



<http://dx.doi.org/10.35596/1729-7648-2024-22-2-55-69>

Оригинальная статья
Original paper

УДК 004.932+004.258

ПЕРСПЕКТИВНЫЕ РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЙ И РАЗРАБОТКИ В ОБЛАСТИ ОБРАБОТКИ ИЗОБРАЖЕНИЙ И РЕЧЕВЫХ СИГНАЛОВ

Д. С. ЛИХАЧЁВ, Ц. МА, Н. А. ПЕТРОВСКИЙ, И. С. АЗАРОВ, В. Ю. ЦВЕТКОВ

*Белорусский государственный университет информатики и радиоэлектроники
(г. Минск, Республика Беларусь)*

Поступила в редакцию 30.01.2024

© Белорусский государственный университет информатики и радиоэлектроники, 2024
Belarusian State University of Informatics and Radioelectronics, 2024

Аннотация. Представлен анализ перспектив развития технологий обработки изображений и речевых сигналов. Приведены основные результаты по данным направлениям, полученные за последние годы в соответствующих научных школах Белорусского государственного университета информатики и радиоэлектроники. Показано, что использование технологий машинного обучения в сочетании с методами цифровой обработки изображений и речевых сигналов позволяет существенно повысить эффективность систем их распознавания и классификации.

Ключевые слова: классификация изображений, скелетизация изображений, классификация речевых сигналов, оценка огибающей спектра авторегрессионного процесса.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Перспективные результаты исследований и разработки в области обработки изображений и речевых сигналов / Д. С. Лихачёв [и др.] // Доклады БГУИР. 2024. Т. 22, № 2. С. 55–69. <http://dx.doi.org/10.35596/1729-7648-2024-22-2-55-69>.

PROMISING RESEARCH AND DEVELOPMENT RESULTS IN THE FIELD OF IMAGE AND SPEECH SIGNAL PROCESSING

DENIS S. LIKHACHOV, JUN MA, NICK A. PETROVSKY, ILYA S. AZAROV,
VIKTAR YU. TSVIATKOU

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 30.01.2024

Abstract. An analysis of the prospects for the development of technologies for processing images and speech signals is presented. The main results in these areas obtained in recent years in the relevant scientific schools of Belarusian State University of Informatics and Radioelectronics are presented. It is shown that the use of machine learning technologies in combination with methods of digital processing of images and speech signals can significantly increase the efficiency of systems for their recognition and classification.

Keywords: image classification, image skeletonization, speech signal classification, estimation of the spectrum envelope of an autoregressive process.

Conflict of interests. The authors declare no conflict of interests.

For citation. Likhachov D. S., J. Ma, Petrovsky N. A., Azarov I. S., Tsviatkou V. Yu. (2024) Promising Research and Development Results in the Field of Image and Speech Signal Processing. *Doklady BGUIR*. 22 (2), 55–69. <http://dx.doi.org/10.35596/1729-7648-2024-22-2-55-69> (in Russian).

Введение

Развитие цифровой обработки сигналов привело к выделению нескольких самостоятельных направлений, таких как обработка изображений и обработка голоса и речи. Цифровая обработка изображений (ЦОИ) связана с их преобразованием, улучшением, анализом и интерпретацией с использованием различных алгоритмов, методов и инструментов. Она включает изменение или извлечение информации из цифровых изображений для улучшения их качества, извлечение полезных данных или повышение их пригодности для конкретных приложений.

Достижения в технологиях обработки изображений (камеры с высоким разрешением и улучшенные алгоритмы обработки) способствуют повышению качества изображений и возможностей их обработки. Рынок обработки изображений стремительно расширяется из-за резкого роста видеоаналитики и необходимости извлечения ценной информации из визуальных данных, поддерживая процессы принятия решений в различных отраслях. В здравоохранении ЦОИ является неотъемлемой частью медицинской визуализации, диагностики и исследований, что вносит значительный вклад в повышение рыночного спроса. Аналогичным образом такие отрасли, как развлечения, игры, дополненная и виртуальная реальность, в значительной степени опираются на эти технологии для создания реалистичных визуальных впечатлений. Приложения безопасности развиваются благодаря внедрению технологий обработки изображений, используемых для обнаружения объектов, распознавания лиц и видеомониторинга. Автомобильные системы безопасности, робототехника, дистанционное зондирование применяют ЦОИ для различных приложений. Поскольку эти технологии становятся все более доступными, их внедрение в различных секторах экономики продолжает расти, способствуя расширению рынка.

В цифровой обработке голоса и речи особое место принадлежит распознаванию, которое относится к технологии аутентификации, помогающей получать и интерпретировать человеческий голос и выполнять устные команды. Эта технология переводит голосовые команды в электрические сигналы, преобразует их в командные коды и отправляет на устройство в цифровом формате для исполнения. Она использует машинное обучение и искусственный интеллект (ИИ) для понимания разговорных выражений и направлена на повышение эффективности бизнеса за счет автоматизации таких задач, как расшифровка, ввод данных, планирование производственных процессов и технологических операций, позволяя пользователям сосредоточиться на более сложных задачах и повышая их общую производительность. Распознавание голоса и речи помогает людям с ограниченными возможностями и тем, кто испытывает трудности с набором текста, общаться более легко и эффективно.

