end{cases}$$

with update for $b_p^{(k+1)}$, $b_q^{(k+1)}$, $b_z^{(k+1)}$:

$$\begin{cases} b_p^{(k+1)} = b_p^{(k)} + \nabla z^{(k+1)} - p^{(k+1)} \\ b_q^{(k+1)} = b_q^{(k)} + \nabla^\alpha z^{(k+1)} - q^{(k+1)} \\ b_z^{(k+1)} = b_z^{(k)} + Kz^{(k+1)} - u^{(k+1)} \end{cases}$$

Hence, we have three subproblems to solve: z , p and q .

Subproblem z . For the z subproblem, optimality condition reads:

$$\begin{aligned} & \theta z + \rho_1 (\nabla)^T (\nabla z + b_p^{(k)} - p^{(k)}) + \\ & + \rho_2 (\nabla^\alpha)^T (\nabla^\alpha z + b_q^{(k)} - q^{(k)}) = 0. \end{aligned}$$

Therefore, we have

$$\begin{aligned} & (\theta + \rho_1 (\nabla)^T \nabla + \rho_2 (\nabla^\alpha)^T \nabla^\alpha) z = \\ & = \rho_1 (\nabla)^T (p^{(k)} - b_p^{(k)}) + \rho_2 (\nabla^\alpha)^T (q^{(k)} - b_q^{(k)}). \end{aligned} \quad (15)$$

The Eq. (15) can be solved efficiently with one Fourier transform operation and one inverse FFT operation as follows:

$$z^{(k+1)} = F^{-1} \left(\frac{F(\rho_1 (\nabla)^T (p^{(k)} - b_p^{(k)}) + \rho_2 (\nabla^\alpha)^T (q^{(k)} - b_q^{(k)}))}{\theta + \rho_1 F((\nabla)^T \nabla) + \rho_2 F((\nabla^\alpha)^T \nabla^\alpha)} \right), \quad (16)$$

where F and F^{-1} are the forward and inverse Fourier transform operators.

Subproblem u . For the u subproblem, optimality condition reads:

$$\beta \frac{u-f}{u} + \rho_3 (u - Kz^{(k+1)} - b_z^{(k)}) = 0.$$

At the $(k+1)$ -th iteration, we compute u by discretization scheme:

$$\beta \frac{u^{(k+1)} - f}{u^{(k)}} + \rho_3 (u^{(k+1)} - Kz^{(k+1)} - b_z^{(k)}) = 0.$$

Therefore, we have

$$u^{(k+1)} = \frac{\beta f + \rho_3 u^{(k)} (Kz^{(k+1)} + b_z^{(k)})}{\beta + \rho_3 u^{(k)}}. \quad (17)$$

Subproblems p and q . The solution of the p subproblem can readily be obtained by applying the soft thresholding operator:

$$\begin{aligned} p^{(k+1)} &= \mathit{shrink} \left(\nabla z^{(k+1)} + b_p^{(k)}, \frac{\mu_1}{\rho_2} \right) = \\ &= \frac{\nabla z^{(k+1)} + b_p^{(k)}}{|\nabla z^{(k+1)} + b_p^{(k)}|} \cdot \max \left(|\nabla z^{(k+1)} + b_p^{(k)}| - \frac{\mu_1}{\rho_2}, 0 \right). \end{aligned} \quad (18)$$

The solution of the q subproblem can also be obtained by applying the soft thresholding operator:

$$\begin{aligned} q^{(k+1)} &= \mathit{shrink} \left(\nabla^\alpha u^{(k+1)} + b_q^{(k)}, \frac{\mu_2}{\rho_2} \right) = \\ &= \frac{\nabla^\alpha u^{(k+1)} + b_q^{(k)}}{|\nabla^\alpha u^{(k+1)} + b_q^{(k)}|} \cdot \max \left(|\nabla^\alpha u^{(k+1)} + b_q^{(k)}| - \frac{\mu_2}{\rho_2}, 0 \right). \end{aligned} \quad (19)$$

The complete method is summarized in Algorithm 1.

Algorithm 1: Adaptive split Bregman method for solving the problem (5)

1. **Initialize:** $z^{(0)} = f$; $p^{(0)} = q^{(0)} = 0$; $b_p^{(0)} = b_q^{(0)} = b_z^{(0)}$; $k = 1$

2. **while** $\left(\frac{\|u^{(k)} - u^{(k-1)}\|_2}{\|u^{(k)}\|_2} < \varepsilon \right) \parallel (k \leq N_{\max})$ **do**

3. Calculate $z^{(k+1)}$ using (16)

4. Calculate $u^{(k+1)}$ using (17)

5. Calculate $p^{(k+1)}$ using (18)

6. Calculate $q^{(k+1)}$ using (19)

7. $b_p^{(k+1)} = b_p^{(k)} + \nabla z^{(k+1)} - p^{(k+1)}$

8. $b_q^{(k+1)} = b_q^{(k)} + \nabla^\alpha z^{(k+1)} - q^{(k+1)}$

9. $b_z^{(k+1)} = b_z^{(k)} + Kz^{(k+1)} - u^{(k+1)}$
10. $k = k + 1$
11. endwhile
12. return z

Experimental results

In this section, we present some numerical results to illustrate the performance of the proposed model for Poisson noise removal. In order to prove the efficiency of the proposed model, we compare our reconstruction results with those of the mentioned models: TV-model, HTV-model and FTV-model. The compared models are implemented by the split Bregman method. We performed all experiments under MATLAB and Windows 10 on a PC with an Intel Core (TM) i5 CPU at 2.4 GHz and 8 GB of RAM. Empirically, all images are processed with the equivalent parameters $\alpha = 1.5$, $\mu_1 = 0.6$, $\mu_2 = 0.4$, $\rho_1 = 0.01$, $\rho_2 = 0.01$, $\rho_3 = 0.01$, $\theta = 0.0001$. We set the stopping condition for Algorithm 1: $\varepsilon = 0.00004$ and $N = 500$. The observed images in our experiments are simulated as follows. Poisson noise is data dependent, the noise level of the observed images depends on the pixel intensity value. To test different noise levels, the noisy image is

simulated by adding Poisson with some fixed value *Peak*. The test images are shown in Fig. 1.

The peak signal-to-noise ratio (PSNR) used in comparison are defined as follows:

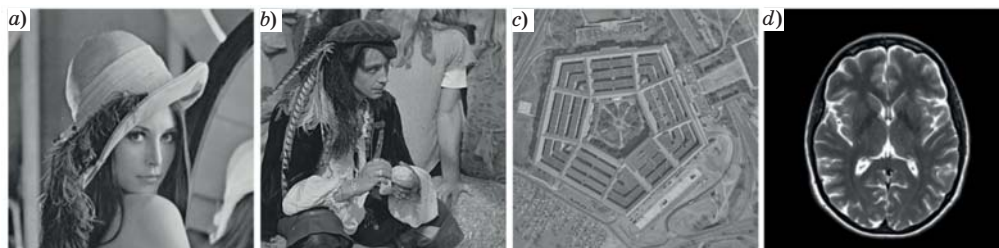
$$PSNR = 10 \log_{10} \left(\frac{255^2 \cdot MN}{\|u^* - u\|_2^2} \right),$$

where M and N are the number of image pixels in rows and columns; u, u^* are the original image, the reconstructed or noisy image accordingly. We also use other popular measure called structural similarity index measure (SSIM) which allows us to get more consistent with human visual characteristics [39]:

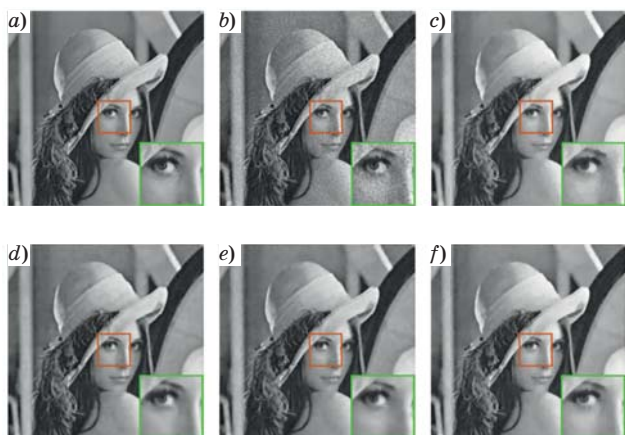
$$SSIM(u, u^*) = \frac{(2\mu_u \mu_{u^*} + c_1)(2\sigma_{u,u^*} + c_2)}{(\mu_u^2 + \mu_{u^*}^2 + c_1)(\sigma_u^2 + \sigma_{u^*}^2 + c_2)},$$

where μ_u, μ_{u^*} are the means of u, u^* respectively; σ_u, σ_{u^*} — their standard deviations; σ_{u,u^*} — the covariance of two images u and u^* ; c_1, c_2 are positive constants.

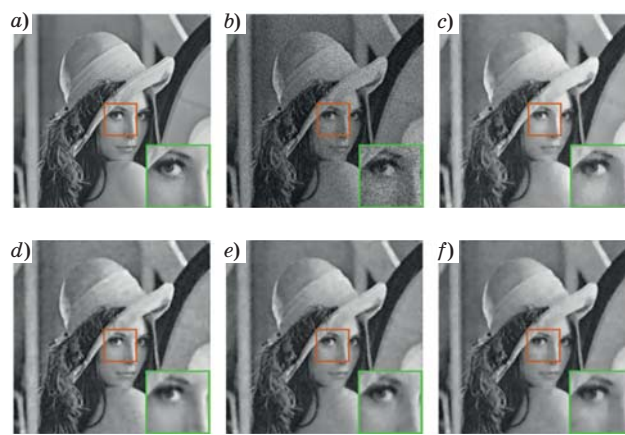
We first deal with Image denoising. In this case, K is an identity matrix. In Figs. 2 and 3, we aim



■ Fig. 1. Test images: a — Lena; b — Man; c — Aerial; d — MRI



■ Fig. 2. Image “Lena”. Recovered images of different methods for image denoising with *Peak* = 255: a — original image; b — noisy image; c — TV; d — HTV; e — FTV; f — ours



■ Fig. 3. Image “Lena”. Recovered images of different methods for image denoising with *Peak* = 100: a — original image; b — noisy image; c — TV; d — HTV; e — FTV; f — ours

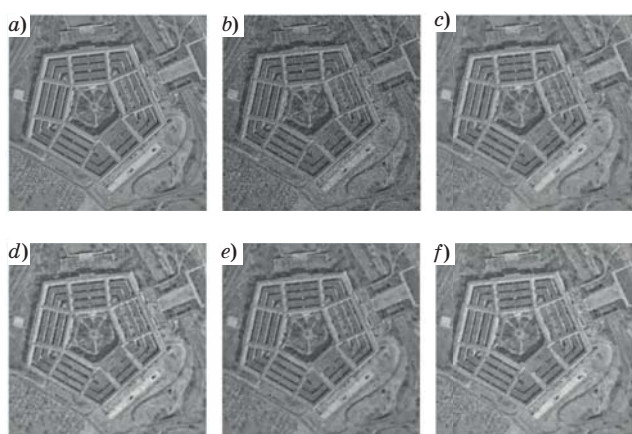
to restore the image “Lena” corrupted by Poisson noise with $Peak = 255$ and $Peak = 100$, respectively. In these figures, we also present the zoom-in on small details of the recovered images.

We see that TV-model is efficient for removing noise and simultaneously preserving the edges. However, the numerous artifact exists in result image recovered by TV-model. Meanwhile, the HTV-model and FTV-model can lead to edge blurring when image denoising. The images recovered by our model are more smooth and distinct than those of another three approaches.

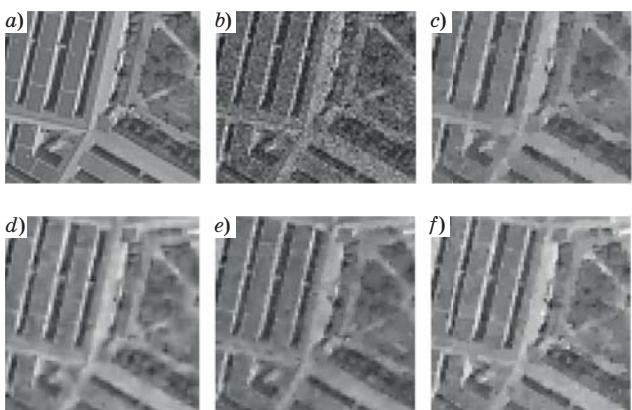
In Figs. 4, 6 and 8, we show the results of compared methods for noise levels $Peak = 100$. In Figs. 4, *b*, 6, *b*, and 8, *b*, we represent the noisy images. In the others, Figs. 4, *c-f*, 6, *c-f*, 8, *c-f*, we show respectively the reconstructions given by TV, HTV, FTV and our proposed approach. In Figs. 5, 7 and 9, we show the zoomed details of the original images, observed images and the zoomed details of the re-

stored images respectively in Figs. 4, 6 and 8. From the details in Figs. 5, 7 and 9, we can see that the our model can get better visual improvement than the others. In Tables 1 and 2, we show the comparison results in terms of SSIM and PSNR (the best result is highlighted in bold). We can clearly see that our method outperforms the other relative methods for restoring images damaged by Poisson noise.

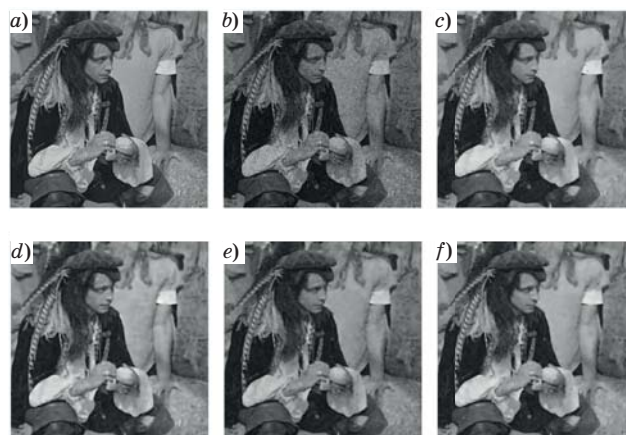
In the case of image deblurring and denoising, we consider blurred images degraded by Poisson noise. For simulation, we use the Gaussian blur with a window size 5×5 and standard deviation of 1. After the blurring operation, we degrade the images by Poisson noise with $Peak = 100$. In Fig. 10, we perform simultaneously image deblurring and denoising on image “Lena”. Fig. 10, *b* denotes corrupted image. In Fig. 10, *c-f*, we show respectively the reconstructions given by TV, HTV, FTV and our approach.



■ Fig. 4. Image “Aerial”. Recovered images of different methods for image denoising with $Peak = 100$: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours



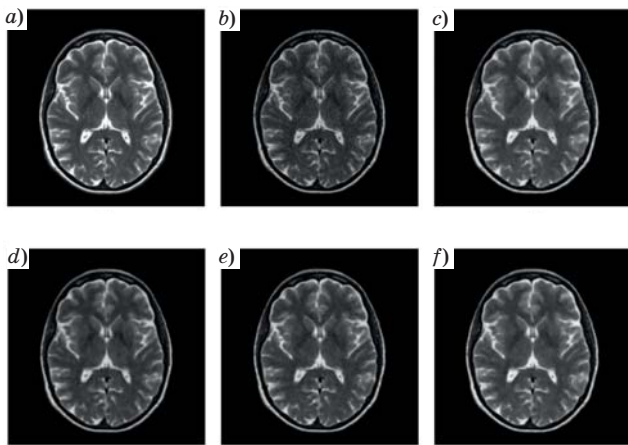
■ Fig. 5. Image “Aerial”. The zoomed-in details of the recovered images in Fig. 4: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours



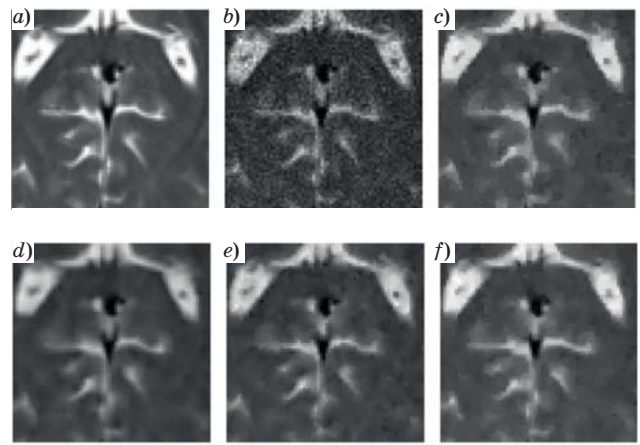
■ Fig. 6. Image “Man”. Recovered images of different methods for image denoising with $Peak = 100$: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours



■ Fig. 7. Image “Man”. The zoomed-in details of the recovered images in Fig. 6: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours



■ Fig. 8. Image “MRI”. Recovered images of different methods for image denoising with $Peak = 100$: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours



■ Fig. 9. Image “MRI”. The zoomed-in details of the recovered images in Fig. (8): *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours

■ Table 1. PSNR values for recovered images given by the compared methods with various noisy levels

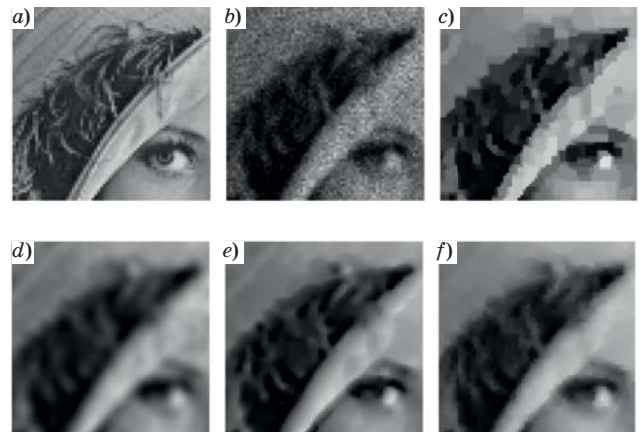
Noise level Peak	PSNR				
	Noisy	TV	HTV	FTV	Ours
Lena					
255	26.9200	31.1786	30.9261	31.0504	31.8640
100	22.8909	29.1995	28.8624	29.2134	29.4694
Man					
255	27.6436	29.6582	29.5528	29.8153	30.1140
100	23.5718	26.5633	26.2971	26.7400	26.8960
Aerial					
255	26.7701	28.4530	28.2417	28.4966	29.006
100	22.6521	25.8654	25.7517	25.6930	26.2966
MRI					
255	30.1130	30.7451	30.7111	30.8487	30.9878
100	26.9175	28.9371	28.9151	28.9780	29.2772

■ Table 2. SSIM values for recovered images given by the compared methods with various noisy levels

Noise level Peak	SSIM				
	Noisy	TV	HTV	FTV	Ours
Lena					
255	0.6721	0.8794	0.8812	0.8780	0.8891
100	0.5234	0.8368	0.8423	0.8411	0.8505
Man					
255	0.8028	0.8770	0.8731	0.8713	0.8835
100	0.6578	0.7830	0.7784	0.7862	0.7986
Aerial					
255	0.8277	0.8649	0.8585	0.8645	0.8805
100	0.7154	0.7787	0.7639	0.7647	0.7882
MRI					
255	0.9025	0.9401	0.9427	0.9344	0.9484
100	0.8416	0.9107	0.9114	0.9095	0.9272



■ Fig. 10. Image “Lena”. Recovered results for the test images: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours



■ Fig. 11. Image “Lena”. The zoomed-in details of the recovered images in Fig. 10: *a* — original image; *b* — noisy image; *c* — TV; *d* — HTV; *e* — FTV; *f* — ours

■ **Table 3.** PSNR values for recovering blurred images with noisy level $Peak = 100$

Image	PSNR				
	Noisy	TV	HTV	FTV	Ours
Lena	20.6275	22.7049	22.5994	22.8684	23.3133
Man	19.4880	20.5501	20.0390	20.4229	20.6916
Aerial	18.3362	18.7176	18.9762	18.8378	19.4537
MRI	24.3167	24.9252	24.231	24.6373	25.0166

Meanwhile, Fig. 11 shows enlarged images recovered by four methods in Fig. 10. In Tables 3 and 4, we report the quantitative measures of PSNR and SSIM values for different images and compared methods. The Figures and Tables demonstrate again the effectiveness of our proposed method for image reconstruction under Poisson noise even in presence of blur.

Conclusions

In this paper, we have researched the hybrid regularizers model, combining the fractional-order and first-order TV for denoising images corrupted by Poisson noise. Computationally, an extended split Bregman method is employed for solving the proposed optimization problem. Finally, compared with the existing state-of-the-art models, the ex-

■ **Table 4.** SSIM values for recovering blurred images with noisy level $Peak = 100$

Image	SSIM				
	Noisy	TV	HTV	FTV	Ours
Lena	0.4354	0.7503	0.7190	0.7647	0.7739
Man	0.4748	0.5702	0.5432	0.5853	0.5994
Aerial	0.5220	0.6047	0.5430	0.6074	0.6214
MRI	0.7956	0.8642	0.8527	0.8690	0.8760

periments demonstrate the efficiency of the proposed method.

Acknowledgments

Pham Cong Thang (corresponding author) would like to thank his colleagues at IT Faculty, DUT, for their helpful comments. The authors also thank reviewers and Editors for their insightful comments and suggestions.

Financial support

This work is supported by The Murata Science Foundation and The University of Danang, University of Science and Technology, code number of Project T2020-02-07MSF.

References

1. Hasinoff S. W. *Photon, poisson noise*. In: *Computer vision*. Ed. K. Ikeuchi. Boston, MA, Springer US, 2014. Pp. 608–610. doi:10.1007/978-0-387-31439-6_482
2. Zhang Y., Zhu Y., Nichols E., Wang Q., Zhang S., Smith C., Howard S. A Poisson — Gaussian denoising dataset with real fluorescence microscopy images. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 11710–11718.
3. Meinel W., Olivo-Marin J., Angelini E. D. Denoising of microscopy images: A review of the state-of-the-art, and a new sparsity-based method. *IEEE Transactions on Image Processing*, 2018, vol. 27, no. 8, pp. 3842–3856. doi:10.1109/TIP.2018.2819821
4. Bubba T. A., Porta F., Zanghirati G., Bonettini S. A nonsmooth regularization approach based on shearlets for Poisson noise removal in ROI tomography. *Applied Mathematics and Computation*, 2018, vol. 318, no. 1, pp. 131–152. doi:10.1016/j.amc.2017.09.001
5. Zeng G. L., Lv L., Huang Q. Poisson-noise weighted filter for time-of-flight positron emission tomography. *Visual Computing for Industry, Biomedicine and Art*, 2020, vol. 3, Article number: 10, 4 p. doi:10.1186/s42492-020-00048-8
6. Ofek E. O., Zackay B. Optimal matched filter in the low-number count Poisson noise regime and implications for X-ray source detection. *The Astronomical Journal*, 2018, vol. 155, no. 4, 8 p. doi:10.3847/1538-3881/aab265
7. Yagoub B., Ibrahim H., Salem A., Suh J. W., Kang H. S. X-ray image denoising for Cargo Dual Energy Inspection System. *International Conference on Electronics, Information, and Communication (ICEIC)*, 2021, pp. 1–4. doi:10.1109/ICEIC51217.2021.9369786
8. Pham C. T., Gamard G., Kopylov A., Tran T. T. T. An algorithm for image restoration with mixed noise using total variation regularization. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2018, vol. 26, no. 6, pp. 2831–2845. doi:10.3906/elk-1803-100
9. Kockentiedt S., Tonnies K. D., Gierke E., Nico D., Thim C., Plitzko S. Poisson shot noise parameter estimation from a single scanning electron microscopy image. *Proc. SPIE 8655, Image Processing: Algorithms and Systems XI*, 2013, 86550N. doi:10.1117/12.2008374
10. Seki T., Ikuhara Y., Shibata N. Theoretical framework of statistical noise in scanning transmission

- electron microscopy. *Ultramicroscopy*, 2018, vol. 193, pp. 118–125. doi:10.1016/j.ultramic.2018.06.014
11. Sanders T. The potential for Poisson image reconstruction models for electron tomography. *Proceedings of SPIE Volume 10772, Unconventional and Indirect Imaging, Image Reconstruction, and Wavefront Sensing*, 2018, 1077200. doi:10.1117/12.2321700
 12. Diwakar M., Kumar M. A review on CT image noise and its denoising. *Biomedical Signal Processing and Control*, 2018, vol. 42, pp. 73–88. doi:10.1016/j.bspc.2018.01.010
 13. Pham C. T., Tran T. T. T., Phan T. D. K., Dinh V. S., Pham M. T., Nguyen M. H. An adaptive algorithm for restoring image corrupted by mixed noise. *Cybernetics and Physics*, 2019, vol. 8, no. 2, pp. 73–82. doi:10.35470/2226-4116-2019-8-2-73-82
 14. Kirti T., Jitendra K., Ashok S. Poisson noise reduction from X-ray images by region classification and response median filtering. *Sadhana*, 2017, vol. 42, no. 6, pp. 855–863. doi:10.1007/s12046-017-0654-4
 15. Sun Y., Liu X., Cong P., Li L., Zhao Z. Digital radiography image denoising using a generative adversarial network. *Journal of X-Ray Science and Technology*, 2018, vol. 26, pp. 523–534. doi:10.3233/XST-17356 IOS Press
 16. Bal A., Banerjee M., Chaki R., Sharma P. An efficient method for PET image denoising by combining multi-scale transform and non-local means. *Multimedia Tools and Applications*, 2020, vol. 79, pp. 29087–29120. doi:10.1007/s11042-020-08936-0
 17. Chen H., Fu S., Wang H., Zhai L., Wang F., Li Y. Fuzzy singular value shrinkage for Poisson image denoising. *Journal of Modern Optics*, 2019, vol. 66, pp. 1569–1578. doi:10.1080/09500340.2019.1648895
 18. Kayyar S. H., Jidesh P. Non-local total variation regularization approach for image restoration under a Poisson degradation. *Journal of Modern Optics*, 2018, vol. 65, no. 19, pp. 2231–2242. doi:10.1080/09500340.2018.1506058
 19. Zhang M., Zhang F., Liu Q., Wang S. VST-Net: Variance-stabilizing transformation inspired network for Poisson denoising. *Journal of Visual Communication and Image Representation*, 2019, vol. 62, pp. 12–22. doi:10.1016/j.jvcir.2019.04.011
 20. Bindilatti A. A., Vieira M. A. C., Mascarenhas N. D. A. Poisson Wiener filtering with non-local weighted parameter estimation using stochastic distances. *Signal Processing*, 2018, vol. 144, pp. 68–76. doi:10.1016/j.sigpro.2017.10.001
 21. Niknejad M., Figueiredo M. A. T. Poisson image denoising using best linear prediction: a post-processing framework. *European Signal Processing Conference*, 2018, pp. 2230–2234. doi:10.23919/EUSIPCO.2018.8553220
 22. Bodduna K., and Weickert J. Poisson noise removal using multi-frame 3D block matching. *8th European Workshop on Visual Information Processing (EUVIP)*, 2019, pp. 58–63. doi:10.1109/EUVIP47703.2019.8946209
 23. Pham C. T., Tran T. T. T. An algorithm for hybrid regularizers based image restoration with Poisson noise. *Kybernetika*, 2021, vol. 57, no. 3, pp. 446–47. doi:10.14736/kyb-2021-3-0446
 24. Le T., Chartrand R., Asaki T. A variational approach to constructing images corrupted by Poisson noise. *Journal of Mathematical Imaging and Vision*, 2007, vol. 27, pp. 257–263. doi:10.1007/s10851-007-0652-y
 25. Liu X., Huang L. Total bounded variation-based Poissonian images recovery by split Bregman iteration. *Mathematical Methods in the Applied Sciences*, 2012, vol. 35, no. 5, pp. 520–529. doi:10.1002/mma.1588
 26. Zhang J., Ma M., Wu Z., Deng C. 2019. High-order total bounded variation model and its fast algorithm for Poissonian image restoration. *Mathematical Problems in Engineering*, 2019, vol. 2019, 11 p. doi:10.1155/2019/2502731
 27. Zhang J., Wei Z., Xiao L. Adaptive fractional-order multi-scale method for image denoising. *Journal of Mathematical Imaging and Vision*, 2012, vol. 43, pp. 39–49. doi:10.1007/s10851-011-0285-z
 28. Chen D., Chen Y. Q., Xue D. Fractional-order total variation image denoising based on proximity algorithm. *Applied Mathematics and Computation*, 2015, vol. 257, pp. 537–545. doi:10.1016/j.amc.2015.01.012
 29. Chowdhury M. R., Zhang J., Qin J., Lou Y. Poisson image denoising based on fractional-order total variation. *Inverse Problems and Imaging*, 2020, vol. 14, no. 1, pp. 77–96. doi:10.3934/ipi.2019064
 30. Zhao W., Lu H. Medical image fusion and denoising with alternating sequential filter and adaptive fractional order total variation. *IEEE Transactions on Instrumentation and Measurement*, 2017, vol. 66, no. 9, pp. 2283–2294. doi:10.1109/TIM.2017.2700198
 31. Huang G., Chen Q., Men T., Zhang X. Q., Qin H. Y., Xu L. Application in image denoising using fractional total variation theory. *The Proceedings of the International Conference on Sensing and Imaging*, 2017, pp. 211–221. doi:10.1007/978-3-319-91659-0_15
 32. Pham C. T., Tran T. T. T., Gamard G. An efficient total variation minimization method for image restoration. *Informatica*, 2020, vol. 31, no. 3, pp. 539–560. doi:10.15388/20-INFOR407
 33. Tran T. T. T., Pham, C. T., Kopylov A. V., Nguyen V. N. An adaptive variational model for medical images restoration. *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2019, vol. XLII-2/W12, pp. 219–224. doi:10.5194/isprs-archives-XLII-2-W12-219-2019
 34. Pham C. T., Tran T. T. T., Nguyen T. C., Vo D. H. Second-order total generalized variation based model for restoring images with mixed Poisson — Gaussian noise. *Informatsionno-upravliaiushchie sistemy [Information and Control Systems]*, 2021, no. 2, pp. 20–32. doi:10.31799/1684-8853-2021-2-20-32
 35. Pham C. T., Tran T. T. T., Vo D. H., Vo D. H. A hybrid regularizers model for multiplicative noise removal.

- Cybernetics and Physics*, 2021, vol. 10, no. 1, pp. 40–50. doi:10.35470/2226-4116-2021-10-1-40-50
36. Goldstein T., Osher S. The split Bregman method for L_1 -regularized problems. *SIAM Journal on Imaging Sciences*, 2009, vol. 2, no. 2, pp. 323–343. doi:10.1137/080725891
37. Zhang C. C., Wang Y., Xiao H. B. Applying improved split Bregman method to fluorescence microscopic image restoration. *Computer Science, Technology and Application*, 2016, pp. 54–60. doi:10.1142/9789813200449_0008
38. Yang Y., Tian D., Jia W., Shu X., Wu B. Split Bregman method based level set formulations for segmentation and correction with application to MR images and color images. *Magnetic Resonance Imaging*, 2019, vol. 57, pp. 50–67. doi:10.1016/j.mri.2018.10.005
39. Bovik A. C., Wang Z. *Modern Image Quality Assessment, Synthesis Lectures on Image, Video, and Multimedia Processing*. Morgan and Claypool Publishers, 2006. 156 p. doi:10.2200/S00010ED1V01Y-200508IVM003

УДК 004.93

doi:10.31799/1684-8853-2021-5-10-19

Комбинированная полная вариация первого и дробного порядка для удаления пуассоновского шума на цифровых изображениях

Фам Конг Тханг^а, PhD, преподаватель, orcid.org/0000-0002-6428-102X, pcthang@dut.udn.vnЧан Тхи Тху Тхао^б, магистр, преподаватель, orcid.org/0000-0001-7705-2405Фам Минь Туан^а, PhD, преподаватель, orcid.org/0000-0001-9843-9676Нгуен Тхань Конг^а, магистр, специалист, orcid.org/0000-0002-8060-0238^аУниверситет Дананга — Университет науки и техники, Нгуен Лыонг Банг, 54, Дананг, 550000, Вьетнам^бУниверситет Дананга — Университет экономики, Нгу Хань Шон, 71, Дананг, 550000, Вьетнам

Введение: известно множество методов для решения проблемы восстановления изображений с помощью пуассоновского шума. Одним из популярных подходов к реконструкции пуассоновского изображения является метод, основанный на полной вариации. С помощью этого метода можно получить весьма резкие границы и визуально четкие изображения, но он приводит к образованию кусочно-постоянных областей в результирующих изображениях. **Цель:** разработка адаптивной модели на основе полных вариаций для реконструкции изображений, искаженных пуассоновским шумом, и алгоритма решения задачи оптимизации. **Результаты:** предложен эффективный метод для восстановления изображений, искаженных пуассоновским шумом. На базе байесовской структуры предложена адаптивная модель, основанная на комбинации полной вариации первого порядка и полной вариации дробного порядка. Восстановление изображения на основе модели полной вариации первого порядка эффективно для шумоподавления и одновременно сохранения острых границ. Однако метод полной вариации первого порядка обычно вызывает проблемы с артефактами в полученных результатах. Чтобы избежать этого недостатка, использованы модели полной вариации высокого порядка, одна из которых является основанной на полной вариации дробного порядка для восстановления изображений. В модели полной вариации дробного порядка производные имеют порядок больше или равный единице. Это приводит к удобству вычислений с компактной дискретной формой. Но методы, основанные на полной вариации дробного порядка, могут вызывать размытие изображения. Таким образом, предложенная модель включает в себя преимущества двух моделей регуляризации полной вариации и существенно влияет на восстановление изображения с сохранением границ. Для решения рассматриваемой оптимизационной задачи используется метод Split Bregman. Приведены экспериментальные результаты, демонстрирующие эффективность предложенного метода. **Практическая значимость:** предлагаемый метод позволяет восстанавливать пуассоновские изображения с сохранением границ. Представленное численное моделирование наглядно демонстрирует конкурентоспособность предложенной модели для восстановления изображений, деградированных пуассоновским шумом. **Обсуждение:** из результатов экспериментов видно, что предложенный алгоритм эффективен в шумоподавлении и сохранении границ изображения. Однако взвешенные параметры, существующие в предложенной модели, не выбирались автоматически на каждой итерации предложенного алгоритма, что требует дополнительных исследований.

Ключевые слова — полная вариация, восстановление изображения, пуассоновский шум, метод минимизации.

Для цитирования: Pham C. T., Tran T. T. T., Pham M. T., Nguyen T. C. Combined total variation of first and fractional orders for Poisson noise removal in digital images. *Информационно-управляющие системы*, 2021, № 5, с. 10–19. doi:10.31799/1684-8853-2021-5-10-19

For citation: Pham C. T., Tran T. T. T., Pham M. T., Nguyen T. C. Combined total variation of first and fractional orders for Poisson noise removal in digital images. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 10–19. doi:10.31799/1684-8853-2021-5-10-19

Методологические подходы и алгоритмы распознавания и подсчета животных на аэрофотоснимках

В. В. Михайлов^а, доктор техн. наук, профессор, ведущий научный сотрудник, orcid.org/0000-0001-6973-3526, mwwcari@gmail.com

Л. А. Колпащиков^б, доктор биол. наук, начальник научного отдела, orcid.org/0000-0003-4871-3360

В. А. Соболевский^а, аспирант, младший научный сотрудник, orcid.org/0000-0001-7685-4991

Н. В. Соловьев^в, канд. техн. наук, доцент, orcid.org/0000-0001-8851-3713

Г. К. Якушев^в, бакалавр, orcid.org/0000-0003-4689-3425

^аСанкт-Петербургский федеральный исследовательский центр РАН, 14-я линия В. О., 39, Санкт-Петербург, 199178, РФ

^бОбъединенная дирекция заповедников Таймыра, Талнахская ул., 22, подъезд 2, Норильск, 663000, РФ

^вСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: сложность процессов распознавания и подсчета объектов на фотоснимке непосредственно связана с изменчивостью сопутствующих факторов: физическим различием объектов одного класса, наличием изображений, сходных с распознаваемыми объектами, неоднородным фоном, изменением условий съемки и положением объектов при получении снимка. В этом смысле сложными являются задачи распознавания людей в скоплениях, животных в природной среде, автомобилей с камер видеонаблюдения, объектов строительства и инфраструктуры на аэрофотоснимках и т. п. Задачи имеют свою специфику факторного пространства, но методические подходы к их решению сходны. **Цель:** разработка методологии и программных реализаций, решающих задачу распознавания и подсчета объектов с высокой изменчивостью, на примере распознавания северных оленей в природной среде. **Методы:** сравнение двух подходов распознавания: по признакам на основе бинарной классификации пикселей и по эталонам с использованием сверточных нейронных сетей. **Результаты:** разработаны методологии и программы для попиксельного распознавания с последующей бинаризацией, кластеризацией изображения и подсчетом числа кластеров и распознавания изображений с использованием сверточной нейронной сети архитектуры Mask R-CNN. При этом сеть первоначально обучается распознаванию животных как класса по массиву изображений MS COCO dataset, затем «доучивается» на массиве аэрофотоснимков стад северных оленей. Анализ результатов показывает, что признаковые методы с попиксельным распознаванием дают хорошие результаты на сравнительно простых изображениях (ошибка распознавания 10–15%). Наличие на снимке артефактов, близких по характеристикам к изображениям оленей, приводит к значительному росту ошибки. Сверточная нейронная сеть показала более высокую точность, которая на тестовой выборке составила 82%, причем ложные срабатывания отсутствовали. **Практическая значимость:** создан программный прототип системы распознавания на основе сверточных нейронных сетей с веб-интерфейсом, а сама программа введена в ограниченную эксплуатацию.

Ключевые слова — распознавание, сверточные нейронные сети, аэрофотоснимки, северные олени.

Для цитирования: Михайлов В. В., Колпащиков Л. А., Соболевский В. А., Соловьев Н. В., Якушев Г. К. Методологические подходы и алгоритмы распознавания и подсчета животных на аэрофотоснимках. *Информационно-управляющие системы*, 2021, № 5, с. 20–32. doi:10.31799/1684-8853-2021-5-20-32

For citation: Mikhailov V. V., Kolpaschikov L. A., Sobolevskii V. A., Soloviev N. V., Yakushev G. K. Methodological approaches and algorithms for recognizing and counting animals in aerial photographs. *Informatsionno-upravliayushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 20–32 (In Russian). doi:10.31799/1684-8853-2021-5-20-32

Введение

Сложность решения задачи автоматического распознавания объектов по их изображению непосредственно зависит от степени изменчивости факторов, сопутствующих распознаванию. Рассмотрим возможный набор факторов при распознавании объектов на фотоснимках.

1. Распознаваемые объекты как фактор распознавания, которые могут существенно различаться по цвету, размерам, форме — признакам, позволяющим отличать объекты разных классов.

2. Артефакты на фотоснимке, по некоторым признакам сходные с объектами распознаваемого класса.

3. Фон на снимке, который может быть однородным или неоднородным по цвету, может существенно отличаться на разных снимках.

4. Условия освещенности. Источники света могут быть естественными или искусственными, объекты могут находиться на солнце или в тени, снимки могут быть сделаны в разное время суток — все это приводит к изменению яркостных и цветовых характеристик объектов и фона.

5. Расположение объектов распознавания при съемке. Объекты могут находиться на разном удалении и под разными углами к камере, изображения объектов на снимках могут перекрывать друг друга, частично закрываться элементами фона.

Сложность решения задачи распознавания зависит от размерности факторного пространства (количества факторов) и от числа различных значений каждого фактора на снимке. В первом приближении сложность можно оценить по количеству факторов с изменчивостью. Наибольшую сложность имеют задачи, в которых все факторы обладают изменчивостью, наименьшую — задачи с фиксированными значениями факторов распознавания. Структуризация исходного множества снимков, на которых происходит обучение системы распознавания, на подмножества в целях фиксации значений каких-либо факторов позволяет упростить решение задачи распознавания на каждом подмножестве данных, но увеличивает объем ручного труда, затрачиваемого на анализ и группирование снимков.

К сложным в указанном выше смысле в полной мере относятся задачи распознавания различного рода природных объектов: людей в скоплениях, животных и птиц в естественной среде, растительных группировок в зональных и интразональных сообществах, аномалий планетарных морфологических структур (например, лунных кратеров) и т. п. В техносфере это, например, распознавание объектов строительства на аэрофотоснимках, автомобилей по снимкам с камер видеонаблюдения, замаскированных военных объектов по снимкам с беспилотников. Методические подходы к решению указанных задач сходны, однако каждая из них обладает своей спецификой факторного пространства и требует индивидуального подхода для учета этих особенностей и выбора конкретного метода распознавания.

Приведенный выше состав факторов соответствует задаче распознавания северных оленей на аэрофотоснимках. Какие-либо ограничения на значения факторов распознавания не накладывались за исключением масштаба снимков, который соответствует принятому при авиаучетных работах.

Выбор северных оленей как объекта распознавания и учета объясняется:

- соответствием в полной мере задаче с высокой изменчивостью факторов распознавания;
- актуальностью как сокращения ручного рутинного труда специалистов по обработке снимков, так и своевременного получения данных для решения задач управления экологическими процессами Крайнего Севера;
- наличием массива исходных данных, содержащего около 5000 архивных аэрофотоснимков

стад северных оленей, сделанных при проведении авиаучетов.

Используемые в настоящее время методики учета численности диких северных оленей тундровых популяций (таймырской, якутских популяций, оленей Чукотки, а также мигрирующих стад северных оленей — карибу Канады и Аляски) основываются на экологической особенности видов, состоящей в том, что олени в жаркую погоду во время вылета кровососущих насекомых собираются в многотысячные скопления на ограниченной территории в северной части летнего ареала [1, 2]. Стада в скоплениях фотографируются с летательного аппарата, и количество животных в них подсчитывается операторами вручную по каждому снимку. На одном фотоснимке могут присутствовать от нескольких сотен до 1000 и более оленей. Самое большое стадо, которое было зафиксировано на одном снимке и просчитано «по головам», содержало около 35 000 животных (Таймыр, район р. Верхняя Таймыра, 28.07.2017). Полное число снимков стад при проведении авиаучета может достигать тысячи. В результате прямого подсчета учитывается основная часть оленей в популяции. Количество животных, не вошедших в скопления, оценивается путем аппроксимации по территории [3].

Снимки, сделанные во время авиаучетов, разнообразны по типам ландшафтов, освещенности, направлению движения стад относительно камеры. Олени на снимках имеют разную величину и окраску, что связано как с физиологией животных, так и с условиями освещенности и положением стад. В скоплениях изображения оленей могут перекрывать друг друга, в результате часто видны лишь фрагменты тел. Помехи в виде камней, кочек, впадин затрудняют обнаружение животных. Ручная обработка результатов съемок специалистами-биологами занимает около трех месяцев, в то время как для экологически обоснованного управления динамикой численности популяции, рационального использования биологических ресурсов вида и определения норм промыслового изъятия оленей данные желательно иметь через 10–15 дней после окончания авиаучета.

О принципах и подходах к распознаванию природных объектов

Задача распознавания объектов живой природы имеет большое познавательное значение, поскольку ее решение позволяет увидеть природу в ее многообразии, выделить в общей мозаике по внешним признакам конкретные объекты, узнать их видовую принадлежность, найти их научное наименование и далее, при желании, узнать

их свойства, место и роль в экосистеме. С научных позиций процесс распознавания и учета дает возможность выявить видовой состав экосистемы и динамику ее компонент. В природоохранной деятельности подобная информация является основой для принятия решений о мерах по охране или по ограничению видов или отдельных популяций. Для промысловых видов животных регулярные учеты численности необходимы при определении лимитов изъятия, обеспечивающих рациональное использование биологических ресурсов.

Исторически задача распознавания биологических объектов решалась путем создания определителей, позволяющих установить место данного объекта в общей системе биологической классификации, т. е. установить видовой, родовой, семейный и другие уровни принадлежности объекта. Методически задача распознавания решается в определителях с использованием двух принципиально различных подходов. Первый — распознавание по признакам. Разработчиками определителей такого рода являются специалисты в области биологической систематики и классификации. Они определяют признаки, позволяющие объединить объекты на некотором уровне классификации, и признаки, по которым объекты переходят в разные группы на более низкий уровень классификации и различаются уже в них. В результате формируется дерево с ветвлениями по ключевым признакам. Распознавание объекта сводится к движению по дереву от корня (или какого-либо другого уровня классификации) к ветвям с выбором направления в точках ветвления в соответствии со значениями признаков. Второй подход — распознавание «по образцу». Определитель содержит множество эталонных изображений объектов разных видов. Методика распознавания состоит в сопоставлении искомого объекта с эталонами для выбора наиболее «похожего» образца. Пользователь выступает здесь в роли эксперта, оценивающего степень сходства изображений. Соответственно, ошибка распознавания будет зависеть от опыта, знаний эксперта и его персональных особенностей визуального восприятия объектов. Для распознавания на видовом уровне данный подход может быть весьма эффективным, поскольку основан на сопоставлении и сравнении целостных зрительных образов, что является естественной рутинной задачей зрительной системы человека. Во многих случаях определители по признакам дополняются набором изображений-эталонов, а человек при оценке «схожести» мысленно выделяет признаки сходства и различия объектов.

Компьютерные системы распознавания объектов живой природы также строятся с использованием этих двух подходов. Системы распознавания по признакам являются «прозрачными»

относительно состава используемых признаков объекта распознавания, методов, технологической цепочки преобразований изображения и результатов, включая промежуточные. Однако при распознавании объектов с высокой изменчивостью в нестационарной среде возникают сложности в связи с изменением параметров факторов распознавания. Появляется необходимость поиска новых систем признаков со сменой методов обработки изображений. В работе [4] высказывается предположение о принципиальной невозможности полной формализации процессов распознавания объектов в подобных условиях и необходимости перехода от автоматической к интерактивной технологии обработки изображений.

В современных системах распознавания по эталонам в качестве распознающего субъекта широко используются искусственные нейронные сети (ИНС) [5, 6]. В таких системах ИНС обучается на множестве изображений, в которых объекты распознавания (эталоны) заранее помечены. Затем обученная ИНС используется для распознавания объектов данного класса в реальной среде, на фотоснимках и видеокдрах. Логика распознавания объектов инкапсулирована в самой ИНС и не требует разработки в явном виде. С другой стороны, из-за сложности структур ИНС интерпретация результатов их работы зачастую затруднена. А сам процесс генерации и «обучения» таких структур происходит, по большей части, эмпирически. На сегодня не существует устоявшихся методик создания ИНС, которые бы подходили для большинства прикладных задач. Чаще всего имеются лишь общие рекомендации и эмпирический опыт разработчиков в решении конкретного класса задач.

Известные по литературным источникам системы распознавания и подсчета объектов по фотоснимкам ориентированы главным образом на человека как объект наблюдений. В меньшей степени это относится к другим видам живых организмов. Рассмотрим некоторые из таких систем.

В статье [7] представлена программа распознавания животных и их поведения в дикой природе по фотоснимкам с 225 камер-ловушек заповедника Серенгети (Танзания). Причина разработки системы заключалась в том, что из-за ограниченности в людских ресурсах реально используется лишь малая часть той информации, которая хранится на серверах. Программа распознавания разрабатывалась совместно в университетах Вайоминга, Миннесоты, Обернска, Гарварда, Оксфорда и в компании Uber AI. В работе использовался метод глубокого обучения различных архитектур сверточных нейронных сетей (СНС) [8]. Были использованы 1,5 млн изображений, из которых 1,4 млн служили для обучения сети, 0,1 млн — для тестирования. По результатам те-

стирования вероятность обнаружения животных с бинарным ответом «да/нет» составила около 96 %. Вид животных из 48 возможных вариантов распознавался с вероятностью 99 %. Однако подсчет животных по кластерам 1, 2, ..., 10, 11–50, > 50 выполнялся программой с точностью лишь 63 %. Это говорит о сложности фиксации отдельного животного на снимке в условиях возможных наложений изображений животных друг на друга, их частичного перекрытия посторонними предметами и других искажений изображения животных.

Сверточные нейронные сети применялись также для распознавания слонов на космоснимках [9]. Для обучения сети использовались архивные снимки со спутников Worldview-3 и Worldview-4 с разрешением 31 см, сделанные в период с 2014 по 2019 год на территории национального парка Addo Elephant в Южно-Африканской Республике. На площади свыше 1,6 тыс. км обитают более 600 саванных слонов. Парк имеет неоднородный ландшафт: там есть кустарники, невысокие леса, открытые луга, водоемы. В течение дня слоны перемещаются по территории и иногда обливают себя грязью, чтобы защититься от жары. Все это затрудняет их обнаружение.

Разработчики протестировали систему на независимой выборке снимков территорий Кении и сравнили точность ее работы с результатами 51 добровольца, которые идентифицировали слонов вручную. Среди участников эксперимента были специалисты по машинному обучению, ученые, работники национальных парков. Вероятность обнаружения слонов на гетерогенной территории с использованием СНС составила 78 %, люди находили слонов с вероятностью 77 %. Для однородной территории результаты составили соответственно 73 и 80 %.

Разработана и протестирована система автоматического распознавания, сегментации и подсчета морских котиков в рамках онлайн соревнования платформы Kaggle по Data Science [10]. Постановка задачи связана с мониторингом популяции котиков Западных Алеутских островов, который ведется с использованием дронов. Обработка полученных фотоснимков выполняется вручную, что занимает около четырех месяцев каждый год. Подсчет особей выполняется по пяти половозрастным классам. Для разработки автоматической системы был сформирован массив из 948 размеченных фотоснимков и еще 18 641 снимок представлен для тестирования системы. Снимки сделаны с различной высоты и в разном масштабе, имеются помехи, сходные с изображениями животных, между некоторыми классами котиков отсутствуют четкие визуальные различия. Эти причины усложняют решение задачи автоматического распознавания и подсчета котиков

на аэрофотоснимках. Предложенный вариант решения основан на использовании базовой архитектуры СНС Inception-Resnet-V2, которая расширена добавлением слоев average-pooling + FC-слой на 256 нейронов + Dropout + финальный FC-слой на пять нейронов. Рейтинг проекта — 4-е место среди 385 участников. Особенность данной статьи состоит в том, что в ней не только дается описание конкретной системы, но детально рассматривается весь процесс ее разработки с анализом правильных и ошибочных решений при подборе архитектуры и параметров нейронной сети.

В программах Seek iNaturalist распознавания различных видов растений и животных для любителей природы использованы СНС на основе SSD (Singl-Shot Multy Box Detector) [11]. Такая архитектура дает возможность применять миниатюрные и мобильные устройства для фото- и видеоанализа изображений в реальном времени.

Можно сделать несколько замечаний, касающихся построения автоматических систем распознавания животных на основе существующих представлений о решении задачи классификации и имеющихся программных разработок:

- животное в природной среде — это сложный объект для распознавания. Изменчивость изображений на фотоснимках, их наложение друг на друга, наличие помех приводят к ошибкам распознавания, допускаемых как человеком, так и автоматической системой;

- наиболее приспособленным к решению задач распознавания изображений природных объектов в настоящее время является аппарат СНС;

- повысить точность распознавания можно за счет увеличения объема обучающей выборки или путем перехода к интерактивному режиму обработки снимков с участием человека;

- современные архитектуры СНС дают возможность создавать мобильные приложения для работы в полевых условиях непосредственно при проведении учетных работ.

Распознавание животных на аэрофотоснимках по признакам

Исходя из общих принципов построения систем распознавания образов [12], решение задачи распознавания и подсчета животных по признакам укрупненно можно провести в несколько этапов:

- 1) предварительная обработка снимка для повышения различимости объектов;
- 2) бинаризация, т. е. нахождение пикселей, относящихся к изображениям объектов;
- 3) сегментация, т. е. выделение связанных групп пикселей, относящихся к отдельному объекту;
- 4) определение числа объектов.

При разработке системы подсчета оленей следует учесть, что для дальнейшего использования результатов точность подсчета должна находиться в пределах 10–15 %. Таким образом, ошибки в указанных пределах при выделении оленей (пропуск отдельных объектов, их ложное выделение, неточность при подсчете числа оленей в слитных сегментах) вполне допустимы.

На первом этапе производится очистка изображения от помех и шумов. В качестве фильтров могут использоваться различные усредняющие пространственные фильтры [13]. На этом этапе могут быть увеличены контрастность и резкость изображения, изменен масштаб, выполнен переход из одной цветовой модели в другую. Если объекты имеют хорошо различимые резкие границы, то предварительная обработка, как правило, включает и методы выделения границ (фильтры Робертса, Кирша, Превитта, Лапласа [13]).

Аэрофотоснимки выполняются в цвете, т. е. каждый пиксель описывается тремя цветовыми составляющими — R (красный), G (зеленый) и B (синий). Приведение изображения к монохромному (в градациях серого) виду может упростить последующий процесс бинаризации [4].

Процесс бинаризации можно представить как классификацию пикселей снимка, т. е. разделение их на два класса. Важной задачей классификации является выбор признаков, которые должны удовлетворять следующим условиям:

- значения признаков для объектов одного класса должны различаться слабо, а для объектов разных классов — существенно;

- набор признаков должен быть полным, т. е. должен обеспечивать идентификацию объектов любого из классов;

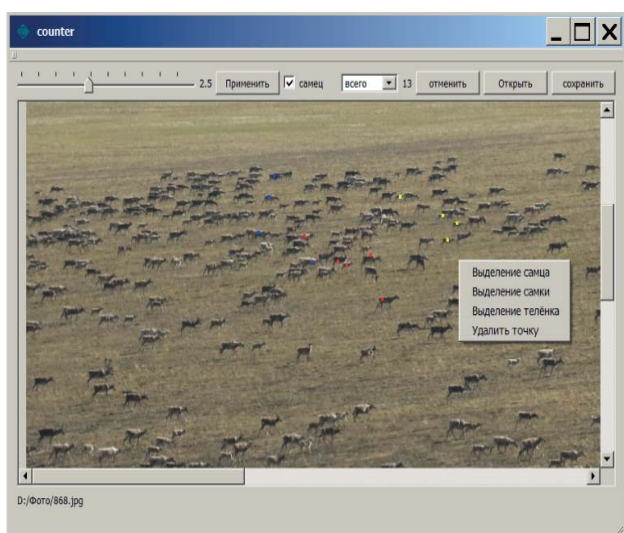
- общее количество признаков должно быть минимальным.

В качестве признаков пикселя можно использовать значения составляющих его цвета, причем как все вместе, так и в различных сочетаниях или по отдельности. Для разделения пространства признаков на два класса необходимо предварительно по обучающей выборке снимков собрать статистические данные о пикселях, относящихся к оленям, в ручном режиме. Для решения этой достаточно трудоемкой задачи была разработана программа [14], позволяющая отображать на дисплее снимок, менять его масштаб, выбирая наиболее подходящий для распознавания объектов, ставить на снимке три разные по цвету метки для самца, самки и теленка, выполнять подсчет выделенных объектов и сохранять результаты (координаты меток и их число) в текстовом файле (рис. 1).

В одной из первых работ по распознаванию и подсчету северных оленей [15] в качестве признака выделения оленей и подстилающей поверхности были приняты различия величин их спектральной яркости. По результатам экспериментов выбрана цветовая пара: зеленый (фон) и голубой (распознаваемый объект), — после чего было составлено неравенство, по которому пиксели объекта отличаются от пикселей фона. Отнесение сегмента к изображению оленя выполнялось по результатам анализа площади и вытянутости сегментированной области. Тестирование показало весьма высокую точность распознавания и подсчета оленей на искусственно сгенерированных «простых» снимках (зеленый фон, отсутствие помех и наложений объектов друг на друга). Ошибка подсчета численности составила около 2 %. Однако при работе с реальными снимками величина ошибки резко возросла, что связано с высокой изменчивостью распознаваемых объектов и условий съемки.

Задачу бинаризации можно представить, как хорошо известную в распознавании образов [12, 13] кластеризацию по k -средним (в данном случае число кластеров $k=2$) в пространстве признаков RGB. Основное преимущество метода — отсутствие необходимости собирать статистические данные перед его применением. Эксперименты на реальных снимках с отсутствием помех показали достаточно хорошие результаты — ошибок около 10 %, но на снимках с неравномерным фоном и наличием помех число ошибок резко выросло. Данный метод наиболее целесообразно использовать для предварительной обработки снимков на этапе сбора статистических данных с последующей корректировкой ошибок в ручном режиме.

Пороговый метод бинаризации пикселей цветного снимка требует установки порогов для отдельных цветовых составляющих с их последующим совместным использованием. Например, для цветовой схемы HSB (H — цветовой тон, S —



■ **Рис. 1.** Интерфейс программы ручной обработки снимков

■ **Fig. 1.** Manual image processing program interface

насыщенность, B — интенсивность) в литературе [16] предлагается для бинаризации использовать только цветовой тон и интенсивность, причем относить к объекту только те пиксели, у которых каждая составляющая меньше соответствующего ей порога, т. е.

$$Z(x, y) = \begin{cases} 1 & H(x, y) < T_H \vee B(x, y) < T_B, \\ 0 & \text{в остальных случаях} \end{cases}, \quad (1)$$

где $Z(x, y)$ — значение яркости пикселя на бинаризованном изображении; $H(x, y)$ — цветовой тон пикселя с координатами (x, y) ; T_H — пороговое значение для цветового тона; $B(x, y)$ — интенсивность пикселя с координатами (x, y) ; T_B — пороговое значение для интенсивности.

Порог T определяется по формуле

$$T = \frac{\sum_x \sum_y I(x, y) \Delta I(x, y)}{\sum_x \sum_y \Delta I(x, y)}, \quad (2)$$

где $I(x, y)$ — значение выбранной для нахождения порога составляющей цветовой модели в пикселе с координатами (x, y) ; $\Delta I(x, y)$ — градиент составляющей цветовой модели:

$$\Delta I(x, y) = \max(\text{abs}(I_x(x, y)), \text{abs}(I_y(x, y))), \quad (3)$$

$$I_x(x, y) = I_x(x + 1, y) - I_x(x - 1, y);$$

$$I_y(x, y) = I_y(x, y + 1) - I_y(x, y - 1). \quad (4)$$

Алгоритм сегментации бинарных изображений хорошо известен [10] и достаточно надежен. В результате его работы каждое связанное множество пикселей получает свой порядковый номер и при необходимости может быть выделено в отдельное изображение для последующего подсчета объектов. Основная проблема, возникающая при подсчете, состоит в том, что на некоторых участках снимка отдельные олени могут сливаться в одно множество сегментированных пикселей. Для решения этой проблемы необходимо проводить анализ распределения сегментов по площади (число пикселей в сегменте), что позволяет оценить среднюю площадь одного объекта и, соответственно, с достаточной точностью найти как число объектов в слитном сегменте, так и общее число объектов на снимке. Сегменты, площадь которых существенно меньше средней площади объекта, целесообразно считать артефактами и не учитывать при подсчете оленей.