Краткий обзор рынка технологий обработки изображений

По результатам исследований GII (Global Information, Inc.), глобальный спрос на рынок ЦОИ достигнет объема 27,26 млрд дол. США к 2030 г. (с 6,28 млрд дол. в 2022-м) со среднегодовым темпом роста 20,14 % в течение 2023–2030 гг. [1]. Ведущие поставщики программного обеспечения по обработке изображений – Nvidia Corporation, Canon Inc., Samsung Electronics Co. Ltd., Panasonic Corporation, Cognex Corporation, Teledyne Technologies, Vegavid Technology, Visionary.ai., IBM, Microsoft. GII отмечает растущую потребность в ЦОИ в секторе здравоохранения, что указывает на обширную тенденцию к технологическому прогрессу для оптимизации оказания медицинской помощи, повышения точности диагностики и улучшения ухода за пациентами. Прогнозируется, что к 2030 г. сегмент здравоохранения будет занимать примерно 25,9 % рынка. Ожидается, что по мере развития технологий дополнительные достижения в области ЦОИ будут иметь все большее значение, влияя на развитие здравоохранения. Некоторые факторы, влияющие на рынок, – это достижения в области ИИ и машинного обучения в ЦОИ, рост уровня использования мобильных устройств и высокая первоначальная стоимость ЦОИ [1, 2].

Достижения в области ИИ и машинного обучения существенно расширили сферу ЦОИ, произведя революцию в анализе, интерпретации и визуализации. Интеграция ИИ и машинного обучения стимулирует спрос на продукты и услуги ЦОИ. Рост качества изображений расширил возможности применения методов их цифровой обработки благодаря высокой детализации визуальных данных. Смартфоны оснащаются все более совершенными камерами, способными снимать изображения с высоким разрешением. На них доступно множество ориентированных на потребителя приложений и программного обеспечения для редактирования изображений. Таким образом, растущий уровень использования мобильных устройств тоже стимулирует рост рынка [1].

Сдерживающим фактором развития рынка ЦОИ, по мнению ГИ, является высокая первоначальная стоимость установки систем ЦОИ, включая установку качественных камер и датчиков, необходимых для захвата и обработки изображений или видео высокого разрешения [1]. По принципу применения ГИ сегментировала рынок ЦОИ на визуальный поиск продуктов, распознавание объектов, зрительную аналитику и др. Наибольшая доля выручки на рынке в 2022 г. зафиксирована в сегменте распознавания объектов, что можно объяснить растущей потребностью в технологиях распознавания объектов в различных секторах и применениях. Например, его можно использовать в розничной торговле для управления запасами, мониторинга полок и анализа поведения покупателей. Распознавание объектов может улучшить контроль качества, обнаружить дефекты и автоматизировать операции сборки на производстве. Распознавание объектов важно для современных систем помощи водителю и беспилотных транспортных средств в автомобилестроении [1].

В зависимости от конечного использования ГИ подразделяет рынок ЦОИ на здравоохранение, автомобили и транспорт, BFSI (банковские, финансовые услуги и страхование), розничную торговлю, электронную коммерцию и т. д. В 2022 г. в сегменте розничной торговли и электронной коммерции зарегистрирована самая высокая доля выручки на рынке. Эти секторы значительно выросли во многом благодаря потребности в улучшенном визуальном контенте, повышению качества обслуживания клиентов и оптимизации операций. Технологии ЦОИ стали ключевыми в онлайн-торговле, увеличивая продажи, улучшая взаимодействие с клиентами и повышая операционную эффективность [1].

Скелетизация изображений

Основу видеоаналитики составляют технологии выделения, параметризации и идентификации объектов. Идентификация использует характеристики распределения яркости (цвета) и формы областей изображения. Для компактного представления формы однородных областей широко применяются скелеты – множество тонких (в пределе однопиксельных) связанных линий, заключенных внутри области и описывающих ее структуру (рис. 1). Скелеты, используемые в распознавании образов, должны быть стабильными при изменении контраста и действии шума: мультипликативного и аддитивного, проявляющегося на границах и внутри скелетизируемой области соответственно.

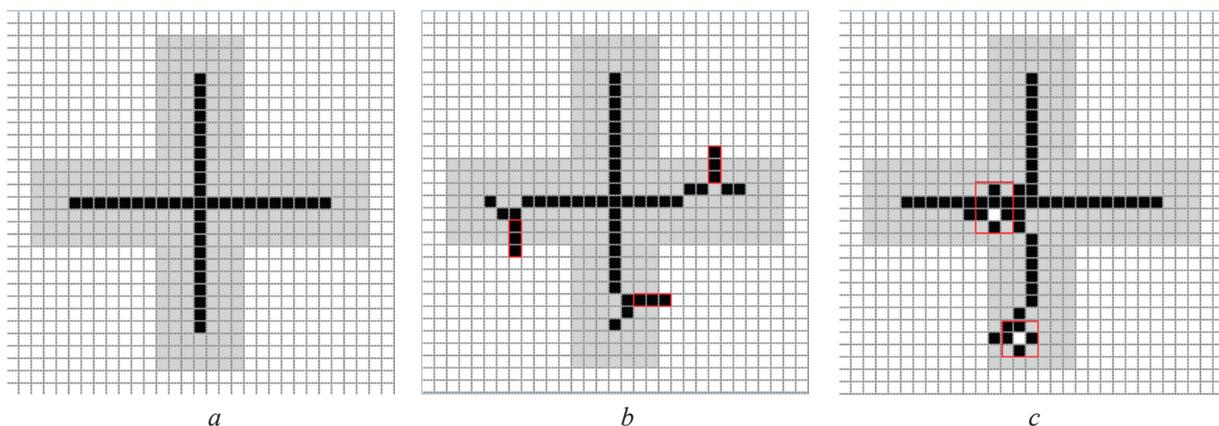


Рис. 1. Скелеты областей бинарных изображений:
a – без шума; *b, c* – с мультипликативным и аддитивным шумами соответственно

Fig. 1. Skeletons of binary image regions:
a – without noise; *b, c* – with multiplicative and additive noise, respectively