Для проведения экспериментов и сравнения результатов применения различных методов распознавания оленей по признакам из имеющихся 100 снимков были сделаны выборки снимков с близкими характеристиками фона и окраса оленей. Для снимков каждой выборки был про-

изведен ручной подсчет оленей и поставлены эксперименты по автоматическому подсчету оленей. Для проведения экспериментов была разработана программа, позволяющая выбирать метод бинаризации, устанавливать необходимые параметры обработки, визуализировать результаты и сохранять количественные показатели для последующей статистической обработки. Сравнение результатов экспериментов с данными ручного подсчета оленей показало, что наиболее близкое совпадение ($\pm 15 \dots 20$ %) дает бинаризация по порогам цветового тона и насыщенности. Наличие на снимке артефактов, близких по яркостным характеристикам, форме и площади к изображениям оленей на данном снимке, приводит к значительному росту ошибки подсчета оленей.

Результаты экспериментов показали, что система распознавания и подсчета оленей на аэрофотоснимках по признакам должна базироваться на детальном анализе и классификации изображений, сборе статистических данных о характеристиках пикселей, относящихся к изображению оленей, определении наиболее информативных признаков и выборе эффективных алгоритмов классификации пикселей и подсчета оленей. Система должна обладать достаточными интеллектуальными способностями для комбинации методов распознавания оленей на предьявляемом снимке.

Распознавание животных на аэрофотоснимках по эталонам с использованием ИНС

СНС как инструмент распознавания изображений

Для решения задачи распознавания объектов живой природы путем обучения по их эталонным изображениям в настоящее время все чаще начинают применять СНС [17]. Данный класс архитектур ИНС является узкоспециализированным инструментом, подходящим прежде всего для работы с изображениями и иными данными, которые возможно представить в матричном виде.

Поскольку в изображениях вся информация хранится в виде двумерных матриц (т. е. в виде пикселей), при работе с ними требуется учитывать значения, поступающие не только от самих нейронов, но и от группы ближайших нейронов. Для этого в сверточных слоях СНС, помимо нейронов, имеется еще один тип элементов, которые применяют определенные линейные операции ко всем входным данным каждого нейрона слоя, — ядро свертки. Ядро свертки представляет из себя сетку, которая «скользит» по изображению (или по сверточному слою предыдущего уровня) и ищет шаблоны и паттерны в данных.

Если будет найдена такая часть изображения, которая совпадет с шаблоном ядра, то оно передаст расчетному нейрону текущего слоя большее положительное значение. Если совпадения не будет, то ядро передаст небольшое значение или нуль.

Из-за того, что ядро свертки применяется к каждой позиции изображения, сверточный слой СНС является крайне эффективным в задачах обработки изображений, поскольку особенности или закономерности в изображениях могут проявляться в любом месте этих изображений. То есть СНС, в отличие от прочих ИНС, способна анализировать контекстно-зависимые данные.

При этом сверточные слои в СНС могут встречаться несколько раз. И тогда каждый последующий слой получает данные от предыдущего сверточного слоя как от входного, поскольку сверточный слой также является двумерным массивом, и вектор его выходных значений может быть представлен в виде изображения. И в таком случае каждый последующий сверточный слой может находить все более сложные шаблоны и паттерны изначального входного массива.

Для поставленной задачи была выбрана архитектура Mask Regions with Convolution Neural Networks (MRCNN) [18] (рис. 2), являющаяся подвидом классической СНС. За счет усложнения архитектура более успешно справляется с задачами семантической и объектной сегментации изображений [19, 20].

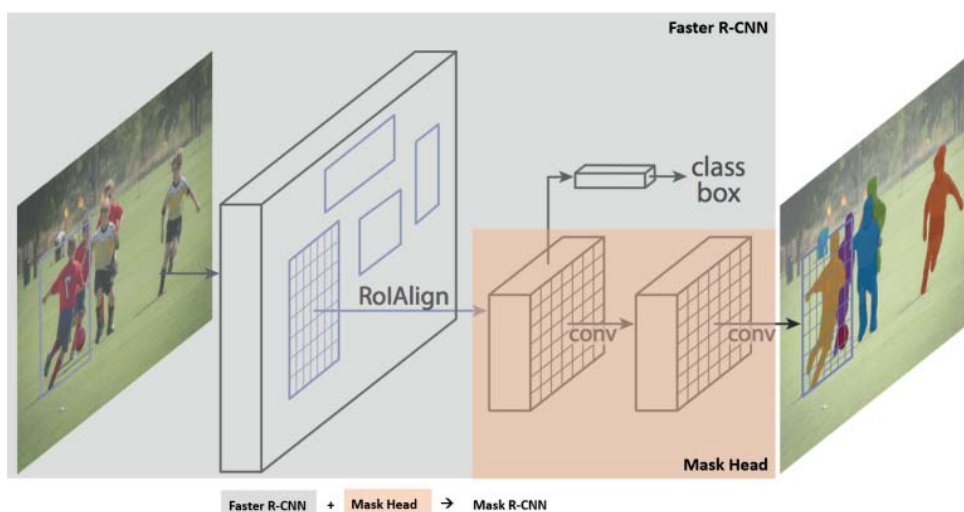
Ключевой особенностью данной архитектуры является совмещение такой архитектуры СНС, как Faster R-CNN (FRCNN), которая отвечает за решение задачи классификации, с модулем Mask Head, который отвечает за решение задачи

сегментации изображения. Результатом работы MRCNN является совмещенный ответ двух описанных составных частей.

Faster R-CNN — это СНС, которая осуществляет поиск на изображении объектов, после чего дополнительно производит классификацию найденного объекта. Результатом работы данной ИНС являются ограничивающие прямоугольники для каждого объекта (т. е. прямоугольная граница, которая ограничивает найденный объект) и метка класса найденного объекта с оценкой достоверности.

Первый этап работы FRCNN состоит из одновременного запуска двух включенных ИНС: магистральной (ResNet, VGG, Inception или аналогичных) и сети регионального позиционирования. Эти сети обрабатывают каждое поступающее на вход FRCNN изображение и на выходе предоставляют трехмерный массив — массив предложенных двумерных областей. Данные двумерные области являются вырезанной частью изображения, на которой магистральная СНС распознала какой-либо объект. Размер данных частей не фиксирован и зависит от размера распознанного объекта.

На втором этапе FRCNN прогнозирует координаты ограничивающих прямоугольников и классы объектов для каждой из предложенных областей, полученных на первом этапе. Каждая предлагаемая область может иметь разный размер, но так как сверточные слои в СНС всегда требуют вектора фиксированного размера для прогнозирования, на данном этапе также производится масштабирование найденных регионов. Размер регионов масштабируется с помощью либо алгоритма RoI, либо метода RoIAlign.



■ Рис. 2. Обобщенная схема MRCNN [18]
 ■ Fig. 2. Generalized MRCNN scheme [18]

MRCNN в свою очередь является расширенной версией FRCNN, дополненной ветвью для прогнозирования масок сегментации для каждой области интереса. На втором этапе работы MRCNN уже используется только RoIAlign, помогающий сохранить оригинальные пространственные координаты, которые смещаются в случае использования RoI. Это необходимо затем, чтобы выходные данные RoIAlign можно было совместить с данными, полученными на первом этапе, и с помощью модуля Mask Head (который в свою очередь также реализован на базе сверточных слоев) сгенерировать маску для каждого ответа RoIAlign. Такие маски представляют из себя двумерную матрицу, которая для каждого пикселя, входящего в границы региона объекта, определяет, принадлежит ли этот пиксель искомому объекту или нет.

Подобный подход, с одной стороны, позволяет более точно определить границы искомого объекта. В идеальных случаях MRCNN может точно вычислить все пиксели изображения, которые отображают искомый объект. С другой стороны, в процессе обучения данные, получаемые в модуле Mask Head, используются для дополнительного обучения магистральной СНС, что позволяет повысить и ее точность работы.

При этом, поскольку для попиксельного поиска используются результаты расчета магистральной СНС, модуль Mask Head во время работы производит семантическую сегментацию только текущего искомого объекта. То есть даже если рядом расположены два объекта одного класса, Mask Head на каждой отдельной итерации своей работы будет распознавать пиксели, принадлежащие только одному из этих объектов. Такой подход позволяет MRCNN корректно решать задачу распознавания, когда искомые объекты расположены близко или даже перекрывают друг друга.

Обучение сети для распознавания северных оленей на аэрофотоснимках

Основным массивом обучения для MRCNN является массив изображений MS COCO dataset (Microsoft Common Objects in Context) [21]. На сегодня это самый крупномасштабный набор данных, используемый для обучения моделей машинного обучения решению задач обнаружения и сегментации [22, 23], состоящий из 328 тысяч изображений. Все изображения уже размечены и сформированы в обучающие выборки. Поэтому использование данного массива для базового обучения MRCNN позволяет задать для нее все основные концепции различных классов объектов, в том числе и животных. Однако изображения оленей не входят в MS COCO dataset, и MRCNN по умолчанию не способна отличить их от ряда других животных (овец, газелей, коров, лошадей).

Результаты распознавания оленей на аэрофотоснимке с помощью базовой MRCNN, обученной на MS COCO dataset, представлены на рис. 3. На снимке правильно распознанные изображения оленей помечены зеленым цветом, их четыре. Красным цветом помечены артефакты, ошибочно принятые за оленей, этих пометок 32.

Таким образом, при обучении сети необходимо учитывать особенности изображений северных оленей на аэрофотоснимках и их отличие от изображений камней, кочек и других объектов на снимках, которые могут быть приняты за оленей. Специфика аэрофотоснимков состоит в том, что фотографирование стад производится с разного расстояния, на разных ландшафтах, при различных условиях освещенности, животные на снимках имеют различную окраску и могут находиться под разными углами к камере, могут перекрывать друг друга. Эти особенности аэрофотоснимков создают дополнительные сложности при решении задачи распознавания оленей.

Задача распознавания северных оленей по аэрофотоснимкам является, таким образом, не тривиальной. Поэтому для ее решения была предложена двухэтапная процедура обучения MRCNN. Первоначальное обучение на массиве MS COCO dataset, который включает изображения других животных, необходимо для того, чтобы сеть научилась распознавать животных как класс объектов. На втором этапе MRCNN дообучается на массиве аэрофотоснимков стад северных оленей. Данный подход известен как transfer learning [24] и сегодня применяется при решении многих задач, для которых невозможно, по тем или иным причинам, собрать массив обучающих данных для конкретной рассматриваемой ситуации. Поэтому сначала модель обучается на большом массиве похожих ситуаций, а после этого уже дообучается на целевых примерах.



■ **Рис. 3.** Результат распознавания оленей с помощью базовой MRCNN, обученной на MS COCO dataset

■ **Fig. 3.** Reindeer recognition result using basic MRCNN trained on MS COCO dataset



■ **Рис. 4.** Результат распознавания оленей с помощью дообученной MRCNN
 ■ **Fig. 4.** Reindeer recognition result using the re-trained MRCNN

Был подготовлен входной массив целевых примеров, содержащий обучающую выборку из 100 аэрофотоснимков стад, на которых все животные помечены, и тестовую выборку из 30 исходных снимков стад, не включенных в обучающую выборку.

При обучении модели были установлены следующие параметры MRCNN: 20 эпох обучения; 60 шагов обучения в эпохе; скорость обучения 0,0058 (безразмерная); порог пропуска обнаружения 0,7 (безразмерный). Поскольку СНС в общем случае являются черными ящиками, нельзя в явном виде описать влияние конкретных значений параметров на результаты работы обученной сети. Данные значения были установлены исходя из практики работы с СНС и с учетом эмпирических наблюдений в процессе обучения.

После дообучения сети эксперимент по распознаванию оленей на фотоснимке, изображенном на рис. 3, был повторен. Результат эксперимента представлен на рис. 4. Модель корректно распознала на снимке 70 оленей из 93 (точность 75 %) и при этом не совершила ни одной ошибки второго рода, т. е. ни разу не приняла за оленя объект, который оленем не является. В среднем же на всем тестовом массиве данных обучения модель корректно распознала 82 % оленей.

Веб-интерфейс системы и результаты ее валидации на независимом массиве данных

На данный момент для разработанной сети создан программный комплекс с веб-интерфейсом (<https://regionview.ru/ai/>), а сама программа введена в ограниченную эксплуатацию.

Поскольку система была создана на основе сервис-ориентированной архитектуры [25], данный программный комплекс можно использовать как стационарно, так и удаленно, посредством сети Интернет.

Для работы с системой пользователь должен загрузить в свой компьютер снимки стад в форматах JPEG (.jpg) или GIF (.gif).

Интерфейс системы содержит набор оконных форм, обеспечивающих:

- загрузку аэрофотоснимков с компьютера пользователя для их обработки;
- запуск программы распознавания и подсчета оленей на снимках;
- представление результатов работы программы с отображением на экране снимка, на котором помечены распознанные системой изображения оленей, и общего числа подсчитанных животных;
- скачивание результатов на компьютер пользователя.

После просмотра помеченного снимка пользователь, если его не удовлетворит точность программного комплекса, может продолжить дальнейшую обработку снимка вручную с использованием программы ручной обработки снимков (см. разд. «Распознавание животных на аэрофотоснимках по признакам») или в любом графическом редакторе, поддерживающем расширение файла с изображением .jpg.

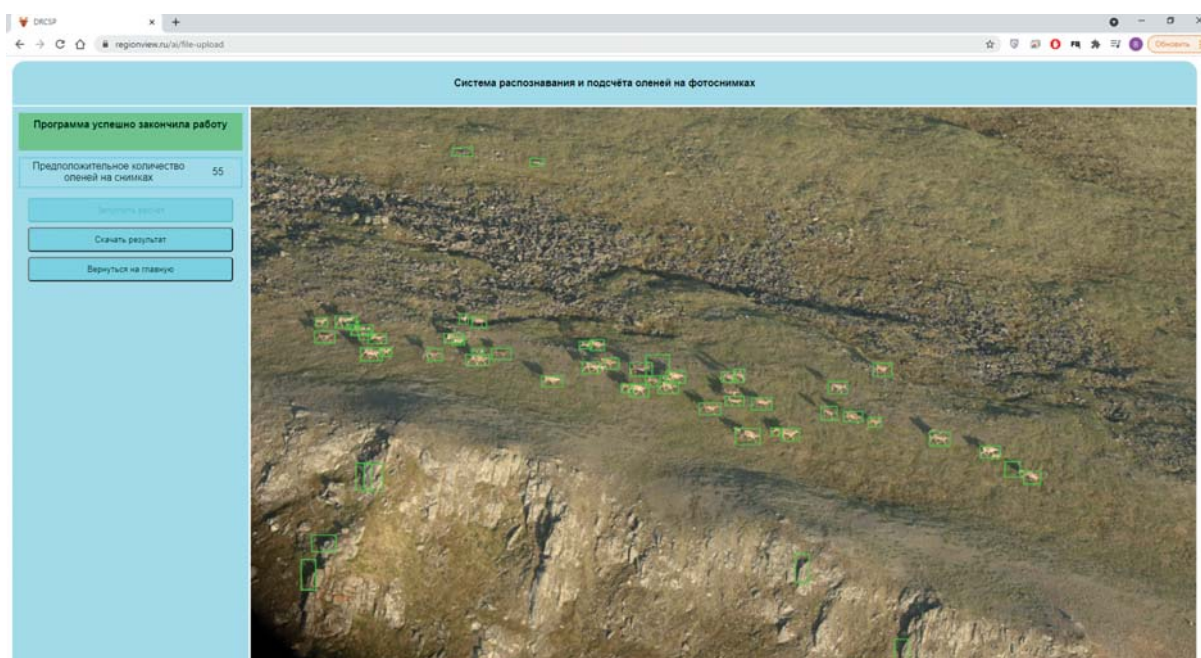
Графический интерфейс системы распознавания и подсчета оленей показан на рис. 5.

Проверка работы программы была проведена на независимом массиве из 10 аэрофотоснимков.

На фотоснимке небольшого стада оленей (рис. 6) при распознавании было совершено четыре ошибки: два оленя были распознаны дважды и два теленка не были распознаны. Всего из 41 оленя корректно было распознано 35 особей. Однако ошибка подсчета количества животных в стаде составила всего 10 %, что является достаточно высоким результатом для автоматических систем распознавания животных в естественных условиях. Причинами этого в данном случае являются высокое качество снимка и высокое разрешение самих оленей, а также сильная контрастность фона с объектами распознавания.

На снимке (рис. 7) видно, что MRCNN хорошо работает с большими стадами. При этом заметно, что сеть с одинаковой эффективностью распознает оленей, находящихся на разном удалении от камеры. Таким образом, сеть не привязана к конкретному разрешению объектов и способна работать с искаженной перспективой. Ошибка распознавания составила около 3 %.

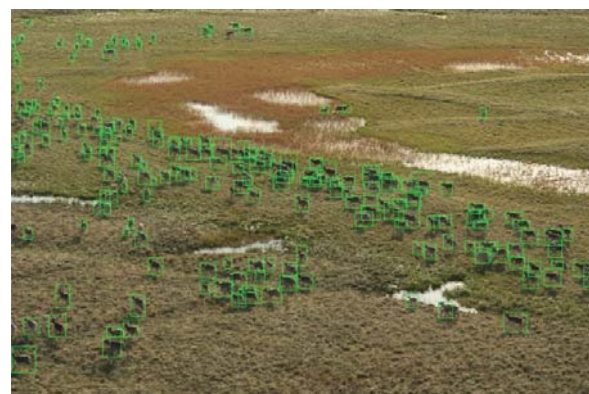
Стадо оленей на фоне типичного неоднородного тундрового пейзажа показано на рис. 8. По разметке изображений оленей видно, что MRCNN может работать со снимками, которые зашумлены фоновыми объектами — лужами, озерами, бугорками, полигонами и т. п. Объекты фона ни разу не были приняты системой распознавания за изображения оленей. Также заметно, что сеть



■ **Рис. 5.** Окно интерфейса системы, режим «Программа успешно закончила работу»
 ■ **Fig. 5.** System interface window, “Program successfully completed work” mode



■ **Рис. 6.** Распознавание оленей на однородном фоне
 ■ **Fig. 6.** Deer recognition in a homogeneous field



■ **Рис. 8.** Крупное стадо на неоднородном фоне
 ■ **Fig. 8.** Large herd against a heterogeneous background



■ **Рис. 7.** Распознавание и подсчет оленей в крупном стаде на однородном фоне
 ■ **Fig. 7.** Recognizing and counting deer in a large herd on a homogeneous background

хорошо работает со стадами, в которых олени собираются в очень плотные группы. Ошибка распознавания составила около 17 %.

Заключение

Выполненные исследования показали, что разработка автоматической системы распознавания и подсчета северных оленей в признаковом пространстве требует больших затрат времени специалистов для анализа снимков, отбора признаков и критериев разделения изображений животных от фона, проведения разного рода тестов

вых компьютерных экспериментов. СНС учится выполнять задачу классификации непосредственно по изображениям без предварительного ручного отбора признаков. Наличие готовых архитектур СНС, обученных на распознавание на изображениях различного рода объектов, дает возможность путем дообучения эффективно создавать приложения, ориентированные на распознавание конкретных видов животных. С использованием данного подхода авторами был создан программный комплекс для распознавания и подсчета северных оленей по аэрофотоснимкам, для которого можно выделить следующие перспективные направления развития.

1. Уменьшение ошибок распознавания. Необходимо увеличить объем обучающей выборки, особенно для сильно «зашумленных» и разномасштабных изображений животных. Однако связь между величиной ошибки и объемом выборки носит экспоненциально затухающий характер, а чрезмерное увеличение выборки может привести к переобучению модели. Поэтому заранее нельзя оценить количество снимков, которые потребуются добавить в обучающую выборку. Кроме того, при учете животных в скоплениях всегда может возникнуть ситуация, с которой автоматическая система распознавания справиться не сможет. В этом случае система распознавания может работать в интерактивном режиме совместно с программой обработки изображений, представленной в разд. «Распознавание животных на аэрофотоснимках по признакам». Эта программа позволяет после получения результатов автоматического распознавания корректировать их в ручном режиме, удаляя ошибочные метки и (или) ставя их на нераспознанные программой объекты.

2. Перенос программного комплекса на высокопроизводительные серверы. Специфика вычислений при использовании MRCNN накладывает ряд требований на аппаратную часть.

А коммерческое использование подразумевает параллельную работу с программным комплексом нескольких пользователей, что может существенно нагрузить систему. Несмотря на заложенную в архитектуру программного комплекса параллельность вычислений, следует провести дополнительные исследования и стресс-тесты, чтобы определить конкретные требования к аппаратной части.

3. Анализ видеоизображений. Сегодня архитектура MRCNN используется при обработке видеоизображений в ряде проектов [26]. Имеются программные библиотеки, позволяющие интегрировать данную архитектуру СНС в программные комплексы, работающие с видеопотоком. Это позволяет модифицировать представленную разработку для распознавания и подсчета оленей в видеопотоке. Подобный подход поможет ускорить процесс получения результатов учета, поскольку исключает процедуру монтажа снимков, необходимую при дискретной фотосъемке стад оленей.

4. Подсчет других видов животных и птиц. Основным массивом для начального обучения MRCNN является массив изображений MS COCO dataset. Изображения северных оленей служат для дообучения СНС. Подобным же образом, без изменения архитектуры СНС и организации интерфейса, система может быть дообучена для распознавания и подсчета стад сайгаков, стай гусей в период линьки или некоторых видов полярных чаек.

Финансовая поддержка

Работа выполнена при поддержке World-Wide Fund for Nature — Contract № BBF003107, РФФИ — грант № 19-37-90112 и бюджетной темы № 0073-2019-0004.

Литература

1. Зырянов В. А., Павлов Б. М., Якушкин Г. Д. *Экологические основы учета численности промысловых животных в тундровой зоне Таймыра*. Проблемы охотничьего хозяйства Красноярского края: материалы совещ. Красноярск, 1971. С. 70–72.
2. Kolpashikov L., Mikhailov V., Russell D. E. The role of harvest, predators, and socio-political environment in the dynamics of the Taimyr wild reindeer herd with some lessons for North America. *Ecology and Society*, 2015, vol. 20, iss. 1, article 9. <https://www.ecologyandsociety.org> (дата обращения: 13.10.20).
3. Челинцев Н. Г. *Математические основы учета животных*. М., ГУ Центрохотконтроль, 2000. 431 с.
4. Соловьев Н. В., Сергеев А. М. Распознавание на изображении множества однотипных объектов с высокой изменчивостью. *Научная сессия ГУАП: сб. докл.: в 3 ч. Ч. II: Технические науки*. СПб., ГУАП, 2019. С. 427–430.
5. Simon J. D. *Principles of computer vision: Models, learning, and inference*. Cambridge, Cambridge University Press, 2012. 598 p.
6. Ayyadevara K., Reddy Y. *Modern computer vision with PyTorch: Explore deep learning concepts and implement over 50 real-world image applications*. Birmingham, Packt Publishing, 2020. 824 p.
7. Norouzzadeh M. S., Nguyen A., Kosmala M., Swanson A., Palmer M., Packer C., Clune J. Automated animal identification using deep learning techniques.

- Proc. of the Nat. Acad. of Sciences of the USA*, 2018, June, vol. 115, iss. 25, pp. 5716–5725. doi.org/10.1073/pnas.17193671155
8. Lecun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, vol. 1, iss. 4, pp. 541–551.
 9. Duporse J., Isupova O., Reece S., Macdonald D., Way T. Using very- high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. *bioRxiv — the preprint server for biology*. doi: <https://doi.org/10.1002/rse2.195>
 10. *Open Data Science. Kaggle: как наши сеточки считали морских львов на Алеутских островах*. <https://habr.com/ru/company/ods/blog/337548/> (дата обращения: 20.11.20).
 11. *Seek by iNaturalist*. https://www.inaturalist.org/pages/seek_app (дата обращения: 21.11.20).
 12. Tou J., Gonzales R. *Pattern Recognition Principles*. Addison-Wesley Publishing Co, 1974. 378 p.
 13. Gonzales R., Woods R. *Digital Image Processing*. Prentis Hall, 2002. 822 p.
 14. *Программа для подсчета однотипных объектов на сжатом изображении*. Свидетельство о гос. рег. программы для ЭВМ № 2019666902 Российская Федерация/ А. М. Сергеев, Н. В. Соловьев, Г. К. Якушев. № 2019665795, заявл. 04.12.19; опубл. 17.12.19; Бюл. № 12. 1 с.
 15. Михайлов В. В., Харин Я. В. К вопросу о построении системы распознавания и подсчета животных по аэрофотоснимкам. Ч. 1. Анализ методов распознавания. *Информационно-управляющие системы*, 2011, № 2, с. 22–28.
 16. Пятаев А. С. Сегментация дерева на изображении методом неполной пороговой обработки с глобальным порогом. *Решетневские чтения*, 2018, т. 2, с. 291–293.
 17. LeCun Y., Haffner P., Bottou L., Bengio Y. *Object Recognition with Gradient-Based Learning*. Shape, Con-
tour and Grouping in Computer Vision. Springer, 1999. Pp. 319–345.
 18. He K., Gkioxari G., Dollar P., Girshick R. *Mask R-CNN*. Computer Vision and Patter Recognition. Cornell University, 2017. <https://arxiv.org/abs/1703/1703.06870>.
 19. Ganesh P., Volle K., Burks T. F., Mehta S. S. Deep orange: Mask R-CNN-based orange detection and segmentation. *IFAC-PapersOnLine*, 2019, vol. 52, iss. 30, pp. 70–75. <https://doi.org/10.1016/j.ifacol.2019.12.499>
 20. Zhao G., Hu J., Xiao W., Zou J. A mask R-CNN based method for inspecting cable brackets in aircraft. *Chinese Journal of Aeronautics*, 2020. <https://doi.org/10.1016/j.cja.2020.09.024>
 21. Lin T. Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollar P., Zitnick C. L. Microsoft COCO: Common objects in context. *Computer Vision — ECCV 2014*, 2014, part 5, pp. 740–755.
 22. Patterson G., Hays J. COCO attributes: Attributes for people, animals, and objects. *Computer Vision — ECCV 2016*, 2016, part 6, pp. 85–100.
 23. Srivastava S., Divekar A. V., Anilkumar C., Naik I., Kulkarni V., Pattabiraman V. Comparative analysis of deep learning image detection algorithms. *Journal of Big Data*, 2021, no. 8, article 66. doi:10.1186/s40537-021-00434-w
 24. Weiss K., Khoshgoftaar T. M., Wang D. D. A survey of transfer learning. *Journal of Big Data*, 2016, vol. 3, article 9. doi:10.1186/s40537-016-0043-6
 25. Niknejad N., Ismail W., Ghani I., Nazari B., Bahari M., Ab Razak Bin Che Hussin. Understanding Service-Oriented Architecture (SOA): A systematic literature review and directions for further investigation. *Information Systems*, 2020, vol. 91, article 101491.
 26. Yang L., Fan Y., Xu N. Video instance segmentation. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 5188–5197.

UDC 004.932.2

doi:10.31799/1684-8853-2021-5-20-32

Methodological approaches and algorithms for recognizing and counting animals in aerial photographsV. V. Mikhailov^a, Dr. Sc., Tech., Professor, Leading Researcher, orcid.org/0000-0001-6973-3526, mwwcari@gmail.comL. A. Kolpaschikov^b, Dr. Sc., Biol., Head of a Research Division, orcid.org/0000-0003-4871-3360B. A. Sobolevskii^a, Junior Researcher, Post-Graduate Student, orcid.org/0000-0001-7685-4991N. V. Soloviev^c, PhD, Tech., Associate Professor, orcid.org/0000-0001-8851-3713G. K. Yakushev^c, Bachelor, orcid.org/0000-0003-4689-3425^aSt. Petersburg Federal Research Center of the RAS, 39, 14 Line, V. O., 199178, Saint-Petersburg, Russian Federation^bDirectorate of Taimyr Nature Reserves, Talnakhskaia St., 22, entrance 2, 663300, Norilsk, Russian Federation^cSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaja St., 190000, Saint-Petersburg, Russian Federation

Introduction: The complexity of recognition and counting of objects in a photographic image is directly related to variability of related factors: physical difference of objects from the same class, presence of images similar to objects to be recognized, non-uniform background, change of shooting conditions and position of the objects when the photo was taken. Most challenging are the problems of

identifying people in crowds, animals in natural environment, cars from surveillance cameras, objects of construction and infrastructure on aerial photo images, etc. These problems have their own specific factor space, but the methodological approaches to their solution are similar. **Purpose:** The development of methodologies and software implementations solving the problem of recognition and counting of objects with high variability, on the example of reindeer recognition in the natural environment. **Methods:** Two approaches are investigated: feature-based recognition based on binary pixel classification and reference-based recognition using convolutional neural networks. **Results:** Methodologies and programs have been developed for pixel-by-pixel recognition with subsequent binarization, image clustering and cluster counting and image recognition using the convolutional neural network of Mask R-CNN architecture. The network is first trained to recognize animals as a class from the array of MS COCO dataset images and then trained on the array of aerial photographs of reindeer herds. Analysis of the results shows that feature-based methods with pixel-by-pixel recognition give good results on relatively simple images (recognition error 10–15%). The presence of artifacts on the image that are close to the characteristics of the reindeer images leads to a significant increase in the error. The convolutional neural network showed higher accuracy, which on the test sample was 82%, with no false positives. **Practical relevance:** A software prototype has been created for the recognition system based on convolutional neural networks with a web interface, and the program itself has been put into limited operation.