Методы формирования скелетов берут начало от [3] и делятся на три основных типа, основанных на: геометрических свойствах симметрии [4], построении непрерывных кривых [5], бинарной эрозии [6]. Скелеты, формируемые методами первых двух типов, характеризуются высокой избыточностью и низкой устойчивостью. Методы третьего типа могут распараллеливаться и подавлять незначительный мультипликативный шум, возникающий из-за деформации границ объектов при изменении условий формирования изображений (освещенность, перемещение камеры) [7]. Однако этим методам, как и методам других типов, свойственна избыточная скелети-

зация (формирование лишних ветвей скелетов). Для повышения устойчивости скелетов к мультипликативному шуму и исключения избыточных ветвей используются методы обрезки, основанные на дискретной эволюционной кривой [8], оценке изгиба [9], оценке значимости ветвей для реконструкции объекта, описываемого скелетом [10], сортировке по значимости элементов скелета [11]. Основным недостатком данных методов является необходимость ручного подбора параметров с учетом статистических особенностей изображения.

В качестве примера на рис. 2 приведены результаты формирования скелетов с помощью метода DCE (Discrete Curve Evolution), основанного на дискретной эволюционной кривой (ДЭК) [12–14]. ДЭК – виртуальная замкнутая ломаная, составленная из ребер k -угольника, вершинами которого (точками излома виртуальной ломаной) являются выступающие пиксели Р-контура скелетизируемой области бинарного изображения. В качестве точек излома ДЭК выбираются k выступающих пикселей, имеющих наибольшие значения параметра $K(r_1, r_2)$, определяемого для смежных ребер r_1 и r_2 ДЭК с помощью выражения:

$$K(r_1, r_2) = \frac{\beta(r_1, r_2)l(r_1)l(r_2)}{l(r_1) + l(r_2)}, \quad (1)$$

где $\beta(S_1, S_2)$ – угол между смежными ребрами r_1 и r_2 ; $l(r_1)$, $l(r_2)$ – длины смежных ребер r_1 и r_2 .

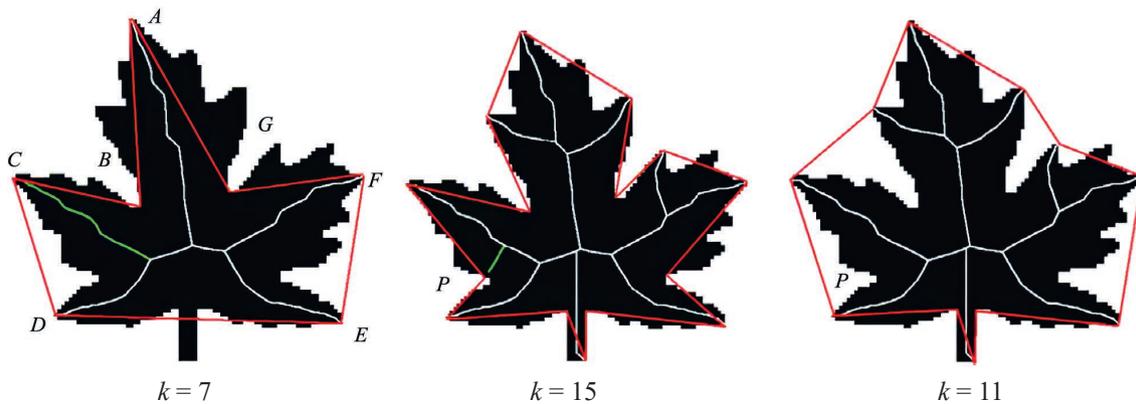


Рис. 2. Скелеты и дискретные эволюционные кривые (красные линии) при различных значениях k

Fig. 2. Skeletons and discrete evolutionary curves (red lines) at different values of k

Согласно методу DCE, в исходном скелете удаляются все ветви, кроме k выбранных. Значение k определяет оператор исходя из результатов визуальной оценки степени обрезки скелета. Из рис. 2 следует, что выбор значения параметра k существенно влияет на количество ветвей скелета и может приводить к недостаточной или избыточной скелетизации. Таким образом, значительные затраты времени при выборе k являются основным недостатком метода DCE.

Для автоматической (без участия оператора) обрезки избыточных ветвей скелетов предлагается метод DCEM (Discrete Curve Evolution Modification), основанный на оценке качества реконструкции изображения по скелету и являющийся модификацией метода DCE. Сущность DCEM состоит в определении формы ДЭК по результатам оценки точности восстановления контура области по скелету. Форма ДЭК определяется количеством и положением точек излома, выбираемых из множества выступающих пикселей P . Для параметризации выступающих пикселей в DCEM используется выражение (1) из метода DCE. Выбор значения k в предлагаемом методе основан на сравнении размеров и форм исходной скелетизируемой области и области, восстановленной по обрзанному скелету, состоящему из k ветвей. Метод DCEM состоит из следующих шагов.

1. Выделяется скелетизируемая область A_B бинарного изображения.
2. Выделяется контур C_B области A_B .
3. На контуре C_B ищутся выступающие пиксели P .
4. Формируется скелет S_B области A_B . Для этого используется метод ОРСА [15].
5. На скелете S_B детектируются конечные точки E ветвей. Фиксируется количество N_E конечных точек E .