Keywords — recognition, convolutional neural networks, aerial images, reindeer.

For citation: Mikhailov V. V., Kolpaschikov L. A., Sobolevskii B. A., Soloviev N. V., Yakushev G. K. Methodological approaches and algorithms for recognizing and counting animals in aerial photographs. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 20–32 (In Russian). doi:10.31799/1684-8853-2021-5-20-32

References

- Zirjanov V. A., Pavlov B. M., Jakushkin G. D. *Ekologitceskie osnovi utcheta chislennosti promislovih zivotnih v tundrovoi zone Taimyra*. In: *Problemi ohotnichogo hoziaistva Krasnoyarskogo kpaia* [Ecological basics of counting the number of hunting animals in the tundra zone of Taimyr. In: Meeting materials “Problems of hunting in Krasnoyarsk region”]. Krasnoyarsk, 1971. Pp. 70–72 (In Russian).
- Kolpashikov L., Mikhailov V., and Russell D. E. The role of harvest, predators, and socio-political environment in the dynamics of the Taimyr wild reindeer herd with some lessons for North America. *Ecology and Society*, 2015, vol. 20, iss. 1, article 9. Available at: <https://www.ecologyandsociety.org> (accessed 13 October 2020).
- Chelintsev N. G. *Matematicheskie osnovi utchea zivotnih* [Mathematical basis of animal counting]. Moscow, GU Centrohtkontrol Publ., 2000. 431 p. (In Russian).
- Soloviev N. V., Sergeev M. B. Image recognition of a set of homogeneous objects with high variability. *Nautshnaia sessija GUAP: sbornik dokladov: v 3 chastiah. Chast II: Tehniceskie Nauki* [Scientific session of SUAE: collection of reports: in 3 parts. Part II: Technical Sciences]. Saint-Petersburg, GUAP Publ., 2019, pp. 427–430 (In Russian).
- Simon J. D. *Prince computer vision: Models, learning, and inference*. Cambridge, Cambridge University Press, 2012. 598 p.
- Ayyadevara K., Reddy Y. *Modern computer vision with PyTorch: Explore deep learning concepts and implement over 50 real-world image applications*. Birmingham, Packt Publishing, 2020. 824 p.
- Norouzzadeh M. S., Nguyen A., Kosmala M., Swanson A., Palmer M., Packer C., Clune J. Automated animal identification using deep learning techniques. *Proc. of the Nat. Acad. of Sciences of the USA*, 2018, June, vol. 115, iss. 25, pp. 5716–5725. doi.org/10.1073/pnas.17193671155
- Lecun Y., Boser B., Denker J. S., Henderson D., Howard R. E., Hubbard W., Jackel L. D. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1989, vol. 1, iss. 4, pp. 541–551.
- Duporse J., Isupova O., Reece S., Macdonald D., Way T. Using very high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. *bioRxiv — the preprint server for biology*. doi: <https://doi.org/10.1002/rse2.195>
- Open Data Science. Kaggle: kak nashi setochki schitali morskih l'vov na Aleutskih ostrovah* [Open Data Science. Kaggle: how our nets counted sea lions on Aleutian Islands]. Available at: <https://habr.com/ru/company/ods/blog/337548/> (accessed 20 November 2020) (In Russian).
- Seek by iNaturalist*. Available at: https://www.inaturalist.org/pages/seek_app (accessed 21 November 2020).
- Tou J., Gonzales R. *Pattern recognition principles*. Addison-Wesley Publishing Co, 1974. 378 p.
- Gonzales R., Woods R. *Digital image processing*. Prentice Hall, 2002. 822 p.
- Sergeev M. B., et al. *Programma dla podscheta odnotypnih ob'ektov na shtom izobrajenii* [A programme for counting of similar objects in a compressed image]. State Registration Certificate for a Computer Programme, no. 2019666902, 2019.
- Mikhailov V. V., Harin Y. V. On the developing of an animal recognition and counting system for aerial photographs. 1. Analysis of recognition methods. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2011, no. 2, pp. 22–28 (In Russian).
- Plataea A. S. Tree segmentation on an image using an incomplete thresholding method with a global threshold. *Reshetnevskie chteniya*, 2018, vol. 2, pp. 291–293 (In Russian).
- LeCun Y., Haffner P., Bottou L., Bengio Y. *Object Recognition with Gradient-Based Learning*. In: *Shape, Contour and Grouping in Computer Vision*. Springer, 1999. Pp. 319–345.
- He K., Gkioxari G., Dollar P., Girshick R. *Mask R-CNN*. In: *Computer Vision and Patter Recognition*. Cornell University, 2017. <https://arxiv.org/abs/1703.06870>.
- Ganesh P., Volle K., Burks T. F., Mehta S. S. Deep orange: Mask R-CNN-based orange detection and segmentation. *IF-AC-PapersOnLine*, 2019, vol. 52, iss. 30, pp. 70–75. <https://doi.org/10.1016/j.ifacol.2019.12.499>
- Zhao G., Hu J., Xiao W., Zou J. A mask R-CNN based method for inspecting cable brackets in aircraft. *Chinese Journal of Aeronautics*, 2020. <https://doi.org/10.1016/j.cja.2020.09.024>
- Lin T. Y., Maire M., Belongie S., Hays J., Perona P., Ramanan D., Dollar P., Zitnick C. L. Microsoft COCO: Common objects in context. *Computer Vision — ECCV 2014*, 2014, part 5, pp. 740–755.
- Patterson G., Hays J. COCO attributes: Attributes for people, animals, and objects. *Computer Vision — ECCV 2016*, 2016, part 6, pp. 85–100.
- Srivastava S., Divekar A. V., Anilkumar C., Naik I., Kulkarni V., Pattabiraman V. Comparative analysis of deep learning image detection algorithms. *Journal of Big Data*, 2021, no. 8, article 66. doi:10.1186/s40537-021-00434-w
- Weiss K., Khoshgoftar T. M., Wang D. D. A survey of transfer learning. *Journal of Big Data*, 2016, vol. 3, article 9. doi:10.1186/s40537-016-0043-6
- Niknejad N., Ismail W., Ghani I., Nazari B., Bahari M., Ab Razak Bin Che Hussin. Understanding Service-Oriented Architecture (SOA): A systematic literature review and directions for further investigation. *Information Systems*, 2020, vol. 91, article 101491.
- Yang L., Fan Y., Xu N. Video instance segmentation. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019, pp. 5188–5197.

Влияние шума и частоты дискретизации на погрешность дискретного представления изображений

А. Л. Тимофеев^а, канд. техн. наук, доцент, [orcid.org / 0000-0003-2137-8030](https://orcid.org/0000-0003-2137-8030), a_l_t@inbox.ru

А. Х. Султанов^а, доктор техн. наук, профессор, [orcid.org / 0000-0002-2830-3498](https://orcid.org/0000-0002-2830-3498)

^аУфимский государственный авиационный технический университет, Карла Маркса ул., 12, Уфа, 450008, РФ

Введение: цифровая регистрация изображений сопровождается не только появлением погрешности, обусловленной конечным пространственным разрешением фотоматрицы, но и воздействием шума, вклад которого в полную погрешность уменьшается при увеличении апертуры фотосенсоров в матрице. Таким образом, изменение частоты дискретизации оказывает противоположное влияние на погрешность дискретизации и погрешность, вызванную шумом. **Цель:** определение оптимальной частоты дискретизации изображений, обеспечивающей минимальную погрешность дискретизации при наличии шума. **Результаты:** исследована зависимость погрешности дискретного представления изображения от частоты дискретизации и уровня шума. Проведено моделирование процесса дискретизации изображения в присутствии шума. Построены зависимости среднеквадратического отклонения погрешности дискретизации, вызванной усечением спектра (уменьшением полосы пропускания фильтра нижних частот), и шумовой составляющей погрешности от частоты дискретизации. Сформулирована теорема о верхней границе теоремы отсчетов: при дискретизации функции конечной длительности при наличии шума существует конечное минимальное значение погрешности дискретизации, определяемое формой спектра функции и уровнем шума. **Практическая значимость:** результаты исследований целесообразно использовать при выборе фотоматрицы по числу пикселей для регистрации изображений в присутствии шума, а также при выборе полосы пропускания фильтра нижних частот для первичной обработки цифрового изображения.

Ключевые слова — дискретизация изображений, частота дискретизации, шум, погрешность дискретизации.

Для цитирования: Тимофеев А. Л., Султанов А. Х. Влияние шума и частоты дискретизации на погрешность дискретного представления изображений. *Информационно-управляющие системы*, 2021, № 5, с. 33–39. doi:10.31799/1684-8853-2021-5-33-39

For citation: Timofeev A. L., Sultanov A. Kh. Influence of noise and sampling rate on the discrete image representation error. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 33–39 (In Russian). doi:10.31799/1684-8853-2021-5-33-39

Введение

Преобразование реального «физического» изображения в цифровую форму состоит из двух процедур — дискретизации и квантования, каждая из которых сопровождается появлением специфических погрешностей. Далее рассматривается дискретизация изображений по пространственным координатам и свойственные ей погрешности, определяющие отличие цифрового изображения от исходного аналогового и, соответственно, устанавливающие предел точности последующей цифровой обработки.

Цифровая регистрация изображений всегда сопровождается появлением шума — дефектов изображения, вносимых фотосенсорами и электронным оборудованием, обусловленных несовершенством технологий и фотонной природой света.

Рассмотрим задачу цифровой регистрации изображения как процесс пространственной дискретизации непрерывной функции яркости исходного изображения. Существенным условием

задачи является конечный размер изображения и, соответственно, бесконечный спектр пространственных частот. Восстановление пространственно-дискретного изображения производится с помощью интерполяционных функций конечной длительности, поэтому в восстановленном изображении появляются ошибки — ошибки дискретизации.

Методы минимизации погрешности дискретизации сигналов в цифровой связи рассмотрены, например, в работе [1]. Цифровая регистрация изображений имеет свои особенности [2]. В реальных системах дискретизации изображений в отличие от идеального случая выполняются следующие условия:

- дискретизирующая решетка имеет конечные размеры;
- ширина дискретизирующих импульсов заметно отличается от нуля.

Практически операция дискретизации осуществляется измерением сигнала с помощью датчика, который проводит усреднение по некоторой конечной площадке (апертуре).

Требования к величине погрешности дискретизации определяются областью применения и стоящими задачами. Достаточно широкой областью обработки изображений является цифровая голография. Применение цифровой голографии открывает широкие возможности для качественного и, более того, точного количественного анализа свойств объекта, таких как смещение точек поверхности при анализе деформаций, определение формы объекта, измерение коэффициента преломления в прозрачных средах, изучение траекторий частиц, микроскопия, радиоголография, помехоустойчивое кодирование и многое другое [3–11]. В упомянутых случаях изменение состояния изучаемого объекта приводит к изменению объектной волны и получаемой интерференционной картины, которая записывается цифровым образом. Влияние дискретности на точность воспроизведения цифровых изображений рассмотрено во многих работах, например, в [12–15], но в них основной задачей является выбор периода дискретизации изображения исходя из ширины пространственного спектра и не учитывается наличие шума в изображении.

Существует широкий круг задач, в которых требуется получение изображений особо высокого пространственного разрешения: медицина [16, 17], астрономия и микроскопия [18], обработка данных дистанционного зондирования Земли [19], компьютерное зрение [20], криминалистика [21] и т. д. В [22] рассмотрен метод оптимального линейного сверхразрешающего восстановления изображений, в котором учтено наличие шума, но не исследована взаимосвязь частоты пространственной дискретизации и уровня шума, фиксируемого в цифровом изображении.

Результаты моделирования

Главным фактором, определяющим выбор частоты дискретизации, является ширина пространственного спектра регистрируемого изображения. Нижнюю границу частоты дискретизации определяет теорема отсчетов, требующая в идеальном случае частоты дискретизации не менее удвоенной граничной частоты спектра изображения. При невыполнении этого условия возникает алиасинг — эффект наложения соседних по оси частот копий дискретного спектра, приводящий к искажению дискретного изображения, в частности, появлению муара. При дискретизации изображений со спектром, не имеющим четко выраженной граничной частоты, выбор частоты дискретизации часто сводится к максимальному значению, которое доступно в данной ситуации с точки зрения аппаратных возможностей, так как действует однозначное правило:

чем больше частота дискретизации, тем меньше погрешность дискретизации, вызванная потерей части спектра сигнала выше половины частоты дискретизации.

При рассмотрении двумерных изображений используется их описание в матричной форме, соответствующей числу и расположению фотосенсоров в прямоугольной фотоматрице. Частота дискретизации по каждой координате определяется шагом сенсоров по строкам и столбцам матрицы. В большинстве случаев шаг их размещения по обеим координатам одинаков, поэтому для уменьшения объема вычислений при исследовании процесса дискретизации достаточно рассматривать одну строку изображения.

При дискретизации с частотой ω_d одномерной функции яркости $s(t)$ изображения размером t_s , имеющей бесконечный спектр $S(j\omega)$, и последующем восстановлении функции $s(t)$ с помощью фильтра нижних частот (ФНЧ) с частотой среза $\omega_d/2$ происходит подавление спектра выше частоты $\omega_d/2$, т. е. восстановленное изображение имеет спектр

$$S_d(j\omega) = \{S(j\omega), -\omega_d/2 < \omega < \omega_d/2; S_d(j\omega) = 0, \omega \leq -\omega_d/2, \omega \geq \omega_d/2\}.$$

Погрешность, обусловленная искажением функции $s(t)$ вследствие усечения спектра, является методической погрешностью, возникающей при дискретизации любой физически реализуемой функции. Эту погрешность далее будем называть погрешностью дискретизации.

Проведем анализ зависимости погрешности дискретизации от частоты дискретизации для элементарного изображения, содержащего одну узкую полосу на черном фоне, путем моделирования в среде MATLAB. Функция яркости выбранного изображения имеет вид

$$s(nT) = \sin(\pi n/16 + \pi) \text{ при } n_1 < n < n_2; s(nT) = 0 \text{ при } n \leq n_1, n \geq n_2,$$

где $n = 1, \dots, 2048$; $T = t_s/2048$; $n_1 = 1016$; $n_2 = 1032$; частота дискретизации $\omega_d = 1/T = 2048/t_s$.

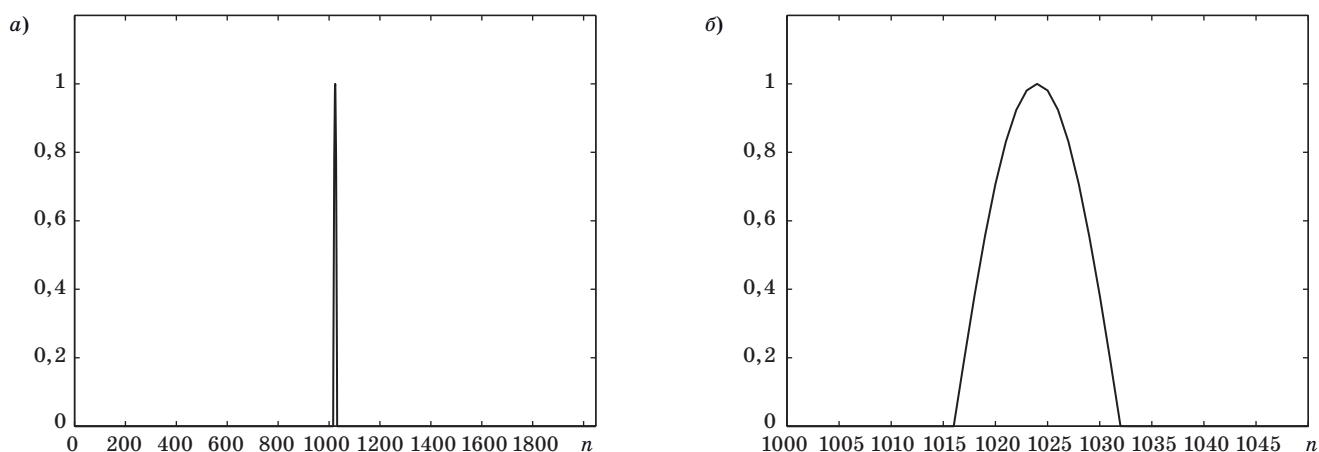
График функции $s(nT)$ приведен на рис. 1.

Вещественная часть спектра $S(k)$ функции $s(nT)$ представлена на рис. 2, а.

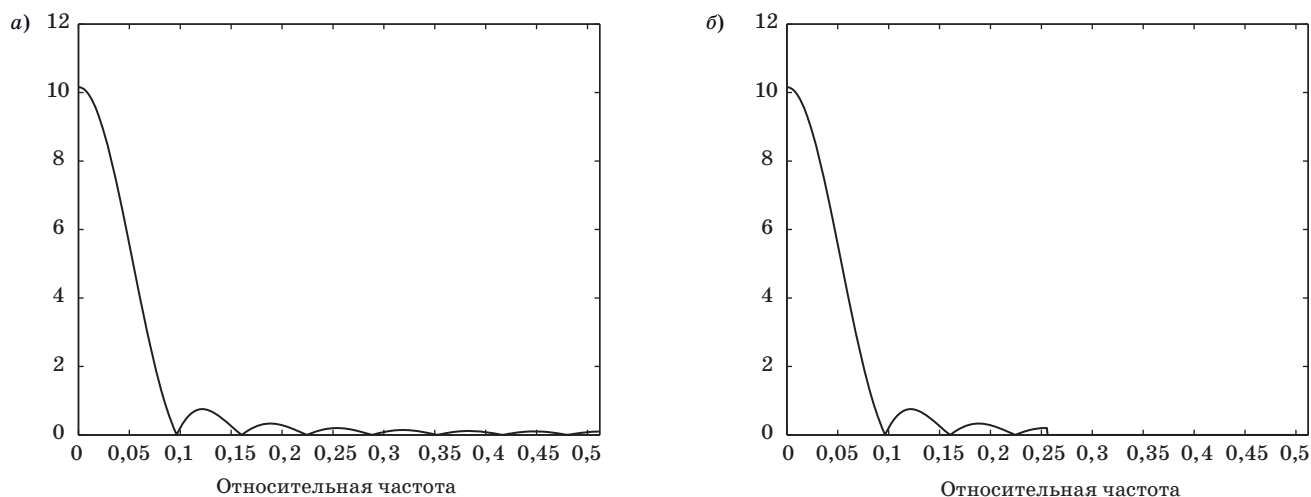
Уменьшим в два раза частоту дискретизации (с соответствующим уменьшением частоты среза ФНЧ), тогда спектр $S_d(k)$ примет вид, показанный на рис. 2, б.

Фрагмент функции $s_d(nT)$ демонстрирует искажения, возникающие в результате подавления высокочастотной части спектра (рис. 3).

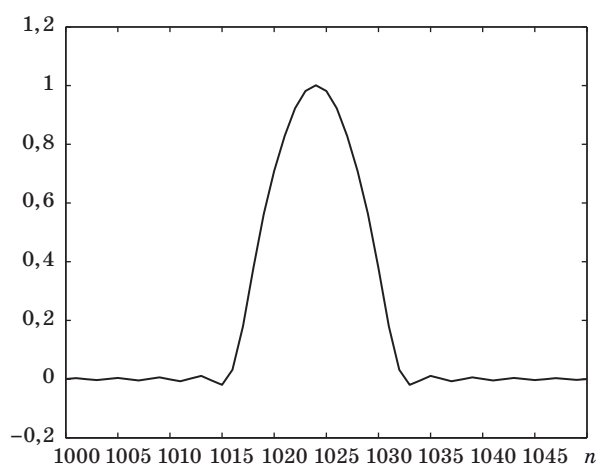
Оценим погрешность дискретизации с помощью среднеквадратического отклонения функции $s_d(nT)$ от $s(nT)$.



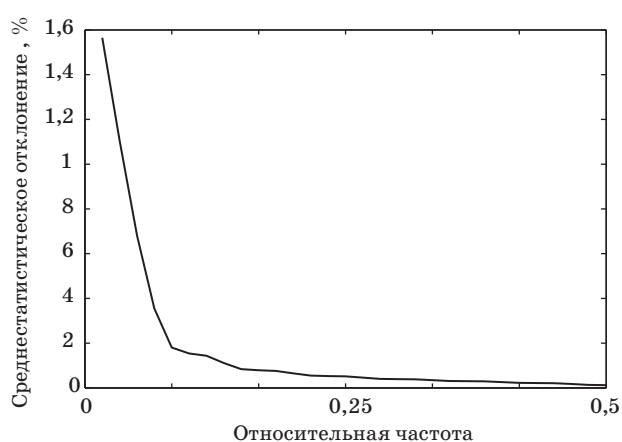
■ **Рис. 1.** Функция яркости изображения: *a* — полный вид функции $s(nT)$; *б* — увеличенный фрагмент
 ■ **Fig. 1.** Image brightness function: *a* — full view of the function $s(nT)$; *б* — enlarged fragment



■ **Рис. 2.** Спектр изображения: *a* — исходный; *б* — после фильтрации
 ■ **Fig. 2.** Image spectrum: *a* — original; *б* — after filtering



■ **Рис. 3.** Функция яркости восстановленного изображения
 ■ **Fig. 3.** Brightness function of the reconstructed image

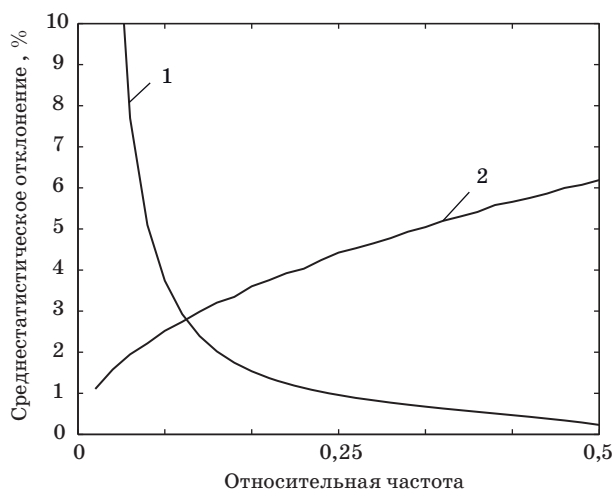


■ **Рис. 4.** Зависимость погрешности дискретизации от частоты дискретизации
 ■ **Fig. 4.** Dependence of the sampling error on the sampling frequency

Зависимость погрешности дискретизации от частоты дискретизации ω_d показана на рис. 4.

Кроме физической нереализуемости изображений с ограниченным спектром, в условие теоремы отсчетов входит и другая нереализуемая идеализация, которая обычно в этой связи не упоминается, — отсутствие шума. В реальных изображениях всегда присутствует шум, но его вклад в полную погрешность цифрового изображения зависит от плотности размещения фотосенсоров в матрице и их апертуры [2]. Увеличение размера фотосенсоров и, соответственно, уменьшение шага их размещения приводит к сглаживанию дискретного изображения и уменьшению шума. Таким образом, изменение частоты дискретизации оказывает противоположное влияние на погрешность дискретизации и погрешность, вызванную шумом. Для исследования этого явления было проведено моделирование в среде MATLAB процесса дискретизации функции $s(nT)$, приведенной на рис. 1, в присутствии шума и последующего восстановления — формирования функции $s_d(nT)$ (см. рис. 3). В качестве источника шума использовался генератор белого шума `awgn` — встроенная функция MATLAB. Рассчитывалось отдельно среднеквадратическое отклонение погрешности дискретизации, вызванной усечением спектра (уменьшением полосы пропускания ФНЧ), и шумовая составляющая погрешности — среднеквадратическое отклонение шума после прохождения ФНЧ.

Зависимости этих погрешностей от частоты дискретизации при отношении сигнал/шум 40 дБм показаны на графиках (рис. 5).



■ Рис. 5. Зависимость среднеквадратических погрешностей от частоты дискретизации: 1 — погрешность дискретизации; 2 — шумовая составляющая

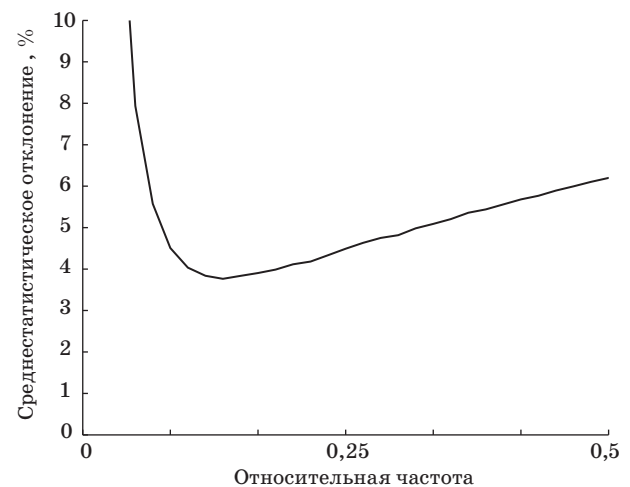
■ Fig. 5. Dependence of the root mean square errors on the sampling frequency: 1 — sampling error; 2 — noise component

Зависимость среднеквадратического значения полной погрешности (дискретизации и шумовой составляющей) от частоты дискретизации приведена на рис. 6.

Из графика видно, что полная погрешность дискретизации имеет выраженный минимум при некоторой частоте дискретизации, зависящей от формы функции яркости изображения и, соответственно, скорости убывания ее спектра и интенсивности шума. Отсюда следует, что при выборе периода дискретизации изображения (выборе шага расположения элементов фотоматрицы) надо учитывать не только характеристики фиксируемого изображения, но и уровень шума. Таким образом, можно сформулировать теорему о верхней границе теоремы отсчетов для функций ограниченной длительности при наличии шума: при дискретизации функции конечной длительности при наличии шума существует конечное минимальное значение погрешности дискретизации, определяемое формой спектра функции и уровнем шума.

Варианты практического применения полученных результатов зависят от области применения и условий получения изображений:

1. Регистрация однотипных изображений в схожих шумовых условиях, например аэрофотосъемка. В этом случае параметры изображения и шума учитываются при выборе фотоматрицы по числу пикселей. Второй способ: при использовании матрицы с большим числом пикселей выбирается ФНЧ для первичной обработки цифрового изображения с полосой пропускания, равной половине оптимальной частоты дискретизации, обеспечивающей минимальную суммарную погрешность.



■ Рис. 6. Зависимость полной погрешности от частоты дискретизации

■ Fig. 6. Dependence of the total error on the sampling frequency

2. Фотоматрица универсального применения. Для минимизации погрешности дискретизации возможна оценка уровня шума и характера изображения путем цифровой предобработки изображения, фиксируемого фотоматрицей, и настройка адаптивного ФНЧ, который формирует выходное изображение при оптимальной граничной частоте.

Заключение

В работе проведено исследование полной погрешности дискретизации изображений в присутствии шума. Показано, что изменение частоты дискретизации оказывает противоположное влияние на погрешность дискретизации и погрешность, вызванную шумом. В результате моделирования процесса дискретизации изображения в присутствии белого шума установлено,

что существует оптимальное значение частоты дискретизации, зависящее от формы пространственного спектра изображения и уровня шума. Полученные результаты изменяют подход к выбору фотоматрицы по количеству пикселей и дают возможность минимизировать погрешность дискретизации.

Финансовая поддержка

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 22-29-00041, а также Министерства науки и высшего образования РФ в рамках выполнения работ по Государственному заданию ФГБОУ ВО УГАТУ # FEUE-2020-0007 по теме «Теоретические основы моделирования и семантического анализа процессов преобразования вихревых электромагнитных полей в инфокоммуникационных системах».