6. Устанавливаются связи между выступающими P и конечными E точками (используется минимум расстояния). Несвязанные выступающие точки удаляются.
 7. Для всех выступающих точек P вычисляются значения параметра $K(r_1, r_2)$ с помощью (1).
 8. Инициализация счетчика k циклов обрезки ветвей скелета: $k = 3$.
 9. Начало цикла обрезки ветвей скелета. Среди выступающих точек P выбираются k точек с наибольшими значениями $K(r_1, r_2)$.
 10. Для выбранных k точек излома формируется обрезанный скелет $S_C(k)$, содержащий только k ветвей, соответствующих k конечным точкам E , связанным с выбранными k выступающими точками P .
 11. Из скелета $S_C(k)$ восстанавливается область $A_R(k)$, соответствующая области A_B . Восстановление осуществляется на основе операции бинарной дилатации. Критерием прекращения дилатации является минимум разности $D_A(k)$ площадей исходной A_B и восстановленной $A_R(k)$ областей.
 12. Увеличение значения k : $k = k + 1$.
 13. Проверка условия окончания цикла обрезки ветвей скелета. Если $k < N_E$, – переход на шаг 9.
 14. Определяется значение n , соответствующее минимальному значению разностей $\{D_A(n)\}_{(n=3, N_E)}$ площадей исходной A_B и восстановленной $A_R(k)$ областей.
 15. Формирование результата: обрезанный скелет $S_C(n)$, содержащий n ветвей.
- На рис. 3 приведены результаты скелетизации бинарных изображений с помощью метода ОРСА [15], результаты обрезки с помощью методов DCEM и DCE скелетов, полученных методом, описанным в [16], превосходящим по точности ОРСА.

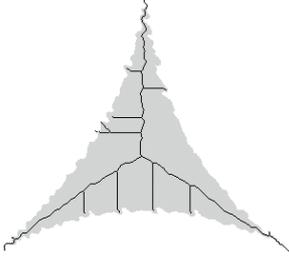
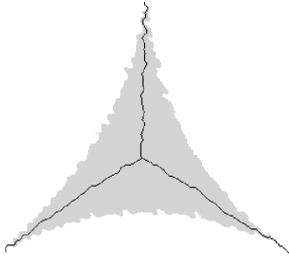
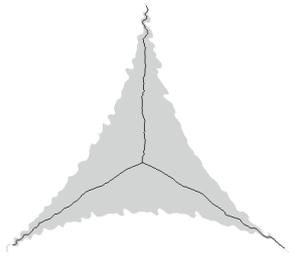
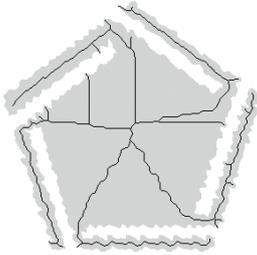
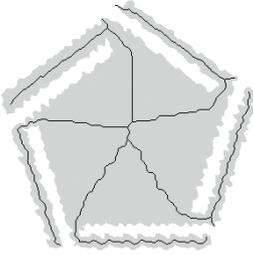
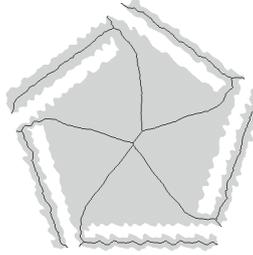
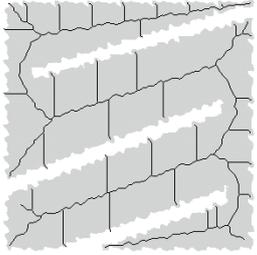
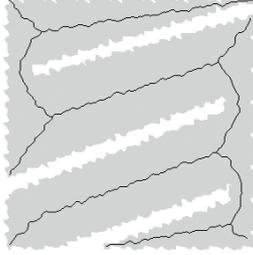
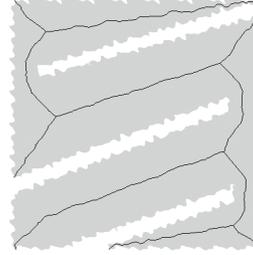
Скелеты, полученные с помощью метода [15]	Результаты метода обрезки DCEM для скелетов, полученных с помощью метода [16]	Результаты метода обрезки DCE для скелетов, полученных с помощью метода [16]
		 $k = 3$
		 $k = 15$
		 $k = 10$

Рис. 3. Результаты обрезки скелетов с помощью методов DCEM и DCE
Fig. 3. Results of skeleton trimming using DCEM and DCE methods

Из рис. 3 следует, что исходные бинарные изображения имеют значительный мультипликативный шум, приводящий к избыточной скелетизации, связанной с формированием большого количества лишних ветвей скелетов. Предложенный метод DCEM автоматической обрезки скелетов успешно справляется с устранением лишних ветвей, обеспечивая высокую точность формы скелетов, соизмеримую с точностью обрезанных скелетов, сформированных с помощью метода DCE на основе результатов более точного метода скелетизации [16].

Основным недостатком предложенного метода DCEM, как и базового DCE, а также других известных методов обрезки скелетов, является чувствительность к аддитивному шуму, возникающему внутри областей. Аддитивный шум, как и мультипликативный, приводит к избыточной скелетизации. Для формирования качественных скелетов в условиях аддитивного шума используются специально разработанные для таких условий методы скелетизации, например, МОРСА [17], а также методы, основанные на предварительной фильтрации изображений, например, АТФМ [18]. Для различных методов скелетизации бинарных изображений в табл. 1 приведены значения показателя R_{SV} искажений скелета в условиях мультипликативного (рис. 4) и аддитивного (рис. 5) шумов. Показатель R_{SV} вычисляется для скелетов S_N и S_C , сформированных в условиях шума и при его отсутствии соответственно, с помощью выражения:

$$R_{SV} = \frac{f_X(S_N, S_C) + f_X(S_C, S_N)}{2f_S(S_C)}, \quad (2)$$

где $f_X(S_A, S_B)$ – функция пересечения, определяющая количество пикселей изображения, которые принадлежат скелету S_A , но не принадлежат скелету S_B ($A = \{C, N\}$; $B = \{C, N\}$); $f_S(S_C)$ – функция, определяющая количество пикселей в скелете S_C .

Таблица 1. Оценка искажений скелетов в условиях мультипликативного и аддитивного шумов
Table 1. Estimate of skeleton distortions under conditions of multiplicative and additive noise

Метод скелетизации / Skeletonization method	R_{SV}	
	Мультипликативный шум (рис. 4) / Multiplicative noise (Fig. 4)	Аддитивный шум (рис. 5) / Additive noise (Fig. 5)
ОРТА	1,14	3,24
МОРСА	1,08	2,98
ОРТА + DCEM	0,74	2,89
ОРТА + АТФМ	0,85	0,01

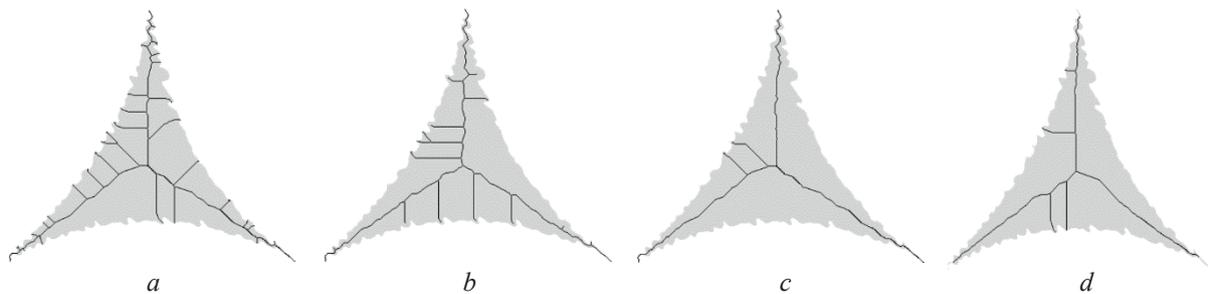


Рис. 4. Результаты скелетизации изображений в условиях мультипликативного шума методами:
 a – ОРТА; b – МОРСА; c – ОРТА + DCEM; d – ОРТА + АТФМ

Fig. 4. Results of image skeletonization in conditions of multiplicative noise at methods:
 a – ORTA; b – MORSA; c – ORTA + DCEM; d – ORTA + ATFM

Из табл. 1 следует, что в условиях мультипликативного шума обрезка скелетов более эффективна по сравнению с другими подходами к скелетизации. Использование метода DCEM в сочетании с наиболее простым и неустойчивым к шуму методом ОРТА [19] позволяет существенно уменьшить значение R_{SV} по сравнению с методами ОРТА, МОРСА и ОРТА в сочетании с АТФМ. Однако в условиях аддитивного шума эффективность сочетания ОРТА и DCEM резко снижается по сравнению с сочетанием ОРТА и АТФМ. Для формирования качественных скелетов в условиях совместного действия аддитивного и мультипликативного шумов необходимо использование сочетания методов DCEM и АТФМ.

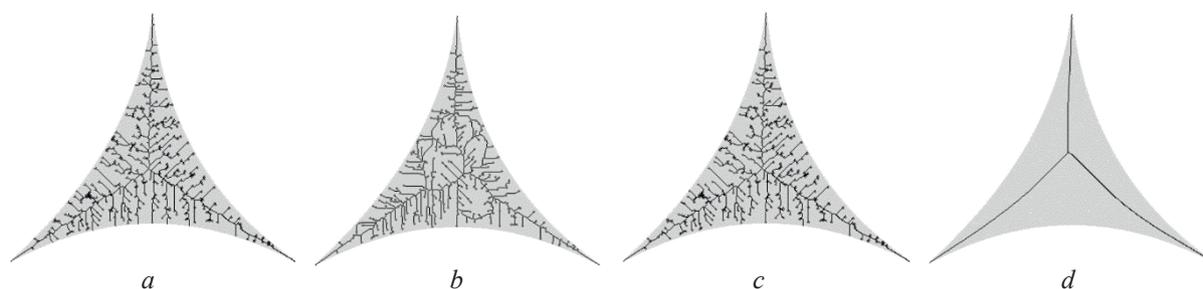


Рис. 5. Результаты скелетизации изображений в условиях аддитивного шума методами:
a – OPCA; *b* – MOPCA; *c* – OPCA + DCEM; *d* – OPCA + ATFM

Fig. 5. Results of image skeletonization under additive noise conditions at methods:
a – OPCA; *b* – MOPCA; *c* – OPCA + DCEM; *d* – OPCA + ATFM

Распознавание жестов рук на основе скелетов

Для приведенных на рис. 6 примеров десяти типов изображений жестов рук в табл. 2 представлены результаты статической классификации на основе различных методов скелетизации с использованием разных классификаторов: дерева решений (DT); *k*-ближайших соседей (KNN), простого Байеса (NB), машины опорных векторов (SVM), ансамблевого обучения (EL) и искусственной нейронной сети (ANN) [20].