Литература

1. Feher D. *Wireless Digital Communication: Modulation and Spread Spectrum Applications*. Prentice Hall PTR, New Jersey, 1995. 520 p.
2. Гужов В. И., Марченко И. О., Хайдуков Д. С., Ильных С. П. Дискретизация изображений в реальных системах с помощью обобщенных функций. *Автоматика и программная инженерия*, 2016, № 4 (18), с. 45–52.
3. Корешев С. Н., Смородинов Д. С., Никаноров О. В. Влияние дискретности синтезированных и цифровых голограмм на их изображающие свойства. *Компьютерная оптика*, 2016, т. 40, № 6, с. 793–801. doi:10.18287/2412-6179-2016-40-6-793-801
4. Шойдин С. А. Синтез голограмм на приемном конце канала связи с объектом голографирования. *Компьютерная оптика*, 2020, т. 44, № 4, с. 547–551. doi:10.18287/2412-6179-СО-694
5. Тимофеев А. Л., Султанов А. Х. Построение помехоустойчивого кода на базе голографического представления произвольной цифровой информации. *Компьютерная оптика*, 2020, т. 44, № 6, с. 978–984. doi:10.18287/2412-6179-СО-739
6. Корешев С. Н., Никаноров О. В., Смородинов Д. С. Изображающие свойства дискретных голограмм. I. Влияние дискретности голограммы на восстановленное изображение. *Оптический журнал*, 2014, т. 81, № 3, с. 14–19.
7. Корешев С. Н., Никаноров О. В., Смородинов Д. С. Изображающие свойства дискретных голограмм. II. Влияние модификации структуры голограммы и высокой, превышающей частоту Найквиста, несущей пространственной частоты голограммной структуры на восстановленное изображение. *Оптический журнал*, 2014, т. 81, № 4, с. 48–53.
8. Тимофеев А. Л. Использование голографического кодирования для повышения помехоустойчивости каналов связи. *ИТпортал*, 2018, т. 18, № 2. <http://itportal.ru/science/tech/ispolzovanie-golograficheskogo-kodi/> (дата обращения: 25.05.2019).
9. Пат. 2450439 РФ. Способ кодирования и декодирования цифровых данных, основанный на принципах цифровой голографии/ С. Л. Френкель (РФ), Ш. Долев (IL). № 2010145892/08; заявл. 11.11.2010; опубл. 10.05.2012, Бюл. № 13. 11 с.
10. Timofeev A. L., Sultanov A. Kh. Holographic method of error-correcting coding. *Proc. SPIE 11146 "Optical Technologies for Telecommunications 2018", Proceedings Volume 11146, 111461A*, 2019. doi:10.1117/12.2526922
11. Timofeev A. L., Sultanov A. Kh., Filatov P. E. Holographic method for storage of digital information. *Proc. SPIE 11516 "Optical Technologies for Telecommunications 2019", 1151604*, 2020. doi:10.1117/12.2566329
12. Чапурский В. В. Получение радиоголографических изображений объектов на основе разреженных антенных решеток типа ММО с одночастотным и многочастотным излучением. *Вестник МГТУ им. Н. Э. Баумана. Сер. Приборостроение*, 2011, № 4, с. 72–91.
13. Семенчик В. Г., Пахомов В. А. Радиоголографическая система формирования многочастотных изображений. *Электроника*. 2004, № 1, с. 50–51.
14. Pahomov V., Semenchik V. G., Kurilo S. Reconstructing reflecting object images using born approximation. *Proc. of 35th European Microwave Conference*, 2005, CNIT la Defense, Paris, France, pp. 1375–1378.
15. Голунов В. А., Коротков В. А., Коротков К. В. Обоснование возможности получения радиоизображе-

- ний объектов методом одномерных голограмм. *Радиотехника и электроника*, 2019, т. 64, № 1, с. 38–43. doi:10.1134/S0033849419010066
16. Isaac J. S., Kulkarni R. Super resolution techniques for medical image processing. *2015 International Conference on Technologies for Sustainable Development (ICTSD)*, 2015, 7095900. doi:10.1109/ICTSD.2015.7095900
 17. Sano Y., Mori T., Goto T., Hirano S., Funahashi K. Super-resolution method and its application to medical image processing. *2017 IEEE 6th Global Conference on Consumer Electronics (GCCE)*, 2017, pp. 1–2. doi:10.1109/GCCE.2017.8229301
 18. Chainais P., Pfennig P., Leray A. Quantitative control of the error bounds of a fast super-resolution technique for microscopy and astronomy. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 2853–2857. doi:10.1109/ICASSP.2014.6854121
 19. Shen H., Ng M. K., Li P., Zhang L. Super-resolution reconstruction algorithm to MODIS remote sensing images. *Computer Journal*, 2009, vol. 52, iss. 1, pp. 90–100. doi:10.1093/comjnl/bxm028
 20. Shi F., Yuan J., Zhu X. Fast super-resolution reconstruction for video-based pattern recognition. *2008 Fourth International Conference on Natural Computation*, 2008, vol. 4, pp. 135–139. doi:10.1109/ICNC.2008.553
 21. Zamani N. A., Darus M. Z. A., Abdullah S. N. H., Nordin M. J. Multiple-frames super-resolution for closed circuit television forensics. *2011 International Conference on Pattern Analysis and Intelligence Robotics*, 2011, vol. 1, pp. 36–40. doi:10.1109/ICPAIR.2011.5976908
 22. Максимов А. И., Сергеев В. В. Метод оптимального линейного сверхразрешающего восстановления изображений. *Компьютерная оптика*, 2021, т. 45, № 5, с. 692–701. doi:10.18287/2412-6179-CO-909

UDC 621.391

doi:10.31799/1684-8853-2021-5-33-39

Influence of noise and sampling rate on the discrete image representation errorA. L. Timofeev^a, PhD, Tech., Associate Professor, orcid.org / 0000-0003-2137-8030, a_l_t@inbox.ruA. Kh. Sultanov^a, Dr. Sc., Tech., Professor, orcid.org / 0000-0002-2830-3498^aUfa State Aviation Technical University, 12, Karl Marks St., 450008, Ufa, Russian Federation

Introduction: Digital registration of images is accompanied not only by an error caused by finite spatial resolution of the photo matrix, but also by the effect of noise whose contribution to the total error decreases with an increase in the aperture of the photosensors in the matrix. Thus, changing the sampling rate has the opposite effect on the sampling error and on the error caused by the noise. **Purpose:** Finding the optimal image sampling rate which would provide the minimum sampling error in the presence of noise. **Results:** We have studied how an image discrete representation error depends on the sampling frequency and noise level. The image sampling process in the presence of noise was simulated in the MATLAB environment. The dependencies of the root-mean-square deviation of the sampling error caused by spectrum truncation (decrease in the passband of the low-pass filter) and the noise component of the error on the sampling frequency were plotted. A theorem is formulated on the upper bound of the sampling theorem: when sampling a function of finite duration in the presence of noise, there is a finite minimum value of the sampling error which is determined by the shape of the spectrum of the function and the noise level. **Practical relevance:** It is advisable to use the research results when choosing a photomatrix by the number of pixels for recording images in the presence of noise, as well as when choosing a low-pass filter passband for primary processing of a digital image.

Keywords — image sampling, sampling rate, noise, sampling error.

For citation: Timofeev A. L., Sultanov A. Kh. Influence of noise and sampling rate on the discrete image representation error. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 33–39 (In Russian). doi:10.31799/1684-8853-2021-5-33-39

Financial support

This work was partially supported by Russian Science Foundation (project No. 22-29-00041) and by the Ministry of Science and Higher Education of the Russian Federation for research under the State Assignment of FSBEI HE USATU No. FEUE-2020-0007 on the topic “Theoretical foundations of modeling and semantic analysis of the processes of transformation of vortex electromagnetic fields in infocommunication systems”.

References

1. Feher D. *Wireless Digital Communication: Modulation and Spread Spectrum Applications*. Prentice Hall PTR, New Jersey, 1995. 520 p.
2. Guzhov V. I., Marchenko I. O., Hajdukov D. S., Ilynikh S. P. Sampling the image in real systems using generalized functions. *Automatics & Software Engineering*, 2016, no. 4 (18), pp. 45–52 (In Russian).
3. Koreshev S. N., Smorodinov D. S., Nikanorov O. V. Influence of the discreteness of synthesized and digital holograms on their imaging properties. *Computer Optics*, 2016, vol. 40, no. 6, pp. 793–801. doi:10.18287/2412-6179-2016-40-6-793-801 (In Russian).
4. Shoidin S. A. Synthesis of holograms at the receiving end of the communication channel with the holographic object. *Computer Optics*, 2020, vol. 44, no. 4, pp. 547–551. doi:10.18287/2412-6179-CO-694 (In Russian).
5. Timofeev A. L., Sultanov A. Kh. Building a noise-tolerant code based on a holographic representation of arbitrary digital information. *Computer Optics*, 2020, vol. 44, no. 6, pp. 978–984. doi:10.18287/2412-6179-CO-739 (In Russian).
6. Koreshev S. N., Nikanorov O. V., Smorodinov D. S. Imaging properties of discrete holograms. I. Influence of the discreteness of the hologram on the reconstructed image. *Optical Journal*, 2014, vol. 81, no. 3, pp. 14–19 (In Russian).

7. Koreshev S. N., Nikanorov O. V., Smorodinov D. S. Imaging properties of discrete holograms. II. Influence of the modification of the structure of the hologram and the high, exceeding the Nyquist frequency, carrier of the spatial frequency of the hologram structure on the reconstructed image. *Optical Journal*, 2014, vol. 81, no. 4, pp. 48–53 (In Russian).
8. Timofeev A. L. The use of holographic coding to increase noise immunity of communication channels. *ITportal*, 2018, vol. 18, no. 2. Available at: <http://itportal.ru/science/tech/ispolzovanie-golograficheskogo-kodi/> (accessed 25 May 2019) (In Russian).
9. Frenkel S. L., Dolev S. *Sposob kodirovaniya I dekodirovaniya cifrovyyh dannyh, osnovannyy na principah cifrovoi golografii*. [A method for encoding and decoding digital data based on the principles of digital holography]. Patent RU, no. 2450439, 2010.
10. Timofeev A. L., Sultanov A. Kh. Holographic method of error-correcting coding. *Proc. SPIE 11146 "Optical Technologies for Telecommunications 2018", Proceedings Volume 11146, 111461A*, 2019. doi:10.1117/12.2526922
11. Timofeev A. L., Sultanov A. Kh., Filatov P. E. Holographic method for storage of digital information. *Proc. SPIE 11516 "Optical Technologies for Telecommunications 2019", 1151604*, 2020. doi:10.1117/12.2566329
12. Chapurskii V. V. Receiving the radio-holographic images of objects on the basis of disperse MIMO-type antenna arrays with the single-frequency and multi-frequency radiations. *Herald of the Bauman Moscow State Technical University. Series Instrument Engineering*, 2011, no. 4, pp. 72–91 (In Russian).
13. Semenchik V. G., Pahomov V. A. Radio holographic multi-frequency imaging system. *Electronics*, 2004, no. 1, pp. 50–51 (In Russian).
14. Pahomov V. A., Semenchik V. G., Kurilo S. Reconstructing reflecting object images using born approximation. *Proc. of 35th European Microwave Conference*, 2005, CNIT la Defense, Paris, France, pp. 1375–1378.
15. Golunov V. A., Korotkov V. A., Korotkov K. V. Substantiation of the possibility of obtaining radio images of objects by the method of one-dimensional holograms. *Journal of Communications Technology and Electronics*, 2019, vol. 64, no. 1, pp. 38–43. doi:10.1134/S0033849419010066 (In Russian).
16. Isaac J. S., Kulkarni R. Super resolution techniques for medical image processing. *2015 International Conference on Technologies for Sustainable Development (ICTSD)*, 2015, 7095900. doi:10.1109/ICTSD.2015.7095900
17. Sano Y., Mori T., Goto T., Hirano S., Funahashi K. Super-resolution method and its application to medical image processing. *2017 IEEE 6th Global Conference on Consumer Electronics (GCCE)*, 2017, pp. 1–2. doi:10.1109/GCCE.2017.8229301
18. Chainais P., Pfennig P., Leray A. Quantitative control of the error bounds of a fast super-resolution technique for microscopy and astronomy. *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 2853–2857. doi:10.1109/ICASSP.2014.6854121
19. Shen H., Ng M. K., Li P., Zhang L. Super-resolution reconstruction algorithm to MODIS remote sensing images. *Computer Journal*, 2009, vol. 52, iss. 1, pp. 90–100. doi:10.1093/comjnl/bxm028
20. Shi F., Yuan J., Zhu X. Fast super-resolution reconstruction for video-based pattern recognition. *2008 Fourth International Conference on Natural Computation*, 2008, vol. 4, pp. 135–139. doi:10.1109/ICNC.2008.553
21. Zamani N. A., Darus M. Z. A., Abdullah S. N. H., Nordin M. J. Multiple-frames super-resolution for closed circuit television forensics. *2011 International Conference on Pattern Analysis and Intelligence Robotics*, 2011, vol. 1, pp. 36–40. doi:10.1109/ICPAIR.2011.5976908
22. Maksimov A. I., Sergeev V. V. Optimal super-resolution linear image recovery method. *Computer Optics*, 2021, vol. 45, no. 5, pp. 692–701. doi:10.18287 / 2412-6179-CO-909

ПАМЯТКА ДЛЯ АВТОРОВ

Поступающие в редакцию статьи проходят обязательное рецензирование.

При наличии положительной рецензии статья рассматривается редакционной коллегией. Принятая в печать статья направляется автору для согласования редакторских правок. После согласования автор представляет в редакцию окончательный вариант текста статьи.

Процедуры согласования текста статьи могут осуществляться как непосредственно в редакции, так и по e-mail (ius.spb@gmail.com).

При отклонении статьи редакция представляет автору мотивированное заключение и рецензию, при необходимости доработать статью — рецензию.

Редакция журнала напоминает, что ответственность за достоверность и точность рекламных материалов несут рекламодатели.

УДК 004.896

doi:10.31799/1684-8853-2021-5-40-50

Адаптация нейронных машин Тьюринга для задачи агрегации лингвистических оценок в нейросимволических системах поддержки принятия решений

А. В. Демидовский^а, аспирант, старший преподаватель, orcid.org/0000-0003-3605-6332, ademidovskij@hse.ru

Э. А. Бабкин^а, канд. техн. наук, заведующий лабораторией, профессор, orcid.org/0000-0003-2597-9043

^аНациональный исследовательский университет «Высшая школа экономики», Нижегородский филиал, ул. Большая Печерская, 25/12, Нижний Новгород, 603155, РФ

Введение: построение интегрированных нейросимволических систем является актуальной и сложной задачей. Построение нейросимволических систем поддержки принятия решений требует новых подходов к представлению знаний о проблемной ситуации и выражению символических рассуждений на субсимволическом уровне. **Цель:** разработка нейросетевых архитектур и методов для эффективного представления знаний в распределенном виде и субсимволических рассуждений в системах поддержки принятия решений в части алгоритмов агрегирования нечетких оценок экспертов для выбора альтернативных решений. **Методы:** представление нечетких и неопределенных оценок в распределенном виде с помощью тензорных представлений, построение обучаемой нейросетевой архитектуры для субсимволической агрегации лингвистических оценок. **Результаты:** предложены два новых метода представления лингвистических оценок в распределенном виде. Первый подход основывается на возможности перевода произвольной лингвистической оценки в численное представление и заключается в переводе этого численного представления в распределенное за счет перевода самого числа в вид битовой строки и дальнейшего формирования матрицы, хранящей распределенное представление всего выражения для агрегации оценок. Второй подход основывается на представлении лингвистической оценки в виде дерева и кодировании этого дерева с помощью метода тензорных представлений, что позволяет избежать шага перевода лингвистической оценки в численный вид и обеспечивает переход между символическим и субсимволическим представлениями лингвистических оценок без потери информации. Структурные элементы лингвистической оценки рассматриваются как заполнители с соответствующими позиционными ролями. Предложен новый субсимволический метод агрегации лингвистических оценок, который заключается в создании обучаемого нейросетевого модуля в виде нейронной машины Тьюринга. **Практическая значимость:** результаты исследования демонстрируют, как символический алгоритм агрегации лингвистических оценок может быть реализован коннективистскими, или субсимволическими, механизмами, что является необходимым требованием при построении распределенных нейросимволических систем поддержки принятия решений.

Ключевые слова – нейронные машины Тьюринга, искусственные нейронные сети, субсимволические системы, гибридные интеллектуальные системы, многокритериальное принятие решений.

Для цитирования: Демидовский А. В., Бабкин Э. А. Адаптация нейронных машин Тьюринга для задачи агрегации лингвистических оценок в нейросимволических системах поддержки принятия решений. *Информационно-управляющие системы*, 2021, № 5, с. 40–50. doi:10.31799/1684-8853-2021-5-40-50

For citation: Demidovskij A. V., Babkin E. A. Adapting Neural Turing Machines for linguistic assessments aggregation in neural-symbolic decision support systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 40–50 (In Russian). doi:10.31799/1684-8853-2021-5-40-50

Введение

Долгое время области искусственного интеллекта и анализа решений развивались параллельно несмотря на то, что обе исследуют формальные модели знаний и опыта человека [1, 2]. Системы поддержки принятия решений (СППР) были разработаны как инструмент для автоматизации сбора и анализа мнений экспертов и заинтересованных сторон в целях оказания помощи лицу, принимающему решения. При построении СППР следует учитывать многочисленные факторы, в том числе сбор неоднозначных экспертных оценок альтернативных вариантов, использование неполных данных и т. п. Такие сложности создали общую основу для междис-

циплинарной коммуникации, и для решения вышеупомянутых вопросов были разработаны различные подходы.

Одним из центральных является вопрос, как представить знания таким образом, чтобы это было эффективно с точки зрения вычислений. То, как знания представляются и моделируются человеческим мозгом, на данный момент определяет два различных течения научной мысли. Многочисленные исследователи мотивированы тем, что когнитивная система человека использует невероятно широкую нейронную сеть в качестве вычислительного двигателя, и предлагают считать, что когнитивная система человека использует распределенное представление информации и обрабатывает ее динамически и осмыс-

ленно. Такой нейронный подход к представлению и обработке информации называют субсимволическим, или коннективистским. Одновременно с этим другие исследователи советуют рассматривать познание человека как метод символических манипуляций [3], а именно, как «вид вычислений, выполняемых в конкретном типе биологической системы» [4]. В этом случае символы выступают в качестве входов в такую систему, а затем с помощью предопределенных правил и наборов инструкций преобразуются в другие символы. Такой подход называют символическим.

Субсимволические подходы по своей природе распределены и имеют высокую эффективность и надежность. Кроме того, они включают в себя важный компонент обучения, позволяя обучать субсимволические методы для решения конкретной задачи и постоянно модифицировать их по мере поступления новой информации. В то же время такие подходы сложно интерпретировать, хотя развивающаяся область, называемая объяснимым искусственным интеллектом (Explainable AI, xAI), посвящена разработке методов извлечения смысла из сложных вычислительных моделей. Другим недостатком субсимволических методов является их неспособность учитывать сложные взаимосвязи из-за ограничений современных методов распределенного представления информации [3, 4]. Некоторые достижения в этой области позволили решить ряд проблем распределенного представления, связанных с субсимволическими архитектурами: тензорные представления [5], способные представлять рекурсию в символьных структурах; дальнейшее развитие идей о более компактных распределенных представлениях с помощью голографических сокращенных представлений [6] и других методов выражения символьных вычислений. Как правило, этот класс методов называют векторными символьными архитектурами [7, 8]. С другой стороны, хотя символические подходы можно интерпретировать, они по определению являются последовательными. Важно, что символические подходы предполагают креативность правил и инструкций по обработке информации, что приводит к двум последствиям. Во-первых, информация, закодированная в символьной системе программистом, отражает его собственный опыт и мировоззрение, делая систему предвзятой. Этот вопрос был идентифицирован как проблема символьного основания (symbol grounding problem) [9]. Во-вторых, информация, хранящаяся в символической СППР, быстро теряет свою актуальность.

В результате интеграции символического и субсимволического подходов формируются такие архитектуры и системы, как ACT-R [10], CLARION [11] и SS-RICS [12, 13]. В пользу построения интегрированных систем также гово-

рит и психофизиологическая правдоподобность такой интеграции, которая была показана и установлена путем демонстрации наличия взаимосвязанных архитектур в нервной системе человека и различных биологических архитектур [3]. Построение нейросимволических интегрированных систем является актуальной задачей в области анализа решений [14, 15]. Такие системы будут сочетать субсимволические рассуждения и вычисления на коннективистском уровне с символическими рассуждениями и вычислениями на символическом уровне. Однако постоянно возникают споры о том, что значит для СППР стать нейросимволической или интегрированной [1, 15, 16]. Особое внимание уделяется тому, какие компоненты СППР могут быть делегированы на субсимволический уровень и какую функцию искусственные нейронные сети могут играть в СППР [17, 18].

Данная статья посвящена специальному аспекту СППР — агрегации оценок экспертов. Любая задача, связанная с принятием решения, описывается, в частности, проблемой, альтернативными решениями, требованиями и экспертными оценками. Агрегирование оценок является сложной задачей по целому ряду причин, включая нечеткий характер оценок, таких как лингвистические, большое число альтернативных решений и неравные степени компетентности экспертов. Агрегирование оценок является общим этапом в различных методах принятия решений, включая TOPSIS [19], ELECTRE [20] и ML-LDM [21]. Таким образом, построение нейросимволической СППР может начинаться с выражения агрегирования нечетких оценок на нейросетевом уровне. Целью данной работы является разработка нового субсимволического метода агрегации лингвистических оценок в виде обучаемого нейросетевого модуля.

Агрегация лингвистических оценок с помощью мультигранулярного оператора агрегации

Современные методы многоаспектного многоуровневого принятия решений используют в качестве базового строительного блока двухкортежную модель представления лингвистической информации — модель 2-tuple. Важной особенностью данной модели является способность выражать как качественные, так и количественные оценки. Модель 2-tuple основана на концепции символического перевода [22].

Определение 1. Структура 2-tuple состоит из пары (s_i, α) , где $s_i \in S = \{s_0, \dots, s_g\}$ — лингвистический терм (концепт); α — численное значение, результат символического перевода, значение

функции принадлежности. Данное значение показывает расстояние до ближайшего концепта $s_i \in S = \{s_0, \dots, s_g\}$, если функция принадлежности не дает точного ответа (s_i).

Определение 2. Лингвистической шкалой S [23] называется конечное упорядоченное множество термов $S = \{s_i | i = 1 \dots T\}$, удовлетворяющее следующим условиям:

- 1) если $i < j$, то s_i предшествует s_j ($s_i < s_j$);
- 2) отрицание терма определяется правилом $Neg(s_i) = s_{T-i+1}$;
- 3) пусть $s_i < s_j$, тогда объединение (дизъюнкция, связка «ИЛИ») термов определяется правилом $s_i \vee s_j = \max\{s_i, s_j\} = s_j$;
- 4) пусть $s_i < s_j$, тогда пересечение (конъюнкция, связка «И») термов определяется правилом $s_i \wedge s_j = \min\{s_i, s_j\} = s_i$.

Определение 3. Правило перевода. Пусть $S = \{s_0, \dots, s_g\}$ — лингвистическая шкала, где $g = \tau + 1$ обозначает уровень гранулярности S . Если $\beta \in [0, 1]$ — это результат символической агрегации, то существует однозначный способ построить соответствующий элемент типа 2-tuple:

$$\begin{aligned} \Delta_g &= [0, 1] \rightarrow S \times [-0,5; 0,5), \\ \Delta_g(\beta) &= (s_i, \alpha), \\ i &= \text{round}(\beta\tau), \\ \alpha &= \beta\tau - i, \alpha \in [-0,5; 0,5). \end{aligned} \quad (1)$$

Определение 4. Правило обратного перевода. Пусть $S = \{s_0, \dots, s_g\}$ — лингвистическая шкала, где $g = \tau + 1$ обозначает уровень гранулярности S . Пусть (s_i, α) — элемент типа 2-tuple, определенный на лингвистической шкале S , где $\alpha \in [-0,5; 0,5)$. Тогда есть способ преобразования данного элемента в численное представление $\beta \in [0, 1]$:

$$\begin{aligned} \Delta_g^{-1} &= S \times [-0,5; 0,5) \rightarrow [0, 1], \\ \Delta_g^{-1}(s_i, \alpha) &= (i + \alpha)/\tau. \end{aligned} \quad (2)$$

Существует множество способов агрегирования оценок, выраженных в форме 2-tuple, их обычно называют операторами: МТWA (Multigranularity 2-tuple Weighted Averaging), МНТWA (Multigranularity Hesitant 2-tuple Weighted Averaging), P2TLWA (Pythagorean 2-tuple Linguistic Weighted Averaging) [24] и т. д. Один из них, МТА, выполняет вычисление средневзвешенного значения по набору 2-tuple.

Определение 5. Оператор МТА (Multigranular 2-tuple Averaging). Пусть (b_i, α_i) — элемент типа 2-tuple, определенный на шкале S_i^g , $i = 1, 2, \dots, n$. Тогда оператор МТА определяется как

$$\begin{aligned} \text{MTA}((b_1, \alpha_1), (b_2, \alpha_2), \dots, (b_n, \alpha_n)) &= \\ = \Delta_{gh}(\sum_{j=1}^n (1/n \Delta_{gj}^{-1}(b_j, \alpha_j))). \end{aligned} \quad (3)$$

Несмотря на то, что модель 2-tuple является фундаментальной, в последние годы были предложены многочисленные методы агрегирования нечетких оценок: Hesitant Fuzzy Linguistic Term Sets (HFLTST) [25], Institutional 2-tuple [26], гибридные модели [27] и т. д. Каждый оператор обычно ассоциируется с набором арифметических операций.

Одним из возможных методов построения нейросимволических СППР является динамическое выражение арифметических операций в виде динамики нейронной сети. Настоящая статья является частью серии статей о выстраивании нейросимволической архитектуры СППР [28]. Однако здесь используется подход к проектированию СППР, отличный от рассмотренного в работе [28], где предложено создание таких нейросетевых архитектур, которые не требуют обучения и работают поверх лингвистических оценок, закодированных с помощью тензорных представлений [29]. В данной статье предлагается агрегировать лингвистические оценки в виде обучаемого нейросетевого модуля, построенного по принципам архитектуры нейронной машины Тьюринга (НМТ, Neural Turing Machine — NTM).

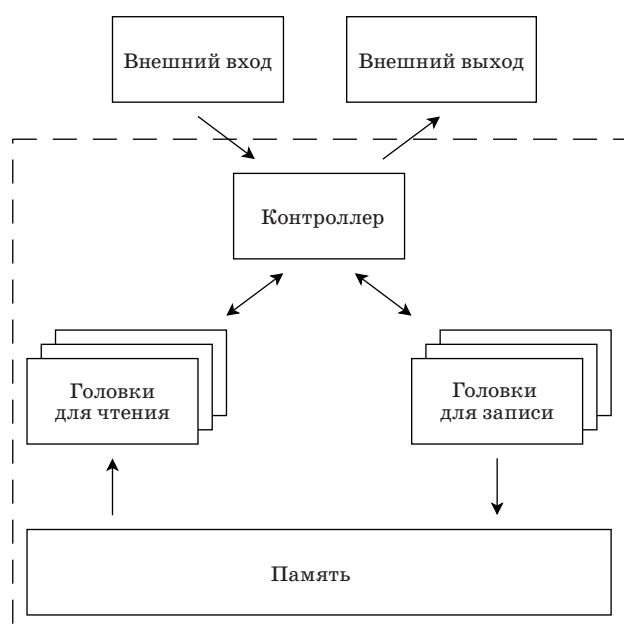
Нейронные машины Тьюринга

Архитектура

Нейронные машины Тьюринга впервые были предложены в работах [30, 31] и с тех пор получили широкое распространение в различных задачах, начиная от простых алгоритмических задач и заканчивая обучением с подкреплением [32], последовательными рекомендациями [33], транскрипцией естественного языка [34] и др.

Нейронные машины Тьюринга есть частный пример отдельного класса ИНС с дополненной памятью, которая играет важнейшую роль и является внешней по отношению к внутреннему состоянию нейросети. Для записи в эту память и чтения из нее существуют специальные абстракции, называемые записывающими и считывающими головками соответственно. Контроллер выполняет координацию этих головок для получения результата. Теоретическая архитектура НМТ показана на рис. 1.

Нейронная машина Тьюринга обучается с учителем. На этапе обучения НМТ ожидает набор векторных последовательностей в качестве входа и набор меток, который является еще одной последовательностью векторов и обозначает ожидаемый выход сети. На выходе НМТ выдает пакет векторных последовательностей, таким образом, НМТ может выполнять одну и ту же задачу над несколькими последовательностями одновременно. Доказано, что НМТ может решать мно-



■ *Рис. 1.* Общий дизайн НМТ
 ■ *Fig. 1.* General design of NTM

жество задач, таких как копирование, повтор, ассоциативный поиск, сложение и умножение двоичных чисел, задачу приоритетной сортировки. В данной работе рассматривается выражение оператора МТА для чисел, представленных в двоичном виде различным количеством бит: 4, 6, 8, 10. Общая схема подачи данных в нейронную сеть показана на рис. 2.