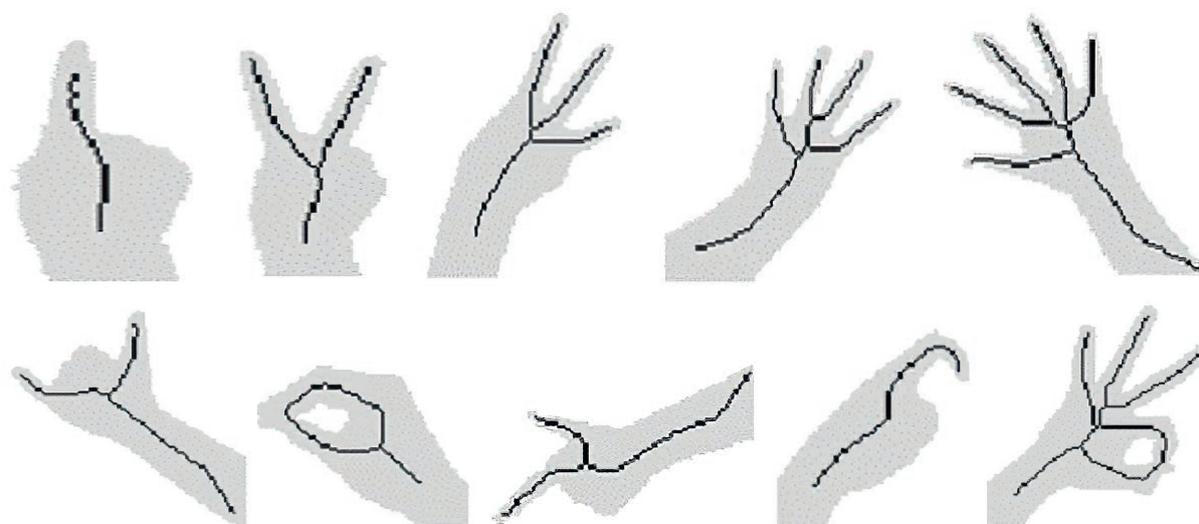


Рис. 6. Примеры скелетов десяти классов жестов рук
Fig. 6. Skeleton examples of ten hand gesture classes

Таблица 2. Оценка точности классификации жестов рук (%) по 10 классам
Table 2. Assessment of hand gesture classification accuracy (%) for 10 classes

Модель классификатора / Classifier model	Обучающий/проверочный наборы / Training/Validation Sets				Тестовый набор / Test set			
	Метод скелетизации / Skeletonization method				Метод скелетизации / Skeletonization method			
	OPCA+ATFM	OPCA+ATFM	MOPCA+ATFM	MOPCA+ATFM+DCEM	OPCA+ATFM	OPCA+ATFM	MOPCA+ATFM	MOPCA+ATFM+DCEM
DT	70,6	83,3	86,7	91,1	73,8	87,3	86,1	91,1
NB	61,8	76,1	77,4	82,9	62,0	81,4	80,2	77,2
SVM	59,9	77,4	80,4	81,6	62,4	75,5	82,3	81,0
KNN	63,9	75,5	83,8	82,7	63,7	75,9	87,8	84,4
EL	76,7	88,3	90,0	92,9	79,3	91,6	87,8	92,8
ANN	59,1	73,0	82,6	82,3	60,8	76,8	83,1	83,5
Среднее	65,2	78,9	83,4	85,5	67,0	81,4	84,5	85,0

Из табл. 2 следует, что сочетание методов МОРСА, АТФМ, DCEM и ансамблевого обучения обеспечивает более высокую точность классификации на тестовом наборе данных (92,8 %, что на 5–20 % лучше по сравнению с другими сочетаниями методов) за счет лучшей дискриминационной способности векторов признаков и устойчивости к совместному проявлению мультипликативного и аддитивного шумов. Время обучения классификаторов находится в диапазоне от 0,6 до 13,3 с (DT – 0,6 с; KNN – 1,1 с; EL – 4,4 с; NB – 5,2 с; SVM – 1,7 с; ANN – 13,3 с).

Краткий обзор рынка технологий распознавания голоса и речи

По результатам исследований ГП, рынок технологий распознавания голоса и речи в США составил 3,4 млрд долл. в 2022 г. [1]. Прогнозируется, что Китай, вторая по величине экономика мира, к 2030 г. достигнет прогнозируемого размера рынка в 5,3 млрд долл. По результатам исследований Fortune Business Insights, мировой рынок распознавания речи и голоса вырастет с 12,62 млрд долл. в 2023-м до 59,62 млрд долл. к 2030-му при среднегодовом темпе роста 24,8 % [2]. Ведущими поставщиками программного обеспечения для распознавания речи являются: 3M Co., Alphabet Inc., Amazon.com Inc., Apple Inc., Baidu Inc., Brainasoft, CastleOS Software LLC, Cisco Systems Inc., iFLYTEK Co. Ltd., International Business Machines Corp., LumenVox GmbH, Meta Platforms Inc., Microsoft Corp., Optimal Technology Solutions LLC, Paragon GmbH and Co. KGaA, Raytheon Technologies Corp., Sensory Inc., talk2web, Verint Systems Inc., Advanced Voice Recognition Systems Inc. [1].

Повышение уровня жизни, растущая индустрия средств массовой информации и развлечений стимулируют спрос на распознавание голоса и речи во всем мире. Ведущие компании-разработчики внедряют расширенные функции, такие как проактивные голосовые помощники и распознавание голоса. Возможности усиления системы безопасности за счет голосовой аутентификации способствуют внедрению технологий распознавания голоса и речи в сфере BFSI. Распространение автомобилей с бортовыми информационно-развлекательными системами, в которых голос используется для управления, также способствует росту данного сегмента рынка. Ожидается, что в ближайшие годы распознавание речи составит самый крупный сегмент рынка речевых технологий благодаря растущему количеству онлайн-транзакций и усилению требований к аутентификации для снижения риска кибератак [1].

Параунитарные банки фильтров на основе алгебры кватернионов

Полифазное представление критически децимированного банка фильтров является одной из возможных реализаций дискретного вейвлет-преобразования, особенно актуального для задач обработки цифровых сигналов. Однако существует ограничение на свойства совершенной реконструкции сигнала таким банком фильтров при числе каналов больше двух. Одна из существующих альтернатив параметризации ортогональной матрицы – представление парой единичных кватернионов, которые позволяют сохранить ортогональность искомого преобразования, несмотря на квантование. Решетчатая структура соответствующего банка показана на рис. 7.