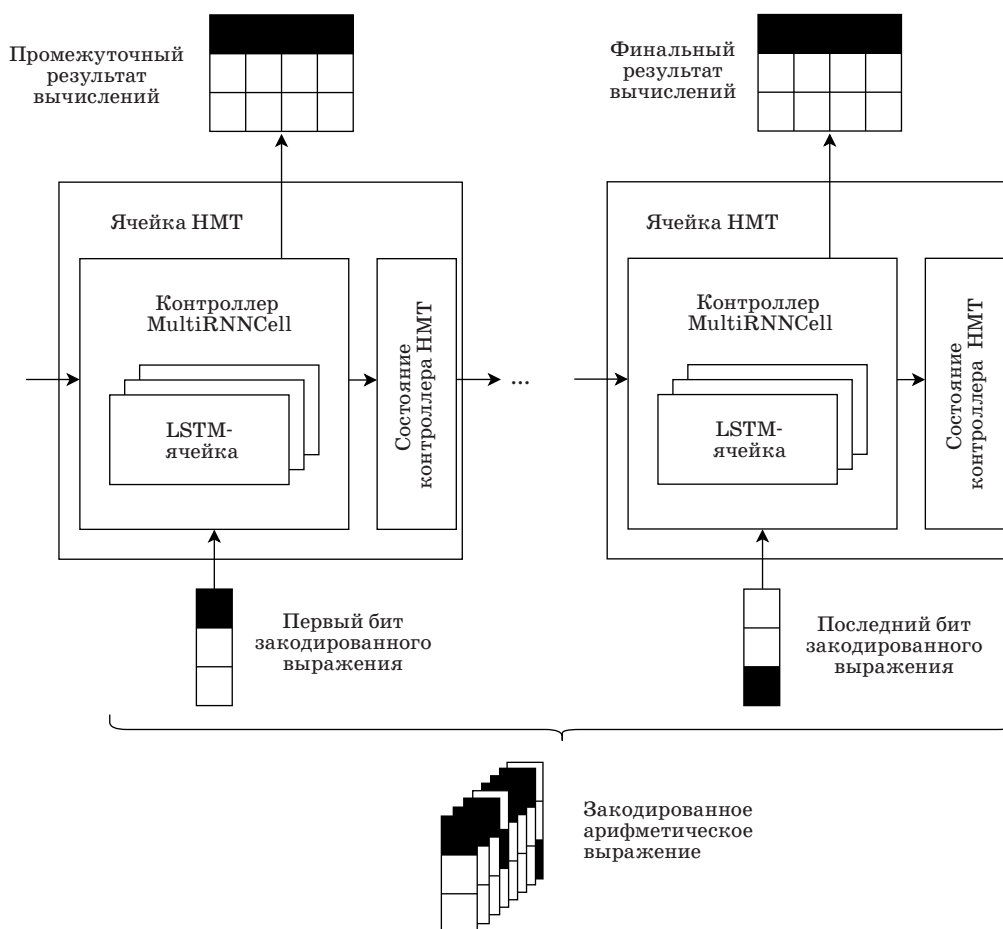
Аспекты реализации

Аспекты внедрения НМТ хорошо описаны в работах [35, 36] и в основном охватывают различные подходы к повышению устойчивости обучения за счет внедрения механизма контроля над градиентом. Важно отметить функцию ошибки, которая была использована как критерий для оценки качества обучения модели:

$$\delta_{i \geq 0,5} = \{(1, i \geq 0,5), (0, i < 0,5)\}; \quad (4)$$

$$e = \sum_{b=1}^B \sum_{n=1}^{N+1} \sum_{m=1}^M |\delta_{X(b,n,m) \geq 0,5} - Y_{b,n,m}| \quad (5)$$

где B — размер партии обучающих данных, поступающих в сеть; N — количество бит на одно



■ *Рис. 2.* Выполнение закодированного выражения для МТА с помощью НМТ
 ■ *Fig. 2.* Executing NTM with encoded expression for MTA operator

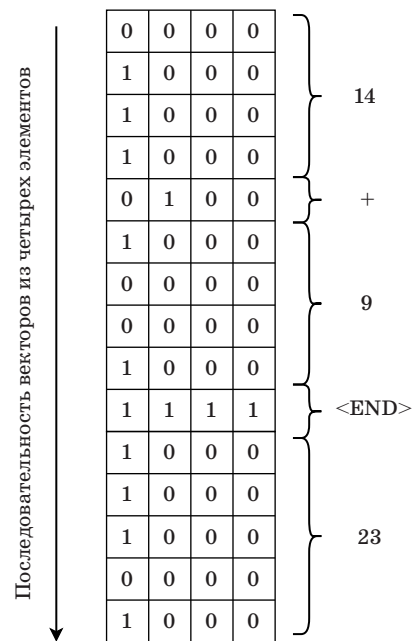
число; M — длина вектора, содержащего один бит выражения.

Архитектура НМТ предполагает создание одной или нескольких НМТ-ячеек, где каждая включает в себя контроллер, внешнюю память и головки для работы с этой памятью. Более того, НМТ способна к обобщению той задачи, для которой она была сконструирована и обучена. В частности, НМТ может обучаться на векторных последовательностях определенной длины, и после обучения сеть может решить задачу для большего набора последовательностей. Все приведенные ниже результаты могут быть воспроизведены с помощью проекта с открытым исходным кодом. Проект построен на основе текущей реализации системы TensorFlow Neural Turing Machine [22]. Однако были произведены значимые улучшения — осуществлен переход на поддерживаемую версию TensorFlow (2.3.0), так как базовая реализация НМТ основывалась на версии TensorFlow, которая более не поддерживается компанией Google.

Оценка НМТ как нейросетевого оператора МТА

В данной работе с помощью НМТ мы выражаем оператор МТА, который агрегирует оценки от фиксированного числа экспертов; каждая оценка преобразуется в числовую форму и представляется в виде двоичной строки [36]. Затем численные представления оценок объединяются в одно арифметическое выражение, которое уже кодируется в виде матрицы с определенным форматом. Каждое число представлено в виде матрицы с N строками и M столбцами, где N обозначает количество бит, используемых для кодирования числа в формате Little-Endian. Мы используем множество «каналов» в данных, чтобы помочь сети различать биты, кодирующие число, и биты, кодирующие маркеры. Первый столбец матрицы содержит биты данных, второй — биты для маркеров, а конечный маркер представлен вектором длины M . Для обучения создан набор данных, в котором каждое выражение закодировано в определенном формате и каждый обучающий набор содержит 32 арифметических выражения. Каждое закодированное выражение содержит либо две, либо три лингвистические оценки, представленные в численном виде и закодированные согласно схеме на рис. 3. В рамках данного исследования проводились эксперименты с различной длиной битовой строки: 4, 6, 8, 10, 16 бит.

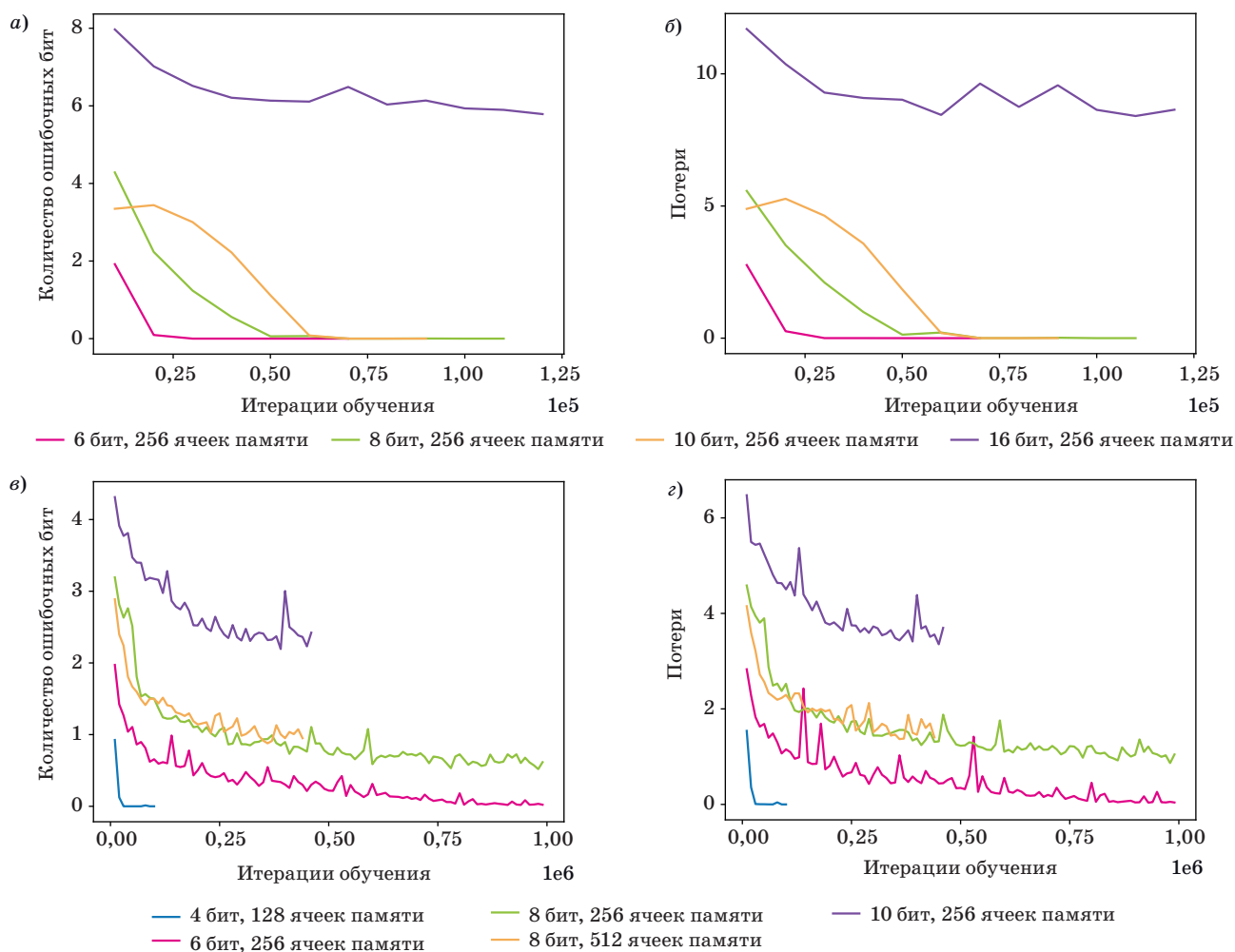
С точки зрения гиперпараметров и внутреннего устройства нейронной сети Тьюринга в рамках данного исследования использовались следующие значения. У НМТ одна считывающая и одна записывающая головка. Контроллер представлен



■ Рис. 3. Закодированное выражение для оператора МТА
 ■ Fig. 3. Encoded expression for the MTA operator

одной LSTM-ячейкой. Размер LSTM-ячейки равен 100 элементам и не менялся в зависимости от входных данных. Размер внешней памяти составлял 256 ячеек, каждая состояла из 128 элементов. Для каждого эксперимента, который характеризуется количеством экспертов и длиной битовой строки, производилось отдельное обучение нейронной сети и отдельная серия экспериментов после обучения. В случае агрегации двух оценок НМТ смогла добиться нулевой ошибки в экспериментах с 4-, 6-, 8- и 10-битными числами (рис. 4, а). При агрегации трех оценок НМТ смогла добиться нулевой ошибки только для 4-битных чисел (рис. 4, в). Для большей длины битовой строки ошибка варьируется от 0,26 для 6 бит до 2,42 для 10-битных чисел. В то же время НМТ смогла существенно минимизировать значение функции потерь во всех вариантах (рис. 4, б, г). Эксперименты проводились на различных конфигурациях CPU (таблица): Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20GHz, AMD EPYC 7282 16-Core Processor @ 2.79 GHz, Intel(R) Xeon(R) Gold 6140 CPU @ 2.30 GHz с нефиксированной частотой и 60 GB оперативной памяти и GPU: NVIDIA Corporation TU104GL [Tesla T4]. Результаты экспериментов можно воспроизвести с использованием авторского проекта с открытым исходным кодом (<https://github.com/demid5111/NeuralTuringMachine>).

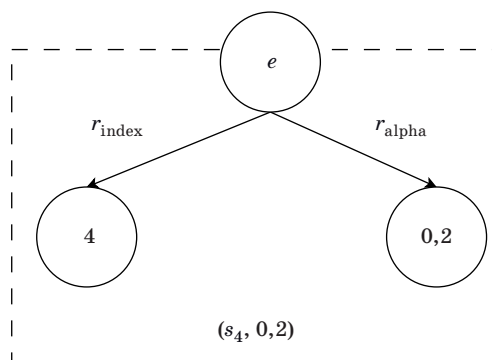
Ограничением подхода, апробированного в первой серии экспериментов, является необходимость промежуточного перевода лингвистических



■ **Рис. 4.** Динамика обучения НМТ для агрегации оценок, представленных в виде битовых строк фиксированной длины. Изменение функции ошибки (а) и потерь (б) НМТ при агрегации двух оценок; изменение функции ошибки (в) и потерь (з) НМТ при агрегации трех оценок

■ **Fig. 4.** NTM training dynamics for bit strings of various length. Change of error per sequence for two assessments aggregation (а), loss function for two assessments aggregation (б), error per sequence for three assessments aggregation (в), loss function for three assessments aggregation (з)

оценок в численное представление перед тем, как передавать их на вход обученному нейросетевому агрегатору оценок. Более того, как уже отмечалось, в рамках задачи построения нейросимволических систем вопрос представления знаний является актуальным. Требуется создание методов для перевода символов в истинно распределенное представление. Поэтому с точки зрения решаемой задачи агрегации лингвистических оценок необходимо построение распределенного представления для каждой такой оценки и обучение нейросетевого агрегатора, который бы мог работать с такими распределенными представлениями. Для этого предлагается рассматривать лингвистическую оценку с точки зрения тензорных представлений [28], а именно как дерево с двумя листьями. Например, для лингвистической оценки, представленной на рис. 5, при заданных $r_{\text{index}} = [1, 0]$ и $r_{\text{alpha}} = [0, 1]$,



■ **Рис. 5.** Представление лингвистической оценки $(s_4, 0,2)$ в виде структуры из двух заполнителей, т. е. листьев дерева, и двух позиционных ролей: r_{index} и r_{alpha}

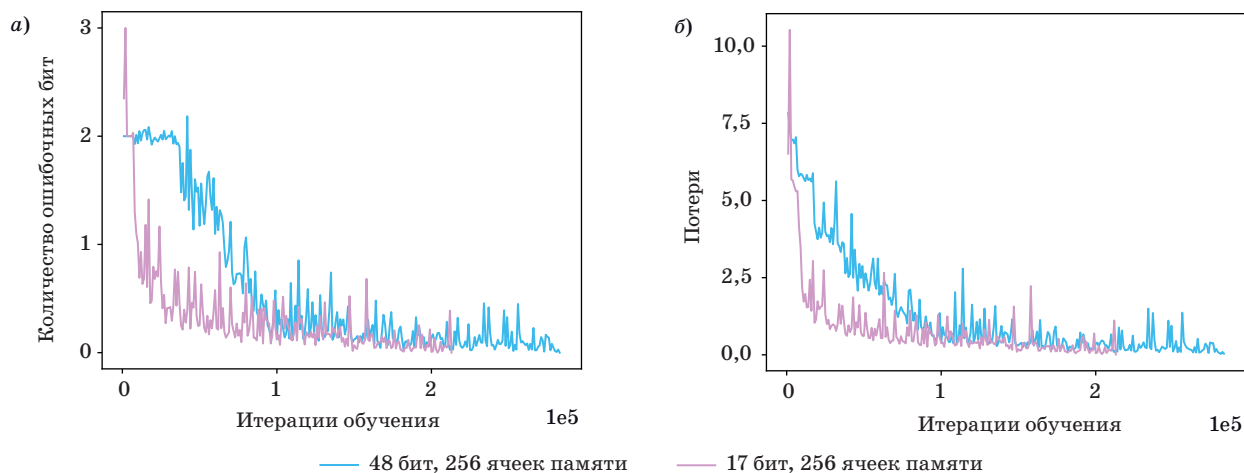
■ **Fig. 5.** Representing linguistic assessment $(s_4, 0,2)$ as a structure with two fillers as leaves and two positional roles: r_{index} and r_{alpha}

размере заданной лингвистической шкалы, равной пяти, и фиксированной точности для представления проекции лингвистической переменной, равной одному знаку после запятой, получаем следующее распределенное представление: $\psi = [[0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0], [0, 0], [0, 0], [0, 0], [0, 0], [1, 0], [0, 0], [0, 0], [0, 0], [0, 0], [0, 0], [0, 0], [0, 0], [0, 0], [0, 1], [0, 0], [0, 0], [0, 0]]$. При вытягивании такого распределенного представления в вектор получается тензор ранга 1, состоящий из 48 элементов.

Вторая серия экспериментов была посвящена подготовке данных и обучению нейросетевого агрегатора, который принимает на вход набор из распределенных представлений лингвистических оценок. Иными словами, по сравнению с первой серией экспериментов больше не требовался предварительный шаг по переводу лингвистических оценок в численное представление перед их упаковкой в качестве входных данных для нейросетевого агрегатора. Общая схема подготовки данных для первого и второго экспериментов выглядит аналогично. Во втором эксперименте вместо распределенного представления чисел в виде битовой строки заданной длины используется распределенное представление лингвистических оценок в виде битовой строки фиксированной длины, равной 48. Эксперименты проводились с агрегацией оценок только от двух экспертов. С точки зрения гиперпараметров и внутреннего устройства нейронной сети Тьюринга в рамках второго эксперимента использовались те же настройки, что и в первом эксперименте. Результат обучения нейросетевого агрегатора представлен на рис. 6, а и б. Вычислительные эксперименты проводились на Intel(R) Xeon(R) Gold 6240 CPU @ 2.30 GHz с нефиксированной частотой и 60 GB оперативной памяти.

Таким образом, нейросетевой агрегатор принимает на вход лингвистические оценки в распределенном виде, закодированные согласно правилам тензорных представлений, и выполняет задачу агрегации лингвистических оценок от двух экспертов с нулевой ошибкой.

В экспериментальном исследовании производилась агрегация оценок от фиксированного количества экспертов. Учитывая устройство МТА и фиксированное количество экспертов, данный оператор можно представить в виде единственного нейрона обычной сети прямого распространения (без функции активации). Использование существенно более сложной нейросетевой архитектуры для реализации этого оператора вызвано рядом причин. Во-первых, с точки зрения развития идей распределенных вычислений в СППР существует более широкая задача, которая заключается в разработке методов построения истинно распределенного представления информации о проблемной ситуации, которая, например, включает и информацию об экспертах, альтернативах и критериях. Веса в таком случае не должны быть закодированы в локальном виде, что происходит, например, при использовании единственного нейрона и некоторого количества обучаемых весов, соответствующих конкретному эксперту или критерию. Во-вторых, исследовательский вопрос заключался в попытке избавиться от необходимости создания новых символических операторов агрегации, чему посвящено большое количество работ, и их число с каждым годом только растет. Вместо этого предлагается использовать нейросетевую архитектуру, которая могла бы функционировать как оператор агрегации, при этом правила агрегации подбирались бы нейронной сетью самостоятельно. В-третьих, существу-



■ **Рис. 6.** Динамика функции ошибки (а) и потерь (б) НМТ для агрегации двух экспертных оценок, представленных в распределенном виде с помощью тензорных представлений

■ **Fig. 6.** Change of error per sequence for aggregation of two assessments encoded with tensor representations in a distributed form

- Таблица экспериментов по обучению НМТ
- NTM training experiments details

Номер серии экспериментов	Количество оценок	Размер битовой строки	Количество слотов внешней памяти	Количество ошибочных бит в результате работы НМТ	Аппаратное обеспечение
1	2	4	128	0	Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz
1	2	6	256	0,26	AMD EPYC 7282 16-Core Processor @ 2.79 GHz
1	2	8	256	0,62	Intel(R) Xeon(R) CPU E5-2650 v4 @ 2.20 GHz
1	2	8	512	0,26	Intel(R) Xeon(R) Gold 6140 CPU @ 2.30 GHz
1	2	10	256	0,62	Intel(R) Xeon(R) Gold 6140 CPU @ 2.30 GHz
1	3	4	128	0	AMD EPYC 7282 16-Core Processor @ 2.79 GHz
1	3	6	256	0	AMD EPYC 7282 16-Core Processor @ 2.79 GHz
1	3	8	256	0	AMD EPYC 7282 16-Core Processor @ 2.79 GHz
1	3	10	256	0	Intel(R) Xeon(R) Gold 6240 CPU @ 2.60 GHz
1	3	16	256	5,58	NVIDIA Corporation TU104GL [Tesla T4]
2	2	17	256	0	Intel(R) Xeon(R) Gold 6240 CPU @ 2.60 GHz
2	2	48	256	0	Intel(R) Xeon(R) Gold 6240 CPU @ 2.60 GHz

ет задача построения компилятора нейросетевых каскадов для различных конфигураций СППР и для работы с произвольным сценарием. В такой постановке задачи использование простейшего варианта в виде нейрона нецелесообразно, в то время как существенно более сложная нейросетевая архитектура позволяет решать поставленную задачу.

Заключение

Цель данного исследования заключалась в разработке нейросетевой архитектуры для эффективного представления субсимволических рассуждений в системах поддержки принятия решений в части алгоритмов агрегирования нечетких оценок экспертов для выбора альтернативных решений. Данная работа является частью серии статей о выстраивании нейросимволической архитектуры СППР [28]. В рамках общего исследования рассматривается шаг агрегации оценок при выполнении многокритериального выбора и осуществляется развитие методов к выражению алгоритмов агрегации в полностью коннективистском виде. Было продемонстрировано [28], что задача агрегации оценок может быть осуществлена с помощью нейросетевых архитектур, не требующих обучения. Однако показано, что такое кодирование чисел требует создания древовидных структур большой вложенности, а одним из известных недостатков тензорных представлений является взрывной рост размерности. Обработка таких распределенных представлений

большой размерности требует значительных вычислительных ресурсов и также является затратной по времени [37].

Настоящая работа показывает, что шаг агрегации оценок может быть реализован на основе обучения нейронной сети, а уже обученная нейронная сеть может достаточно быстро и точно агрегировать оценки. При этом мы считаем важным использовать тензорные представления для кодирования проблемной ситуации, а также самих оценок в силу отсутствия потерь информации. Иными словами, вклад данной работы в общее исследование построения нейросимволических СППР заключается в демонстрации возможности создания и практического использования обучаемого нейросетевого модуля на основе нейронных машин Тьюринга для агрегации оценок. Кроме того, предлагается использовать метод тензорных представлений для кодирования лингвистических оценок в распределенный вид. Это позволяет при необходимости переходить с символического уровня на субсимволический и обратно, что создает возможность для интерпретации работы нейросимволической СППР. В качестве дальнейших направлений исследования предлагается рассмотреть возможности НМТ для выражения другого алгоритма агрегации — МТWA оператора, так как он является самым распространенным оператором, используемым при разработке лингвистических СППР. Также интерес представляет исследование возможности создания распределенного представления проблемной ситуации и дальнейшего выражения всех этапов агрегации

экспертных оценок в виде динамики нейронной сети. Разработка таких решений, на наш взгляд, обеспечивает фундаментальные строительные блоки для распределенных гибридных нейросимволических СППР.

Литература

1. Matzkevich I., Abramson B. Decision analytic networks in artificial intelligence. *Management Science*, 1995, vol. 41, no. 1, pp. 1–22. doi:10.1287/MNSC.41.1.1
2. Silverman B. G. Unifying expert systems and the decision sciences. *Operations Research*, 1994, vol. 42, no. 3, pp. 393–413. doi:10.1287/opre.42.3.393
3. Kelley T. D. Symbolic and sub-symbolic representations in computational models of human cognition: What can be learned from biology? *Theory & Psychology*, 2003, vol. 13, no. 6, pp. 847–860. doi:10.1177/0959354303136005
4. Pylyshyn Z. W. *Computing in cognitive science. Foundations of cognitive science*. Cambridge, MA, MIT Press, 1989. Pp. 51–91.
5. Smolensky P., Legendre G. *The harmonic mind: From neural computation to optimality-theoretic grammar. Vol. 1: Cognitive architecture*. Cambridge, MA, MIT Press, 2006. 590 p.
6. Plate T. A. Holographic reduced representations. *IEEE Transactions on Neural Networks*, 1995, vol. 6, no. 3, pp. 623–641. doi:10.1109/72.377968
7. Kanerva P. Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*, 2009, no. 1, pp. 139–159. doi:10.1007/s12559-009-9009-8
8. Schlegel K., Neubert P., Protzel P. A comparison of vector symbolic architectures. *arXiv:2001.11797*, 2020.
9. Harnad S. The symbol grounding problem. *arXiv:cs/9906002v1*, 1999.
10. Anderson J. R., Lebiere C. J. *The atomic components of thought*. Psychology Press, 2014. 504 p.
11. Sun R., Peterson T. A hybrid agent architecture for reactive sequential decision making. In: *Connectionist symbolic integration: From unified to hybrid approaches*. New York, Psychology Press, 1997. Pp. 113–138.
12. Avery E., Kelley T., Davani D. Using cognitive architectures to improve robot control: Integrating production systems, semantic networks, and sub-symbolic processing. *Proc. 15th Annual Conference on Behavioral Representation in Modeling and Simulation (BRIMS)*, 2006, pp. 190–198.
13. Kelley T., Avery E., Long L., Dimperio E. A hybrid symbolic and sub-symbolic intelligent system for mobile robots. *Proc. AIAA Infotech@Aerospace Conference*, 2009, p. 1976. doi:10.2514/6.2009-1976
14. Parvar J., Lowe D., Emsley M., Duff R., Akintoye A. Neural networks as a decision support system for the decision to bid process. *Proc. 16th Annual ARCOM Conference*, 2000, pp. 209–217.
15. Delen D., Sharda R. Artificial neural networks in decision support systems. In: *Handbook on decision support systems 1*. Springer, 2008. Pp. 557–580. doi:10.1007/978-3-540-48713-5_26
16. Power D. J., Sharda R. Model-driven decision support systems: Concepts and research directions. *Decision Support Systems*, 2007, vol. 43, no. 3, pp. 1044–1061. doi:10.1016/j.dss.2005.05.030
17. Power D. J. *Decision support systems: Concepts and resources for managers*. Greenwood Publishing Group, 2002. 284 p.
18. Golmohammadi D. Neural network application for fuzzy multi-criteria decision making problems. *International Journal of Production Economics*, 2011, vol. 131, no. 2, pp. 490–504. doi:10.1016/j.ijpe.2011.01.015
19. Hwang C., Yoon K. *Multiple attribute decision making methods and applications: a state-of-the-art survey*. New York, Springer-Verlag Berlin Heidelberg, Springer Publications, 1981. 259 p. doi:10.1007/978-3-642-48318-9
20. Figueira J., Greco S., Ehrgott M. *Multiple criteria decision analysis: State of the art surveys*. New York, Springer-Verlag Berlin Heidelberg, Springer Publications, 2005. 1267 p. doi:10.1007/978-1-4939-3094-4
21. Demidovskij A. V., Babkin E. A. Developing a distributed linguistic decision making system. *Business Informatics*, 2019, vol. 13, no 1, pp. 18–32. doi:10.17323/1998-0663.2019.1.18.32
22. Herrera F., Martinez L. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 2000, vol. 8, no. 6, pp. 746–752. doi:10.1109/91.890332
23. Pogoyan K. S. The algorithm for generating optimal linguistic scale for group decision-making. *Proceedings of Voronezh State University: Systems Analysis and Information Technologies*, 2015, no. 2, pp. 97–104.
24. Wei G., Gao H. Pythagorean 2-tuple linguistic power aggregation operators in multiple attribute decision making. *Economic Research-Ekonomika Istraživanja*, 2020, vol. 33, no. 1, pp. 904–933. doi:10.1080/1331677X.2019.1670712
25. Rodriguez R. M., Martinez L., Herrera F. Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, 2011, vol. 20, no. 1, pp. 109–119. doi:10.1109/TFUZZ.2011.2170076

Финансовая поддержка

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 19-37-90058.

26. Liu P., Chen S.-M. Multiattribute group decision making based on intuitionistic 2-tuple linguistic information. *Information Sciences*, 2018, vol. 430, pp. 599–619. doi:10.1016/j.ins.2017.11.059
27. Wang J.-H., Hao J. A new version of 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 2006, vol. 14, no. 3, pp. 435–445. doi:10.1109/TFUZZ.2006.876337
28. Demidovskij A., Babkin E. Designing a neural network primitive for conditional structural transformations. *Lecture Notes in Computer Science*, 2020, vol. 12412, pp. 117–133. doi:10.1007/978-3-030-59535-7_9
29. Demidovskij A., Babkin E. Designing arithmetic neural primitive for sub-symbolic aggregation of linguistic assessments. *Journal of Physics: Conference Series*, 2020, vol. 1680, no. 1, article ID 012007. doi:10.1088/1742-6596/1680/1/012007
30. Graves A., Wayne G., Danihelka I. Neural turing machines. *arXiv:1410.5401*, 2014.
31. Graves A., Wayne G., Reynolds M., Harley T., Danihelka I., Grabska-Barwińska A., Colmenarejo S. G., Grefenstette E., Ramalho T., Agapiou J., Badia A. P., Hermann K. M., Zwols Y., Ostrovski G., Cain A., King H., Summerfield C., Blunsom P., Kavukcuoglu K., Hassabis D. Hybrid computing using a neural network with dynamic external memory. *Nature*, 2016, no. 538, pp. 471–476. doi:10.1038/nature20101
32. Zaremba W., Sutskever I. Reinforcement learning neural Turing machines-revised. *arXiv:1505.00521*, 2015.
33. Chen X., Xu H., Zhang Y., Tang J., Cao Y., Qin Z., Zha H. Sequential recommendation with user memory networks. *Proc. Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*, 2018, pp. 108–116. doi:10.1145/3159652.3159668
34. Grefenstette E., Hermann K. M., Suleyman M., Blunsom P. Learning to transduce with unbounded memory. *arXiv:1506.02516*, 2015.
35. Collier M., Beel J. Implementing neural Turing machines. *Lecture Notes in Computer Science*, 2018, vol. 11141, pp. 94–104. doi:10.1007/978-3-030-01424-7_10
36. Castellini J. Learning numeracy: Binary arithmetic with neural Turing machines. *arXiv:1904.02478*, 2019.
37. Demidovskij A. Encoding and decoding of recursive structures in neural-symbolic systems. *Optical Memory and Neural Networks*, 2021, vol. 30, no. 1, pp. 37–50.

UDC 004.896

doi:10.31799/1684-8853-2021-5-40-50

Adapting Neural Turing Machines for linguistic assessments aggregation in neural-symbolic decision support systems

A. V. Demidovskij^a, Post-Graduate Student, Senior Lecturer, orcid.org/0000-0003-3605-6332, ademidovskij@hse.ru

E. A. Babkin^a, PhD, Tech., Acting Head of Labs, Professor, orcid.org/0000-0003-2597-9043

^aNational Research University Higher School of Economics, Nizhegorodsky branch, 25/12, Bolshaya Pecherskaya St., 603155, Nizhny Novgorod, Russian Federation

Introduction: The construction of integrated neurosymbolic systems is an urgent and challenging task. Building neurosymbolic decision support systems requires new approaches to represent knowledge about a problem situation and to express symbolic reasoning at the subsymbolic level. **Purpose:** Development of neural network architectures and methods for effective distributed knowledge representation and subsymbolic reasoning in decision support systems in terms of algorithms for aggregation of fuzzy expert evaluations to select alternative solutions. **Methods:** Representation of fuzzy and uncertain estimators in a distributed form using tensor representations; construction of a trainable neural network architecture for subsymbolic aggregation of linguistic estimators. **Results:** The study proposes two new methods of representation of linguistic assessments in a distributed form. The first approach is based on the possibility of converting an arbitrary linguistic assessment into a numerical representation and consists in converting this numerical representation into a distributed one by converting the number itself into a bit string and further forming a matrix storing the distributed representation of the whole expression for aggregating the assessments. The second approach to translating linguistic assessments to a distributed representation is based on representing the linguistic assessment as a tree and coding this tree using the method of tensor representations, thus avoiding the step of translating the linguistic assessment into a numerical form and ensuring the transition between symbolic and subsymbolic representations of linguistic assessments without any loss of information. The structural elements of the linguistic assessment are treated as fillers with their respective positional roles. A new subsymbolic method of aggregation of linguistic assessments is proposed, which consists in creating a trainable neural network module in the form of a Neural Turing Machine. **Practical relevance:** The results of the study demonstrate how a symbolic algorithm for aggregation of linguistic evaluations can be implemented by connectionist (or subsymbolic) mechanisms, which is an essential requirement for building distributed neurosymbolic decision support systems.

Keywords — Neural Turing Machine, artificial neural networks, subsymbolic systems, hybrid intelligent systems, multi-attribute decision making.

For citation: Demidovskij A. V., Babkin E. A. Adapting Neural Turing Machines for linguistic assessments aggregation in neural-symbolic decision support systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 40–50 (In Russian). doi:10.31799/1684-8853-2021-5-40-50

Financial support

The reported study was funded by RFBR, project number 19-37-90058.

References

1. Matzkevich I., Abramson B. Decision analytic networks in artificial intelligence. *Management Science*, 1995, vol. 41, no. 1, pp. 1–22. doi:10.1287/MNSC.41.1.1
2. Silverman B. G. Unifying expert systems and the decision sciences. *Operations Research*, 1994, vol. 42, no. 3, pp. 393–413. doi:10.1287/opre.42.3.393
3. Kelley T. D. Symbolic and sub-symbolic representations in computational models of human cognition: What can be learned from biology? *Theory & Psychology*, 2003, vol. 13, no. 6, pp. 847–860. doi:10.1177/0959354303136005
4. Pylyshyn Z. W. *Computing in cognitive science*. Foundations of cognitive science. Cambridge, MA, MIT Press, 1989. Pp. 51–91.
5. Smolensky P., Legendre G. *The harmonic mind: From neural computation to optimality-theoretic grammar. Vol. 1: Cognitive architecture*. Cambridge, MA, MIT Press, 2006. 590 p.
6. Plate T. A. Holographic reduced representations. *IEEE Transactions on Neural Networks*, 1995, vol. 6, no. 3, pp. 623–641. doi:10.1109/72.377968
7. Kanerva P. Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive Computation*, 2009, no. 1, pp. 139–159. doi:10.1007/s12559-009-9009-8
8. Schlegel K., Neubert P., Protzel P. A comparison of vector symbolic architectures. *arXiv:2001.11797*, 2020.
9. Harnad S. The symbol grounding problem. *arXiv:cs/9906002v1*, 1999.
10. Anderson J. R., Lebiere C. J. *The atomic components of thought*. Psychology Press, 2014. 504 p.
11. Sun R., Peterson T. *A hybrid agent architecture for reactive sequential decision making*. In: *Connectionist symbolic integration: From unified to hybrid approaches*. New York, Psychology Press, 1997. Pp. 113–138.
12. Avery E., Kelley T., Davani D. Using cognitive architectures to improve robot control: Integrating production systems, semantic networks, and sub-symbolic processing. *Proc. 15th Annual Conference on Behavioral Representation in Modeling and Simulation (BRIMS)*, 2006, pp. 190–198.
13. Kelley T., Avery E., Long L., Dimperio E. A hybrid symbolic and sub-symbolic intelligent system for mobile robots. *Proc. AIAA Infotech@Aerospace Conference*, 2009, p. 1976. doi:10.2514/6.2009-1976
14. Parvar J., Lowe D., Emsley M., Duff R., Akintoye A. Neural networks as a decision support system for the decision to bid process. *Proc. 16th Annual ARCOM Conference*, 2000, pp. 209–217.
15. Delen D., Sharda R. *Artificial neural networks in decision support systems*. In: *Handbook on decision support systems 1*. Springer, 2008. Pp. 557–580. doi:10.1007/978-3-540-48713-5_26
16. Power D. J., Sharda R. Model-driven decision support systems: Concepts and research directions. *Decision Support Systems*, 2007, vol. 43, no. 3, pp. 1044–1061. doi:10.1016/j.dss.2005.05.030
17. Power D. J. *Decision support systems: Concepts and resources for managers*. Greenwood Publishing Group, 2002. 284 p.
18. Golmohammadi D. Neural network application for fuzzy multi-criteria decision making problems. *International Journal of Production Economics*, 2011, vol. 131, no. 2, pp. 490–504. doi:10.1016/j.ijpe.2011.01.015
19. Hwang C., Yoon K. *Multiple attribute decision making methods and applications: a state-of-the-art survey*. New York, Springer-Verlag Berlin Heidelberg, Springer Publications, 1981. 259 p. doi:10.1007/978-3-642-48318-9
20. Figueira J., Greco S., Ehrgott M. *Multiple criteria decision analysis: State of the art surveys*. New York, Springer-Verlag Berlin Heidelberg, Springer Publications, 2005. 1267 p. doi:10.1007/978-1-4939-3094-4
21. Demidovskij A. V., Babkin E. A. Developing a distributed linguistic decision making system. *Business Informatics*, 2019, vol. 13, no. 1, pp. 18–32. doi:10.17323/1998-0663.2019.1.18.32
22. Herrera F., Martinez L. A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 2000, vol. 8, no. 6, pp. 746–752. doi:10.1109/91.890332
23. Pogosyan K. S. The algorithm for generating optimal linguistic scale for group decision-making. *Proceedings of Voronezh State University: Systems Analysis and Information Technologies*, 2015, no. 2, pp. 97–104.
24. Wei G., Gao H. Pythagorean 2-tuple linguistic power aggregation operators in multiple attribute decision making. *Economic Research-Ekonomika Istraživanja*, 2020, vol. 33, no. 1, pp. 904–933. doi:10.1080/1331677X.2019.1670712
25. Rodriguez R. M., Martinez L., Herrera F. Hesitant fuzzy linguistic term sets for decision making. *IEEE Transactions on Fuzzy Systems*, 2011, vol. 20, no. 1, pp. 109–119. doi:10.1109/TFUZZ.2011.2170076
26. Liu P., Chen S.-M. Multiattribute group decision making based on intuitionistic 2-tuple linguistic information. *Information Sciences*, 2018, vol. 430, pp. 599–619. doi:10.1016/j.ins.2017.11.059
27. Wang J.-H., Hao J. A new version of 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, 2006, vol. 14, no. 3, pp. 435–445. doi:10.1109/TFUZZ.2006.876337
28. Demidovskij A., Babkin E. Designing a neural network primitive for conditional structural transformations. *Lecture Notes in Computer Science*, 2020, vol. 12412, pp. 117–133. doi:10.1007/978-3-030-59535-7_9
29. Demidovskij A., Babkin E. Designing arithmetic neural primitive for sub-symbolic aggregation of linguistic assessments. *Journal of Physics: Conference Series*, 2020, vol. 1680, no. 1, article ID 012007. doi:10.1088/1742-6596/1680/1/012007
30. Graves A., Wayne G., Danihelka I. Neural Turing machines. *arXiv:1410.5401*, 2014.
31. Graves A., Wayne G., Reynolds M., Harley T., Danihelka I., Grabska-Barwińska A., Colmenarejo S. G., Grefenstette E., Ramalho T., Agapiou J., Badia A. P., Hermann K. M., Zwols Y., Ostrovski G., Cain A., King H., Summerfield C., Blunsom P., Kavukcuoglu K., Hassabis D. Hybrid computing using a neural network with dynamic external memory. *Nature*, 2016, no. 538, pp. 471–476. doi:10.1038/nature20101
32. Zaremba W., Sutskever I. Reinforcement learning neural Turing machines-revised. *arXiv:1505.00521*, 2015.
33. Chen X., Xu H., Zhang Y., Tang J., Cao Y., Qin Z., Zha H. Sequential recommendation with user memory networks. *Proc. Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18)*, 2018, pp. 108–116. doi:10.1145/3159652.3159668
34. Grefenstette E., Hermann K. M., Suleyman M., Blunsom P. Learning to transduce with unbounded memory. *arXiv:1506.02516*, 2015.
35. Collier M., Beel J. Implementing neural Turing machines. *Lecture Notes in Computer Science*, 2018, vol. 11141, pp. 94–104. doi:10.1007/978-3-030-01424-7_10
36. Castellini J. Learning numeracy: Binary arithmetic with neural Turing machines. *arXiv:1904.02478*, 2019.
37. Demidovskij A. Encoding and decoding of recursive structures in neural-symbolic systems. *Optical Memory and Neural Networks*, 2021, vol. 30, no. 1, pp. 37–50.

UDC 004.728.3.057.4

doi:10.31799/1684-8853-2021-5-51-58

Signal power and energy-per-bit optimization problems in mMTC systems

A. A. Burkov^a, Assistant Professor, orcid.org/0000-0002-0920-585X, a.burkov@k36.org

^aSaint-Petersburg State University of Aerospace Instrumentation, 67, B. Morskaya St., 190000, Saint-Petersburg, Russian Federation

Introduction: Currently, the issues of Internet of Things technology are being actively studied. The operation of a large number of various self-powered sensors is within the framework of a massive machine-type communication scenario, using random access methods. Topical issues in this type of communication are how to reduce the transmission signal power and to increase the device lifetime by reducing the consumed energy per bit. **Purpose:** Formulation and analysis of the problems of minimizing the transmission power and consumed energy per bit in systems with or without retransmissions in order to obtain the achievability bounds. **Results:** A model of the system is described, within which four problems are formulated and described, concerning the signal power minimization and energy consumption for given parameters (the number of information bits, the spectral efficiency of the system, and the Packet Delivery Ratio). The numerical results of solving these optimization problems are presented. They make it possible to obtain the achievability bounds for the considered characteristics in systems with or without losses. The lower bounds obtained by the Shannon formula are also presented, assuming that the message length is not limited. The obtained results showed that solving the minimization problem with respect to one of the parameters (signal power or consumed energy per bit) does not minimize the second parameter. This difference is most significant for information messages of a small length, which is common in IoT scenarios. **Practical relevance:** The results obtained allow you to assess the potential for minimizing the transmission signal power and consumed energy per bit in random multiple access systems with massive machine-type communication scenarios. **Discussion:** The presented problems were solved without taking into account the average delay of message transmission; the introduction of such a limitation should increase the transmitted signal power and consumed energy per bit.

Keywords – random multiple access, spectral efficiency, signal-to-noise ratio, energy per bit, mMTC, IoT, minimization problem.

For citation: Burkov A. A. Signal power and energy-per-bit optimization problems in mMTC systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 51–58. doi:10.31799/1684-8853-2021-5-51-58

Introduction

In the framework of the currently unfolding 5G communication standard, as well as the development of the next generation 6G, scenarios of the Internet of Things (IoT) are considered [1, 2]. Soon, the number of IoT devices will be on the order of one million per square kilometer, and each of the devices will periodically transmit a small amount of data [3–5]. Due to the large volume of devices, it is not possible to use scheduling methods to access the channel share. Therefore, within the framework of IoT systems, the use of random access methods is assumed [6–9]. With regard to the requirements put forward with the systems (transmission rate, delay, etc.), the following types of IoT are distinguished: massive IoT, critical IoT, broadband IoT and industrial IoT [10]. Massive IoT operates in a massive machine-type communications (mMTC) scenario and describes data acquisition systems with a large number of low-power end devices (such as sensors) that periodically transmit a small amount of data. Examples of the mass Internet of Things are temperature, pressure, light sensors, and meters in smart home technology. The amount of data transferred is small, but the number of IoT

devices is very large [11]. The main requirements for this scenario are a large number of devices, stability, low power consumption, delivery of messages with given reliability, as well as a limitation on average delay [12].

The work considers the following scenario. There are a large number of stand-alone user devices and one base station. The level of attenuation in the channel between the base station and all user devices is the same. User devices can transmit data to the base station and receive service messages from the base station. All user devices, at random moments in time, have a small piece of data of the same information length, which the device must transmit to the base station. The total input arrival rate of messages in the system is set. There are two options for this scenario.

In the first variant, the system does not provide for the presence of a feedback communication channel. In this case, the user device transmits the message once and deletes it, regardless of the success of the transmission. An important characteristic is the Packet Delivery Ratio (PDR), which can be set in accordance with the requirements of the scenario.

In the second variant, the system has a feedback channel, and the user device, according to some al-

gorithm, repeats the transmission until it receives confirmation from the base station that data has been received from this user device.

In what follows, the first variant of the scenario will be called transmission without retransmissions, and the second variant — transmission with retransmissions.

In view of some of the design features of transmission networks, it may be necessary to limit the signal power of data transmission from the user device. This, for example, can be caused by the conditions of the controlling organizations. In this scenario, reducing the transmit signal power is equivalent to decreasing the signal-to-noise ratio to achieve the desired system performance. As noted earlier, in IoT systems, most devices can be powered from an autonomous power source. Therefore, the question arises about increasing the operating time of the device without additional maintenance (replacing the battery or charging). In the described scenario, the increase in the operating time of the user device is reduced to the task of reducing the consumption energy per bit during data transmission.

Accordingly, for both scenarios, two minimization problems (task) can be considered.

First task. Minimizing signal-to-noise ratio (SNR). In fact, this is the minimization of the transmission signal power for the given parameters: the length of the transmitted message in bits, the intensity of the input arrival rate, packet delivery ratio, and the noise power.

Second task. Minimizing consumption energy per bit $\frac{E_b}{N_0}$. In fact, this is the minimization of energy consumption for the transmission of a message with fixed parameters: the length of the transmitted message in bits, the intensity of the input arrival rate, packet delivery ratio, and the noise power spectral density.

It should be noted that when solving the second task within the framework of the feedback system, it is necessary to take into account the number of retransmissions of the message sent by the user device.

As a rule, in works devoted to random multiple access in systems with mMTC, varieties of algorithms such as ALOHA and its modifications are considered [13–15]. Within the framework of this work, minimization problems are formulated for SNR and $\frac{E_b}{N_0}$ for systems with and without retransmissions, where the optimal ALOHA algorithm will be analyzed as an access algorithm.

System model

A random multiple access system is considered. Additive White Gaussian Noise (AWGN) is pres-

ent in the channel. For this system, we introduce a number of assumptions.

Assumption 1. The system has one base station and many user devices. The system has an input message arrival rate per unit of time (slot), which is a Poisson distribution with the parameter Λ [message / slot]. Each message contains k bits of information. Using the modulation and coding scheme A , on the basis of k bits, a signal is generated that contains n samples.

Assumption 2. There is a potentially infinite number of user devices (user devices and messages are equal). A user device with a message ready to send is called active.

Assumption 3. The user devices and the base station have synchronization, both by samples and by slots. The slot is the time it takes to transmit one message and lasts n samples.

Assumption 4. We consider a time-discrete communication channel with additive white Gaussian noise (Gaussian Multiple Access Channel), defined as $Y = \sum_{i=1}^{K_t} X_i + Z$, where Y is the channel output signal; K_t is the number of user devices transmitting a signal per channel in the slot with number t (which is a random variable); X_i is the signal of the i -th user device; Z is additive white Gaussian noise, and $Z \sim \mathcal{N}(0,1)$. The signals transmitted by the user devices contain n samples and the maximum energy limitation of each signal $X_i = nSNR$ [16, 17].

The described system model is characterized by a set of parameters: Λ [messages/slot], k [bits], n [samples], SNR [times] or $\frac{E_b}{N_0}$ [energy/bit].

Additionally, we introduce into consideration the value defined as:

$$\rho \triangleq \frac{\Lambda k}{n}.$$

This value characterizes the average number of bits transmitted per sample. In what follows, it will be called spectral efficiency.

In the next section, taking into account the described model, the analysis of the system, both with guaranteed message delivery and with losses, will be considered. The tasks of minimizing signal power (SNR) and consumption energy per bit $\left(\frac{E_b}{N_0}\right)$ will be formulated and described.

System analysis

First, consider a system operating in without acknowledged mode, with each user device who has a message transmits it at the beginning of the next slot. The number of information bits k to be trans-

mitted to user devices, the PDR P_d , and spectral efficiency ρ are indicated.

Add the following assumption to the general model:

Assumption 5. The user device transmits a message at the beginning of the next slot, immediately after the message appears, and leaves the system without waiting for confirmation.

The main parameter for this system is the PDR value. There are two ways to determine the PDR (by analogy with the definition of the delay in [18]):

Actual PDR value. All messages in the system are numbered, and the algorithm works for some time. Then the total number of transmitted messages is divided by the total number of messages that were transmitted during the given period of time:

$$P_{deliv,a} = \lim_{t \rightarrow \infty} \frac{N_{deliv}(t)}{N(t)},$$

where t is the operating time of the system; $N_{deliv}(t)$ is the number of messages delivered during time t ; $N(t)$ is the number of messages sent during time t .

Virtual PDR value. A target user device is added to the system at a random moment in time and the probability with which his message will be successfully transmitted is determined:

$$P_{deliv,v} = \lim_{t \rightarrow \infty} \Pr \left\{ \begin{array}{l} \text{the packet of the target user added} \\ \text{at time } t \text{ was successfully deliver} \end{array} \right\}.$$

It was shown in [18] that adding a target user device to the system with a Poisson input stream does not violate the input stream itself and its characteristics. Also in the case of a Poisson input stream, both PDR definitions give the same meaning ($P_{deliv,a} = P_{deliv,v}$).

In this work, to analyze and solve the described minimization problem, the second definition will be used to find the PDR value, and we will denote it as P_d .

With this in mind, let's analyze the PDR in general. Let the number of samples n , the number of bits transmitted by the user device k , the modulation and coding scheme A , and the SNR value is given. In a random slot, we add a target user device with a message ready for transmission. Then the probability that the message of the target user device will be delivered successfully, in accordance with the second definition of PDR, is calculated as follows:

$$P_d = \Pr \left\{ \begin{array}{l} 0 \text{ user devices} \\ \text{appeared in the system} \end{array} \right\} \times \Pr \left\{ \begin{array}{l} \text{message decoded} \\ \text{successfully} \end{array} \right\}.$$

Using the formula for the probability of occurrence of i messages with a Poisson input stream with a parameter Λ , we can write:

$$\Pr \left\{ \begin{array}{l} i \text{ user devices} \\ \text{appeared in the system} \end{array} \right\} = \frac{(\Lambda)^i}{i!} e^{-\Lambda}.$$

Let there be some function $P_e(A, n, k, SNR)$ that allows you to determine the probability of decoding error for given parameters A , SNR, n and k , then:

$$\Pr \left\{ \begin{array}{l} \text{message decoded} \\ \text{successfully} \end{array} \right\} = 1 - P_e(A, n, k, SNR).$$

Then, taking into account the fact that $\Lambda = \frac{\lambda n}{k}$, we get the expression:

$$P_d = e^{-\rho \frac{n}{k}} (1 - P_e(A, n, k, SNR)). \quad (1)$$

Polyansky's formula [16]

$$p_e(n, k, SNR) = Q \left(\frac{\frac{n}{2} \log_2(1 + SNR) + \frac{1}{2} \log_2 n - k}{\sqrt{\frac{n}{2} \frac{SNR}{SNR + 1} \log_2 e}} \right)$$

allows determining the probability of decoding error for given parameters SNR, n and k . Then we define the probability of successful decoding as $1 - p_e(A, n, k, SNR)$. Substitute in formula (1) and get:

$$P_d = e^{-\rho \frac{n}{k}} (1 - p_e(k, n, SNR)). \quad (2)$$

Then the *first minimization task* can be formulated:

$$\begin{array}{l} \text{given: } k, \rho, P_d \\ \text{minimize: } SNR \text{ in } n \\ \text{subject to:} \end{array} \quad (3)$$

$$P_d = e^{-\rho \frac{n}{k}} (1 - p_e(k, n, SNR))$$

The solution to this task will be the minimum SNR value at which the required PDR is achieved for a given number of transmitted bits k and spectral efficiency ρ .

Consider the formulation of a similar task but from the point of view of minimizing the expended energy per bit. It is known that SNR and $\frac{E_b}{N_0}$ are related through the following expression:

$$\frac{E_b}{N_0} = \frac{n SNR}{2k}. \quad (4)$$

With this in mind, expression (2) can be written as

$$P_d = e^{-\rho \frac{n}{k}} \left(1 - p_e \left(k, n, \frac{2k E_b}{n N_0} \right) \right). \quad (5)$$

Then the *second minimization task* (consumption energy per bit, for a given number of transmitted bits and spectral efficiency) can be formulated as follows:

$$\begin{aligned} &\text{given: } k, \rho, P_d \\ &\text{minimize: } \frac{E_b}{N_0} \text{ in } n \\ &\text{subject to:} \end{aligned} \quad (6)$$

$$P_d = e^{-\rho \frac{n}{k}} \left(1 - p_e \left(k, n, \frac{2k E_b}{n N_0} \right) \right)$$

Consider the second system, in which the user device transmits his message according to some algorithm until it is successfully delivered. The number of information bits k and the spectral efficiency ρ are given.

It is necessary to change the model of the system by adding an assumption to it:

*Assumption 5**. By the beginning of each slot, all user devices and base station know how many user devices have a message ready for transmission M_a . Each user device transmits his message with probability $p = \frac{G}{M_a}$, where G is the algorithm parameter ($0 < G \leq 1$), until it is successfully delivered.

The critical input arrival rate can be limited by the following inequality:

$$\Lambda \leq \Pr \left\{ \begin{array}{l} \text{probability of transmission} \\ \text{by one user device} \end{array} \middle| G \right\} \times \Pr \left\{ \begin{array}{l} \text{message decoded} \\ \text{successfully} \end{array} \right\}. \quad (7)$$

Then, taking into account the spectral efficiency [bit/sample] and by analogy with the system without retransmissions, the expression can be written as

$$\rho = \frac{k}{n} G e^{-G} \left(1 - p_e(k, n, SNR) \right). \quad (8)$$

It is worth noting that for the algorithm with retransmissions, taking into account SNR, one can find the value of G that maximizes the spectral efficiency. To do this, you need to solve the following optimization problem:

$$G_{\text{opt}} = \max_G G e^{-G}. \quad (9)$$

It is well known that the solution to this optimization problem is $G = 1$ [19].

Then the third minimization task can be formulated:

$$\begin{aligned} &\text{given: } k, \rho \\ &\text{minimize: } SNR \text{ in } n \\ &\text{subject to:} \end{aligned} \quad (10)$$

$$\rho = \frac{k}{n} e^{-1} \left(1 - p_e(k, n, SNR) \right)$$

Consider the same system, but from the point of view of minimizing consumption energy per bit. In contrast to solving the problem in terms of signal power, it is necessary to take into account the average number of messages S sent by a user device. From work [20] it follows that the average number of transmissions is determined by the expression:

$$S = \frac{kG(1 - \pi_0)}{\rho n}, \quad (11)$$

where S is the average number of messages sent by the user device; k is the number of information symbols; π_0 is the stationary probability that there are no active user devices in the system; ρ is the spectral efficiency; n is the number of samples.

Considering that the value ρ specified in the solution of the task is critical, then the value $\pi_0 \rightarrow 0$. Then, taking into account the number of transmissions, the SNR values and $\frac{E_b}{N_0}$ are related by the following expression:

$$\frac{E_b}{N_0} = \frac{SnSNR}{2k} = 2 \frac{E_b}{N_0} \frac{\rho}{G}. \quad (12)$$

In accordance with expression (12) and task (10), we can formulate the fourth *minimization task* (for $\frac{E_b}{N_0}$):

$$\begin{aligned} &\text{given: } k, \rho \\ &\text{minimize: } \frac{E_b}{N_0} \text{ in } n, G \\ &\text{subject to:} \end{aligned} \quad (13)$$

$$\rho = \frac{k}{n} G e^{-G} \left(1 - p_e \left(k, n, 2 \frac{E_b}{N_0} \frac{\rho}{G} \right) \right)$$

All considered minimization tasks are presented in Table 1.

The solutions to these tasks allow us to determine the achievability bounds of signal power and consumption energy per bit in random multiple access systems.

If we consider the solution of the presented tasks for $k \rightarrow \infty, n \rightarrow \infty$ and the speed $R = k/n$, taking into ac-

■ **Table 1.** Formulation of minimization tasks

System	Minimizing signal power (SNR)	Minimization of energy per bit costs $\left(\frac{E_b}{N_0}\right)$
Without retransmissions	<p>Task 1 given: k, ρ, P_d minimize: SNR in n subject to: $P_d = e^{-\rho \frac{n}{k}} (1 - p_e(k, n, SNR))$</p>	<p>Task 2 given: k, ρ, P_d minimize: $\frac{E_b}{N_0}$ in n subject to: $P_d = e^{-\rho \frac{n}{k}} \left(1 - p_e\left(k, n, \frac{2k E_b}{n N_0}\right)\right)$</p>
With retransmissions	<p>Task 3 given: k, ρ minimize: SNR in n subject to: $\rho = \frac{k}{n} e^{-1} (1 - p_e(k, n, SNR))$</p>	<p>Task 4 given: k, ρ minimize: $\frac{E_b}{N_0}$ in n, G subject to: $\rho = \frac{k}{n} G e^{-G} \left(1 - p_e\left(k, n, 2 \frac{E_b \rho}{N_0 G}\right)\right)$</p>

■ **Table 2.** Lower bounds for solving minimization tasks

System	Minimizing signal power (SNR)	Minimization of energy per bit costs $\left(\frac{E_b}{N_0}\right)$
Without retransmissions	<p>Task 1 $SNR > 2^{\frac{2\rho}{\ln(P_d)} - 1}$</p>	<p>Task 2 $\frac{E_b}{N_0} > \frac{2^{\frac{2\rho}{\ln(P_d)} - 1}}{2\rho / \ln(P_d)}$</p>
With retransmissions	<p>Task 3 $SNR > 2^{\frac{2\rho}{e^{-1}} - 1}$</p>	<p>Task 4 $\frac{E_b}{N_0} > \frac{\left(2^{\frac{2\rho}{2Ge^{-G}} - 1}\right) G}{2\rho}$</p>

count the Shannon theorem for discrete channels, then we can obtain the lower bounds presented in Table 2.