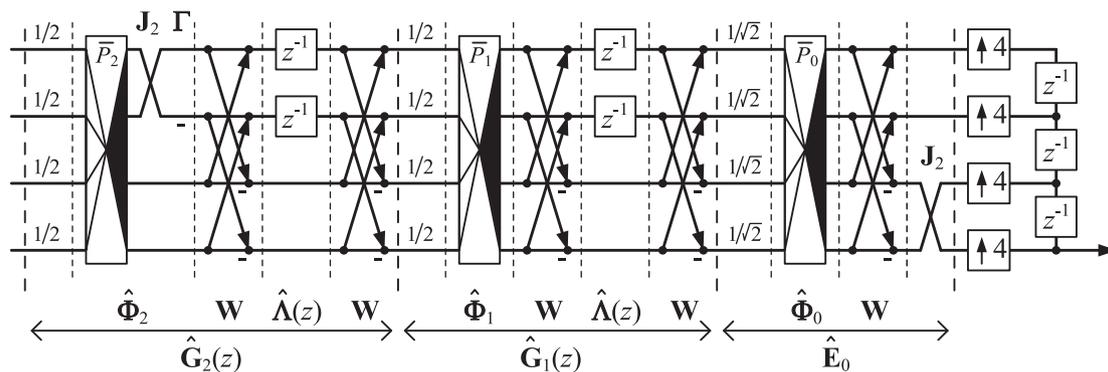


Рис. 7. Полифазная структура одномерного параунитарного банка фильтров
Fig. 7. Polyphase structure of a one-dimensional paraunitary filter bank

В свою очередь, одномерные параунитарные банки фильтров в алгебре кватернионов с линейной фазочастотной характеристикой (Q-ПУБФ) могут быть применены для создания многомерных разделимых систем, где 2D-сигналы отдельно преобразуются в вертикальном и горизонталь-

ном направлениях. Таким образом, Q-PUFB – это системы без потерь, несмотря на квантование коэффициентов, условие регулярности первого рода зафиксировано в значениях коэффициентов решетчатой структуры банка и легко удовлетворяется для арифметики с фиксированной запятой [21]. Однако многомерные сигналы, как правило, имеют неразделимую природу, и 2D-неразделимые банки фильтров более эффективны с точки зрения характеристики эффективности кодирования (Coding Gain) в отличие от разделимых, так как позволяют лучше учитывать пространственные изменения. Принимая во внимание преимущества Q-ПУБФ для задач сжатия и обработки, получена новая техника факторизации двумерных неразделимых кватернионных параунитарных банков фильтров (2D-NS Q-ПУБФ). Цель разработанной техники – обеспечить возможность формализовать описание 2D-неразделимого преобразования, имея полифазную передаточную матрицу синтеза и анализа Q-ПУБФ для одномерного случая.

Исходная факторизация (рис. 7) 4-канального разделимого банка фильтров Q-ПУБФ может быть преобразована в 2D NS (Non-Separable) Q-ПУБФ, в свою очередь, коэффициенты двумерной структуры могут быть синтезированы неразделимым образом. По аналогии есть возможность получить факторизацию 8-канального неразделимого 2D Q-ПУБФ. Разработанные факторизации 2D-неразделимых банков фильтров названы 16in-16out- и 64in-64out-системами, соответственно для 4- и 8-канального Q-ПУБФ прототипов.

Задача синтеза 2D NS Q-ПУБФ сформулирована, как задача поиска условного экстремума: найти коэффициенты 2D-банка фильтра, которые обеспечивают локальный минимум целевой функции $f(x) = \min(-CG_{2D}(x))$ при заданном ограничении ослабления в полосе пропуска каналов банка фильтров. Синтез 2D Q-ПУБФ, строится на основе численных методов. Коэффициент эффективности кодирования CG_{2D} для полученного, например, 4-канального неразделимого банка фильтров 2D NS Q-ПУБФ, и переход на новую итерацию алгоритма синтеза определяются в соответствии со следующей процедурой [22]:

- 1) генерируется матрица белого шума размером $N \times N$;
- 2) формируются входные данные $x_{n,n}$ на основе двумерной модели AR(1);
- 3) проводится обработка $x_{n,n}$ синтезированным 2D NS Q-ПУБФ;
- 4) вычисляется коэффициент эффективности кодирования неразделимого банка фильтров 2D NS Q-ПУБФ:

$$CG_{2D} = 10 \lg \frac{\frac{1}{M^2} \sum_{k=0}^{M^2-1} \sigma_{xk}^2}{\left(\prod_{k=0}^{M^2-1} \sigma_{xk}^2 \right)^{\frac{1}{M^2}}}, \quad (3)$$

где σ_{xk}^2 – дисперсия сигнала в k -м канале банка фильтров.

Получены амплитудно-частотные характеристики восьми диагональных каналов структуры 64in-64out 2D-неразделимого банка фильтров PMI LP Q-ПУБФ: $N = 2$, $CG_{2D} = 17,12$ дБ [23]. Коэффициенты 2D Q-ПУБФ с зеркальной симметрией частотных характеристик приведены в табл. 3.