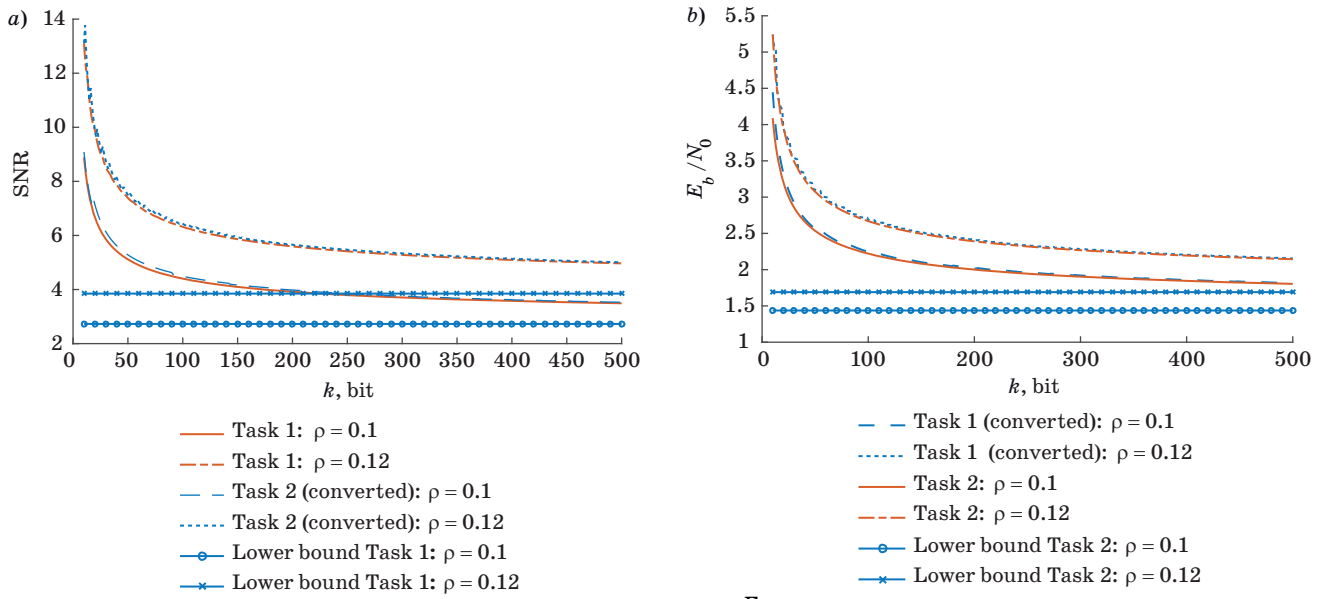
In the next section, the numerical results of solving these tasks will be presented and qualitative analysis of the results will be carried out.

Numerical results

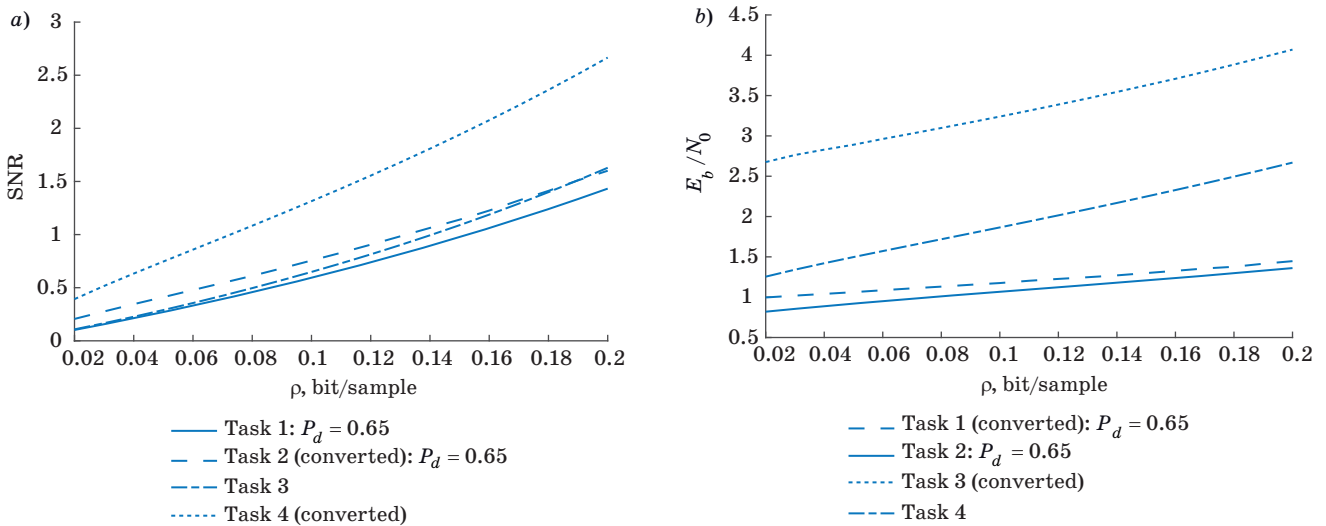
Figure 1, *a, b* show the results of solving minimization tasks for a system without retransmissions (tasks 1 and 2), for a given PDR $P_d = 0.9$ at spectral efficiency $\rho = 0.1$ and $\rho = 0.12$. The values marked

“converted” mean that the minimization task was solved by one of the parameters and was recalculated for the second parameter in accordance with formula (4). It should be noted that in the case of an increase in the required PDR, the obtained values of the achievability bounds increase. It also follows from the graphs that the solution of the problem of minimization in terms of signal power does not give a solution minimization in terms of consumption energy per bit and vice versa. As k grows, these differences decrease. The lower bounds are obtained in accordance with the expressions from Table 2. As can be seen from the figure, with an increase in the number of information bits, the achievability bound tends to the lower bound. As $k \rightarrow \infty$, the results will be the same. However, the IoT has short message lengths. The tasks posed earlier can be changed by minimizing one of the parameters with a constraint on the other, but such solutions will lie above the obtained bounds.

Figure 2, *a, b* show the results of solving the minimization problems for all tasks, for a given number of information bits k equal to 50. The PDR for the system without repetitions is 0.65 and was chosen for the convenience of the scale on the graph. The values marked “converted” also mean that the minimization task was solved by one of the parameters and was recalculated for the second parameter in accordance with formulas (4) — for tasks 1 and 2, (12) — for tasks 3 and 4. As noted earlier, solving the task for the system with retransmissions by SNR, the optimal system parameter G is equal to 1. However, when solving this task by $\frac{E_b}{N_0}$,



■ Fig. 1. Achievability bounds and lower bounds for SNR (a) and $\frac{E_b}{N_0}$ (b) from k for tasks 1 and 2 at $P_d = 0.9$



■ Fig. 2. Achievability bounds SNR (a) and $\frac{E_b}{N_0}$ (b) from all tasks for ρ at $k = 50$

the values of this parameter turn out to be below 1 and tend to this value with increasing spectral efficiency. By optimizing this parameter, there is a gain in consumption energy per bit in systems with retransmissions. If this parameter is equal to 1 in both tasks, they give similar results.

Conclusion

The paper describes a model of a system with a potentially unlimited number of user devices in a Gaussian Multiple Access Channel. Within the

framework of this model, systems without retransmissions and with retransmissions are considered. Four tasks of minimizing signal power and consumption energy per bit for given system parameters were formulated and described: the number of information bits, the spectral efficiency of the system, and the packet delivery ratio. For the problems formulated, the lower bounds obtained by the Shannon formula are presented under the assumption that the message length is not limited. With an increase in the number of information bits, the results of solving tasks and the corresponding achievability bounds will tend to lower bounds. Numerical

solutions for the assigned tasks are presented. The results obtained showed that solving the minimization problem with respect to one of the parameters (signal power or consumption energy per bit) does not minimize the second parameter. This difference is most significant for small information message lengths, which corresponds to IoT scenarios. Difference a decibel at typical value of information bits $k = 50$ bits and spectral efficiency $\rho = 0.1$ according to SNR: for tasks 1 and 2 with a PDR = 0.9 is 0.13 dB; for tasks 3 and 4 is 3.07 dB. Results for

$\frac{E_b}{N_0}$ in similar conditions: for tasks 1 and 2 with a

PDR = 0.9 is 0.12 dB; for tasks 3 and 4 is 2.4 dB. However, in IoT scenarios, the number of information bits is assumed to be small. For a system with retransmissions, with minimization of energy per bit, the optimal transmission parameter G is less

than 1 and tends to it with an increase in the required spectral efficiency. When minimizing signal power in such a system, this parameter is always 1.

The results obtained allow us to assess the potential for minimizing transmission signal power and consumption energy per bit in random multiple access systems in the framework of IoT scenarios. Sometimes, in practice, the transmitter signal power can be limited and it is required to minimize the energy consumption, therefore, one of the parameters can be minimized with a restriction on the other, then the boundaries obtained in the work will be a lower estimate for solving such problems.

Financial support

The research was supported by RFBR, project number 19-37-90041.

References

1. Letaief K. B., Chen W., Shi Y., Zhang J., and Zhang Y. J. A. The roadmap to 6G: AI empowered wireless networks. *IEEE Communications Magazine*, 2019, vol. 57(8), pp. 84–90. doi:10.1109/MCOM.2019.1900271
2. Piran M. J., and Suh D. Y. Learning-driven wireless communications, towards 6G. *International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, IEEE, 2019, pp. 219–224. doi:10.1109/iCCECE46942.2019.8941882
3. Lema M. A., Laya A., Mahmoodi T., Cuevas M., Sachs J., Markendahl J., and Dohler M. Business case and technology analysis for 5G low latency applications. *IEEE Access*, 2017, vol. 5, pp. 5917–5935. doi:10.1109/ACCESS.2017.2685687
4. Andreev S., Galinina O., Pyattaev A., Hosek J., Masek P., Yanikomeroğlu H., and Koucheryavy Y. Exploring synergy between communications, caching, and computing in 5G-grade deployments. *IEEE Communications Magazine*, 2016, vol. 54(8), pp. 60–69. doi:10.1109/MCOM.2016.7537178
5. Recommendation ITU-R M.2083-0. IMT Vision — Framework and overall objectives of the future development of IMT for 2020 and beyond. Series M. 2083 p.
6. Popovski P., Trillingsgaard K. F., Simeone O., and Durisi G. 5G wireless network slicing for eMBB, URLLC, and mMTC: A communication-theoretic view. *IEEE Access*, vol. 6, pp. 55765–55779. doi:10.1109/ACCESS.2018.2872781
7. Clazzer F., Munari A., Liva G., Lazaro F., Stefanovic C., and Popovski P. From 5G to 6G: Has the time for modern random access come? *arXiv preprint arXiv:1903.0306*, 2019.
8. Galinina O., Andreev S., Tyurlikov A. On multiple access of a large number of machine-type devices in cellular networks. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 4, pp. 105–114 (In Russian) doi:10.31799/1684-8853-2018-4-105-114
9. Wu Y., Gao X., Zhou S., Yang W., Polyanskiy Y., and Caire G. Massive access for future wireless communication systems. *IEEE Wireless Communications*, 2020, vol. 27(4), pp. 148–156. doi:10.1109/MWC.001.1900494
10. Zaidi A., Hussain Y., Hogan M., and Kuhlins C. Cellular IoT evolution for industry digitalization. *Ericsson White Paper, GFMC-19*, 2019, p. 17. Available at: <https://www.ericsson.com/en/reports-and-papers/white-papers/cellular-iot-evolution-for-industry-digitalization> (accessed 23 August 2021).
11. Vaezi M., Azari A., Khosravirad S. R., Shirvanimoghaddam M., Azari M. M., Chasaki D., and Popovski P. Cellular, Wide-Area, and Non-Terrestrial IoT: A Survey on 5G Advances and the Road Towards 6G. *arXiv preprint arXiv:2107.03059*, 2021.
12. Borisovskaya A. V., and Turlikov A. M. Reducing energy consumption in the IoT systems with unlimited number of devices. *Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)*, IEEE, 2021, pp. 1–6. doi:10.1109/WECONF51603.2021.9470596
13. Matveev N. V., Turlikov A. M. Review of random multiple access methods for massive machine type communication. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2019, no. 6, pp. 54–67 (In Russian). doi:10.31799/1684-8853-2019-6-54-67
14. Matveev N. V., Turlikov A. M. Slotted ALOHA with iterative procedure for resolving collisions. Stability and non-stability. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2018, no. 3, pp. 89–97 (In Russian) doi:10.15217/issn1684-8853.2018.3.89

15. Foss S., Turlikov A., and Grankin M. Spatial random multiple access with multiple departure. *IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 2728–2731. doi:10.1109/ISIT.2017.8007025
16. Polyanskiy Y. A perspective on massive random-access. *IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 2523–2527. doi:10.1109/ISIT.2017.8006984
17. Ordentlich O., Polyanskiy Y. Low complexity schemes for the random access Gaussian channel. *IEEE International Symposium on Information Theory (ISIT)*, IEEE, 2017, pp. 2528–2532. doi:10.1109/ISIT.2017.8006985
18. Vvedenskaya N. D., Tsybakov B. S. Packet delay in the case of a multiple-access stack algorithm. *Problems Inform. Transmission*, 1984, no. 20:2, pp. 77–97.
19. Tsybakov B. Survey of USSR contributions to random multiple-access communications. *IEEE Transactions on Information Theory*, 1985, vol. 31, no. 2, pp. 143–165. doi:10.1109/TIT.1985.1057023
20. Burkov A., Shneer S., Turlikov A. An achievability bound of energy per bit for stabilized massive random access Gaussian channel. *IEEE Communications Letters*, 2020, vol. 25, no. 1, pp. 299–302. doi:10.1109/LCOMM.2020.3023461

УДК 004.728.3.057.4

doi:10.31799/1684-8853-2021-5-51-58

Задачи оптимизации мощности сигнала и энергии на бит в системах mMTC

А. А. Бурков^а, ассистент, orcid.org/0000-0002-0920-585X, a.burkov@k36.org

^аСанкт-Петербургский государственный университет аэрокосмического приборостроения, Б. Морская ул., 67, Санкт-Петербург, 190000, РФ

Введение: в настоящее время активно изучаются вопросы работы технологии Интернета вещей. Работа большого числа различных датчиков с автономным питанием находится в рамках сценария massive machine-type communications с применением методов случайного доступа. Актуальными вопросами в данном типе связи являются уменьшение мощности передачи и увеличение продолжительности работы устройства за счет снижения затрат энергии на бит. **Цель:** формулирование и анализ задач минимизации мощности передачи и затрачиваемой энергии на бит в системах без повторных передач и с повторными передачами для получения границ достижимости. **Результаты:** описана модель системы, в рамках которой сформулированы и рассмотрены четыре задачи минимизации мощности и затрат энергии при заданных параметрах (число информационных бит, спектральная эффективность системы и требуемая достоверность доставки). Представлены численные результаты решения данных оптимизационных задач, позволяющие получить границы достижимости для указанных характеристик в системах с потерями и без потерь. Также определены нижние границы для данных характеристик, полученные по формуле Шеннона в предположении, что длина сообщения не ограничена. Результаты показали, что решение задачи минимизации по одному из параметров (мощность сигнала или энергия потребления на бит) не минимизирует второй параметр. Эта разница наиболее существенна для информационных сообщений небольшой длины, что соответствует сценариям Интернета вещей. **Практическая значимость:** полученные результаты позволяют оценить потенциальные возможности для минимизации мощности передачи и затрат энергии на бит в системах случайного множественного доступа со сценариями massive machine-type communications. **Обсуждение:** представленные задачи решались без учета средней задержки передачи сообщения, введение такого ограничения должно повысить мощность передаваемого сигнала и затраты энергии на бит.

Ключевые слова — случайный множественный доступ, спектральная эффективность, отношение сигнал/шум, энергия на бит, mMTC, Интернет вещей, задача минимизации.

Для цитирования Burkov A. A. Signal power and energy-per-bit optimization problems in mMTC systems. *Информационно-управляющие системы*, 2021, № 5, с. 51–58. doi:10.31799/1684-8853-2021-5-51-58

For citation: Burkov A. A. Signal power and energy-per-bit optimization problems in mMTC systems. *Informatsionno-upravliaiushchie sistemy* [Information and Control Systems], 2021, no. 5, pp. 51–58. doi:10.31799/1684-8853-2021-5-51-58

БАБКИН
Эдуард
Александрович



Профессор кафедры информационных систем и технологий, заведующий лабораторией теории и практики систем поддержки принятия решений Нижегородского филиала ВШЭ. В 1993 году окончил Нижегородский государственный университет им. Н. И. Лобачевского по специальности «Системы научно-технической информации». В 2001 году защитил диссертацию на соискание ученой степени кандидата технических наук, в 2007 году — на соискание ученой степени PhD of Computer Science. Является автором 116 научных публикаций. Область научных интересов — архитектура распределенных систем, информационное моделирование, методы представления и обработки знаний. Эл. адрес: eababkin@hse.ru

ВОСТРИКОВ
Антон
Александрович



Доцент кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2000 году окончил Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Вычислительные машины, комплексы, системы и сети». В 2004 году защитил диссертацию на соискание ученой степени кандидата технических наук. Является автором 45 научных публикаций и двух свидетельств о регистрации программного продукта. Область научных интересов — распределенные и встраиваемые информационно-управляющие системы, обработка визуальной информации, опτικο-информационные системы. Эл. адрес: vostricov@mail.ru

КОЛПАЩИКОВ
Леонид
Александрович



Руководитель научного отдела ФГБУ «Заповедники Таймыра», Норильск. Член-корреспондент Международной академии минеральных ресурсов, РАЕН по отделению «Охрана природы и биоразнообразия». В 1970 году окончил Кировский сельскохозяйственный институт по специальности «Биолог-охотовед». В 2001 году защитил диссертацию на соискание ученой степени доктора биологических наук. Является автором 250 научных публикаций и восьми монографий. Область научных интересов — зоология, промысловая биология, зоогеография, экологический мониторинг с использованием спутниковых ошейников, традиционное природопользование, методические вопросы организации авиаучетов животных и птиц. Эл. адрес: ntnt69@yandex.ru

БУРКОВ
Артем
Андреевич



Ассистент кафедры инфокоммуникационных систем Санкт-Петербургского государственного университета аэрокосмического приборостроения. В 2017 году окончил с отличием Санкт-Петербургский государственный университет аэрокосмического приборостроения по специальности «Инфокоммуникационные технологии и системы связи». Является автором десяти научных публикаций. Область научных интересов — беспроводные системы передачи данных, алгоритмы случайного множественного доступа, системы с гибридной решающей обратной связью, цифровая обработка изображений, сжатие данных. Эл. адрес: a.burkov@k36.org

ДЕМИДОВСКИЙ
Александр
Владимирович



Аспирант, старший преподаватель кафедры информационных систем и технологий Нижегородского филиала ВШЭ. В 2017 году окончил магистратуру Нижегородского филиала ВШЭ по специальности «Бизнес-информатика». Является автором 34 научных публикаций. Область научных интересов — представление и обработка данных и знаний с неопределенностью, разработка нейросетевых архитектур, методы принятия решений. Эл. адрес: ademidovskij@hse.ru

МИХАЙЛОВ
Владимир
Валентинович



Профессор кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения, ведущий научный сотрудник Санкт-Петербургского института информатики и автоматизации РАН. В 1957 году окончил Ленинградский электротехнический институт по специальности «Математические и счетно-решающие приборы и устройства». В 1999 году защитил диссертацию на соискание ученой степени доктора технических наук. Является автором более 190 научных публикаций. Область научных интересов — компьютерное моделирование в области экологии и робототехники, разработка систем автоматизации моделирования. Эл. адрес: mwwcari@gmail.com

НГУЕН
Тхань Конг



Эксперт департамента науки, технологий и международного сотрудничества Данангского университета науки и технологий, Данангский университет, Дананг, Вьетнам.
В 2013 году окончил магистратуру Университета Юань Цзе, Тайвань, по специальности «Машиностроение».
Область научных интересов — обработка изображений.
Эл. адрес: nthcongk@dut.udn.vn

СОБОЛЕВСКИЙ
Владислав
Алексеевич



Младший научный сотрудник, аспирант Санкт-Петербургского федерального исследовательского центра РАН.
В 2017 году окончил магистратуру Санкт-Петербургского государственного технологического института (Технический университет) по специальности «Информатика и вычислительная техника».
Является автором 16 научных публикаций.
Область научных интересов — искусственный интеллект, искусственные нейронные сети, глубокое обучение, системы data mining.
Эл. адрес: arguzd@yandex.ru

СОЛОВЬЕВ
Николай
Владимирович



Доцент кафедры вычислительных систем и сетей Санкт-Петербургского государственного университета аэрокосмического приборостроения, почетный работник высшего профессионального образования РФ.
В 1979 году окончил Ленинградский политехнический институт по специальности «Автоматизация и комплексная механизация машиностроения».
В 2004 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором более 40 научных публикаций и четырех патентов на изобретения.
Область научных интересов — распознавание образов, компьютерная обработка и анализ изображений, техническое зрение роботов.
Эл. адрес: famsol@yandex.ru

СУЛТАНОВ
Альберт
Ханович



Профессор, заведующий кафедрой телекоммуникационных систем Уфимского государственного авиационного технического университета, заслуженный деятель науки Республики Башкортостан, заслуженный работник высшей школы РФ.
В 1973 году окончил Новосибирский электротехнический институт связи по специальности «Многоканальная электросвязь».
В 1997 году защитил диссертацию на соискание ученой степени доктора технических наук.
Является автором более 200 научных публикаций и 16 патентов на изобретения.
Область научных интересов — полупроводниковая и антенная техника, оптика, связь, обработка изображений.
Эл. адрес: tks@ugatu.ac.ru

ТИМОФЕЕВ
Александр
Леонидович



Доцент кафедры электроники и биомедицинских технологий Уфимского государственного авиационного технического университета.
В 1977 году окончил Уфимский авиационный институт по специальности «Информационно-измерительная техника».
В 1984 году защитил диссертацию на соискание ученой степени кандидата технических наук.
Является автором 52 научных публикаций и 13 авторских свидетельств на изобретения.
Область научных интересов — помехоустойчивое кодирование, связь, обработка изображений, цифровая голография.
Эл. адрес: a_1_t@inbox.ru

ФАМ
Конг Тханг



Преподаватель факультета информационных технологий Данангского университета науки и технологий, Дананг, Вьетнам.
В 2013 году окончил Тульский государственный университет по специальности «Вычислительные машины, комплексы, системы и сети».
В 2016 году защитил диссертацию на соискание ученой степени кандидата технических наук в Тульском государственном университете.
Является автором 20 научных публикаций.
Область научных интересов — обработка изображений, машинное обучение, наука о данных.
Эл. адрес: pcthang@dut.udn.vn

ФАМ
Минь Туан



Преподаватель факультета информационных технологий Данангского университета науки и технологий, Дананг, Вьетнам. В 2007 году окончил университет Миэ, Япония, по специальности «Информационные технологии». В 2012 году защитил диссертацию на соискание ученой степени кандидата технических наук в Нагойском университете, Япония. Является автором 50 научных публикаций. Область научных интересов — искусственный интеллект, мягкие вычисления, геометрическая алгебра.
Эл. адрес: pmtuan@dut.udn.vn

ЧАН
Тхи Тху Тхао



Преподаватель факультета статистики и информатики Экономического университета, Данангский университет, Дананг, Вьетнам. В 2018 году окончила магистратуру Тульского государственного университета по специальности «Прикладная математика и информатика». Является автором десяти научных публикаций. Область научных интересов — обработка изображений, машинное обучение.
Эл. адрес: thaotran@due.udn.vn

ЯКУШЕВ
Георгий
Кириллович



Инженер-программист ООО «Специальный технологический центр», Санкт-Петербург. В 2021 году окончил бакалавриат Санкт-Петербургского государственного университета аэрокосмического приборостроения по специальности «Информатика и вычислительная техника». Область научных интересов — распознавание образов, нейронные сети, компьютерная обработка и анализ растровых изображений.
Эл. адрес: georgyyakushev04@yandex.ru

УВАЖАЕМЫЕ АВТОРЫ!

Научная электронная библиотека (НЭБ) продолжает работу по реализации проекта SCIENCE INDEX. После того как Вы регистрируетесь на сайте НЭБ (<http://elibrary.ru/defaultx.asp>), будет создана Ваша личная страничка, содержание которой составят не только Ваши персональные данные, но и перечень всех Ваших печатных трудов, имеющих в базе данных НЭБ, включая диссертации, патенты и тезисы к конференциям, а также сравнительные индексы цитирования: РИНЦ (Российский индекс научного цитирования), h (индекс Хирша) от Web of Science и h от Scopus. После создания базового варианта Вашей персональной страницы Вы получите код доступа, который позволит Вам редактировать информацию, помогая создавать максимально объективную картину Вашей научной активности и цитирования Ваших трудов.

Уважаемые авторы!

При подготовке рукописей статей необходимо руководствоваться следующими рекомендациями.

Статьи должны содержать изложение новых научных результатов. Название статьи должно быть кратким, но информативным. В названии недопустимо использование сокращений, кроме самых общепринятых (РАН, РФ, САПР и т. п.).

Текст рукописи должен быть оригинальным, а цитирование и самоцитирование корректно оформлено.

Объем статьи (текст, таблицы, иллюстрации и библиография) не должен превышать эквивалента в 20 страниц, напечатанных на бумаге формата А4 на одной стороне через 1,5 интервала Word шрифтом Times New Roman размером 13, поля не менее двух сантиметров.

Обязательными элементами оформления статьи являются: индекс УДК, заглавие, инициалы и фамилия автора (авторов), ученая степень, звание (при отсутствии — должность), полное название организации, аннотация и ключевые слова на русском и английском языках, ORCID и электронный адрес одного из авторов. При написании аннотации не используйте аббревиатур и не делайте ссылок на источники в списке литературы. Предоставляйте подрисовочные подписи и названия таблиц на русском и английском языках.

Статьи авторов, не имеющих ученой степени, рекомендуется публиковать в соавторстве с научным руководителем, наличие подписи научного руководителя на рукописи обязательно; в случае самостоятельной публикации обязательно предоставляйте заверенную по месту работы рекомендацию научного руководителя с указанием его фамилии, имени, отчества, места работы, должности, ученого звания, ученой степени.

Простые формулы набирайте в Word, сложные с помощью редактора MathType или Equation. Для набора одной формулы не используйте два редактора; при наборе формул в формульном редакторе знаки препинания, ограничивающие формулу, набирайте вместе с формулой; для установки размера шрифта в MathType никогда не пользуйтесь вкладкой Other, Smaller, Larger, используйте заводские установки редактора, не подгоняйте размер символов в формулах под размер шрифта в тексте статьи, не растягивайте и не сжимайте мышью формулы, вставленные в текст; пробелы в формуле ставьте только после запятой при перечислении с помощью Ctrl+Shift+Space (пробел); не отделяйте пробелами знаки: + = - ×, а также пространство внутри скобок; для выделения греческих символов в MathType полужирным начертанием используйте Style → Other → bold.

Для набора формул в Word никогда не используйте вкладки: «Уравнение», «Конструктор», «Формула» (на верхней панели: «Вставка» — «Уравнение»), так как этот ресурс предназначен только для внутреннего использования в Word и не поддерживается программами, предназначенными для изготовления оригинал-макета журнала.

При наборе символов в тексте помните, что символы, обозначаемые латинскими буквами, набираются светлым курсивом, русскими и греческими — светлым прямым, векторы и матрицы — прямым полужирным шрифтом.

Подробнее см. pdf-файл «Правила подготовки рукописей» (стр. 11) на сайте <https://guar.ru/ric>

Иллюстрации:

— рисунки, графики, диаграммы, блок-схемы предоставляйте в виде отдельных исходных файлов, поддающихся редактированию, используя векторные программы: Visio (*.vsd, *.vsdx); Adobe Illustrator (*.ai); CorelDraw (*.cdr, версия не выше 15); Excel (*.xls); Word (*.docx); AutoCad, Matlab (экспорт в PDF, EPS, SVG, WMF, EMF); Компас (экспорт в PDF), веб-портал DRAW.IO (экспорт в PDF);

— фото и растровые — в формате *.tif, *.png с максимальным разрешением (не менее 300 pixels/inch).

Наличие подрисовочных подписей и названий таблиц на русском и английском языках обязательно (желательно не повторяющих дословно комментарии к рисункам в тексте статьи).

В редакцию предоставляются:

— сведения об авторе (фамилия, имя, отчество, место работы, должность, ученое звание, учебное заведение и год его окончания, ученая степень и год защиты диссертации, область научных интересов, количество научных публикаций, домашний и служебный адреса и телефоны, e-mail), фото авторов: анфас, в темной одежде на белом фоне, должны быть видны плечи и грудь, высокая степень четкости изображения без теней и отблесков на лице, фото можно представить в электронном виде в формате *.tif, *.png, *.jpg с максимальным разрешением — не менее 300 pixels/inch при минимальном размере фото 40×55 мм;

— экспертное заключение.

Список литературы составляется по порядку ссылок в тексте и оформляется следующим образом:

— для книг и сборников — фамилия и инициалы авторов, полное название книги (сборника), город, издательство, год, общее количество страниц, doi;

— для журнальных статей — фамилия и инициалы авторов, полное название статьи, название журнала, год издания, номер журнала, номера страниц, doi;

— ссылки на иностранную литературу следует давать на языке оригинала без сокращений;

— при использовании web-материалов указывайте адрес сайта и дату обращения.

Список литературы оформляйте двумя отдельными блоками по образцам lit.dot на сайте журнала (<http://i-us.ru/paperrules>): Литература и References.

Более подробно правила подготовки текста с образцами изложены на нашем сайте в разделе «Руководство для авторов».

Контакты

Куда: 190000, Санкт-Петербург,
Б. Морская ул., д. 67, ГУАП, РИЦ
Кому: Редакция журнала «Информационно-управляющие системы»
Тел.: (812) 494-70-02
Эл. почта: ius.spb@gmail.com
Сайт: www.i-us.ru