Таблица 3. Коэффициенты 2D-неразделимого одномерного параунитарного банка фильтров ($N = 2$, $CG_{2D} = 17,12$ дБ)

Table 3. Coefficients of 2D inseparable one-dimensional paraunitary filter bank ($N = 2$, $CG_{2D} = 17.12$ dB)

P_0	P_1	P_2	Q_0	Q_1	Q_2
-0,2991	-0,9980	+0,3604	-0,0398	-0,0167	-0,5669
-0,8952i	+0,0607i	-0,3153i	+0,2928i	+0,9998i	+0,5818i
-0,2010j	+0,0071j	+0,4282j	-0,0691j	-0,0065j	-0,2653j
-0,2622k	-0,0129k	+0,7663k	+0,9528k	+0,0094k	-0,5404k

Полученная эффективность кодирования для 16in-16out 2D-банка фильтров превышает одномерный случай почти на 2 дБ ($CG = 8,23$ дБ в сравнении с $CG_{MD} = 11,55$ дБ). Таким обра-

зом, неразделимые двумерные структуры банка фильтров в алгебре кватернионов 16in-16out и 64in-64out, полученные в результате новой факторизации полифазных передаточных матриц соответствующих одномерных 4- и 8-канальных банков фильтров, позволяют проводить обработку непосредственно в двумерной области по схеме кодирования без потерь в ограничениях целочисленной арифметики. Сохранение свойств перфективной реконструкции дает возможность реализовать решетчатую структуру в виде сопроцессора для кристалла программируемой логической интегральной схемы для различного формата представления коэффициентов в числах с фиксированной точкой [23].

Генеративный метод спектрального оценивания на базе авторегрессионной модели анализируемого сигнала

В настоящее время большое количество задач обработки речевых сигналов решается с помощью машинного обучения. Обычно в таких системах используются признаки на основе спектральной оценки сигнала. При этом эффективность обработки прямо пропорционально связана со спектральным разрешением и точностью оценки. В большинстве применяемых сегодня методах вычисления признаков для машинного обучения в частотной области используется дискретное преобразование Фурье (ДПФ), которое подходит для частотно-временного анализа периодических или частично периодических сигналов. Спектральная оценка ДПФ предполагает использование временного окна, что приводит к значительным искажениям полученной спектральной оценки. Однако недостатки ДПФ можно в значительной степени нивелировать путем спектрального оценивания на базе авторегрессионной (АР) модели анализируемого сигнала, используя генеративный подход [24], который предполагает вычисление ДПФ последовательности отчетов, сгенерированной в соответствии с использованием АР-модели входного речевого сигнала. Сгенерированная последовательность учитывает периодичность преобразования, позволяя тем самым повысить точность спектральной оценки ДПФ. Примером успешного использования данного подхода в области машинного обучения является его применение для вычисления малоразмерных спектральных признаков на основе дискретной оценки огибающей спектра авторегрессионного процесса (ОСАП) в системах автоматического выявления патологий голосовой функции человека [25, 26].

Основная идея метода получения дискретной оценки ОСАП состоит в том, чтобы использовать ДПФ без применения оконной функции для обработки входных данных, обеспечив при этом согласованность периода анализируемого сигнала и размера ДПФ. С этой целью в качестве входных данных для ДПФ используется специальным образом сгенерированная последовательность, которая имеет равный с размером ДПФ период и соответствующий анализируемому сигналу спектральный состав. Признаки на основе ОСАП получаются в результате ДПФ сгенерированной последовательности малой длины. Процесс вычисления дискретной оценки ОСАП включает следующие три этапа [24].

1. По входному кадру анализируемого сигнала с использованием автокорреляционного метода определяются параметры АР-модели P -го порядка a_k , $k = \overline{1, P}$.

2. С помощью АР-модели генерируется последовательность с периодом, равным размеру ДПФ. Размер ДПФ зависит от целевой размерности вектора признаков и покрываемого им частотного диапазона. Пусть требуется сгенерировать ненулевую последовательность $x(m)$ длиной M

$$\sum_{m=1}^M |x(m)| \neq 0. \quad (4)$$

Используя параметры АР-модели анализируемого процесса, получаем следующее соотношение:

$$x(m) = -\sum_{k=1}^P a_k x(m-k), \quad m = \overline{1, M+P}. \quad (5)$$

Положив $x(m) = 0$ при $m = \overline{M+1, M+P}$ и $m = \overline{-P+1, 0}$, выражение (5) для $m = \overline{1, M}$ можно записать в виде:

$$\begin{pmatrix} -1 & 0 & 0 & \dots & 0 & 0 \\ -a_1 & -1 & 0 & \dots & 0 & 0 \\ -a_2 & a_1 & -1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & -a_p \end{pmatrix} \cdot \begin{pmatrix} x(1) \\ x(2) \\ x(3) \\ \vdots \\ x(M) \end{pmatrix} = 0. \quad (6)$$

Для выполнения условия (4) установим $x(M/2) = 1$. Решив систему линейных уравнений (3), найдем искомую последовательность $x(m)$, $m = 1, M$, причем для нее будет характерно затухание к краям кадра, что позволяет добиться того же эффекта, как и в случае применения оконной функции.

3. Вычисляется ДПФ одного периода последовательности, сгенерированной на этапе 2:

$$S_{est}(k) = \left| \sum_{m=0}^{M-1} x(m) e^{-i \frac{2\pi}{M} km} \right|, \quad k = 1, \frac{M}{2}. \quad (7)$$

Полученный дискретный спектр $S_{est}(k)$ используется в качестве вектора признаков для анализа голосового сигнала.

На рис. 8 проиллюстрированы основные этапы вычисления малоразмерных спектральных признаков на основе ОСАП [26]. На рис. 8, *a* показан один фрейм исходного сигнала, на рис. 8, *b* – вычисленные параметры АР-модели сигнала. В данном случае использовался порядок фильтра-предсказателя, равный 280. Модель достаточно точно описывает гармоническую структуру сигнала, что показано на рис. 8, *c*. На рис. 8, *d* представлена последовательность, которая получается в результате решения (3). Видно, что центральный отсчет равен единице, а по краям интервала наблюдения сигнал стремится к нулю, что дает при спектральном оценивании тот же эффект, что и применение оконной функции. На рис. 8, *e* показано вычисление ДПФ от сгенерированной последовательности. Полученные спектральные составляющие являются дискретной оценкой ОСАП и используются в качестве вектора признаков для анализа голоса.

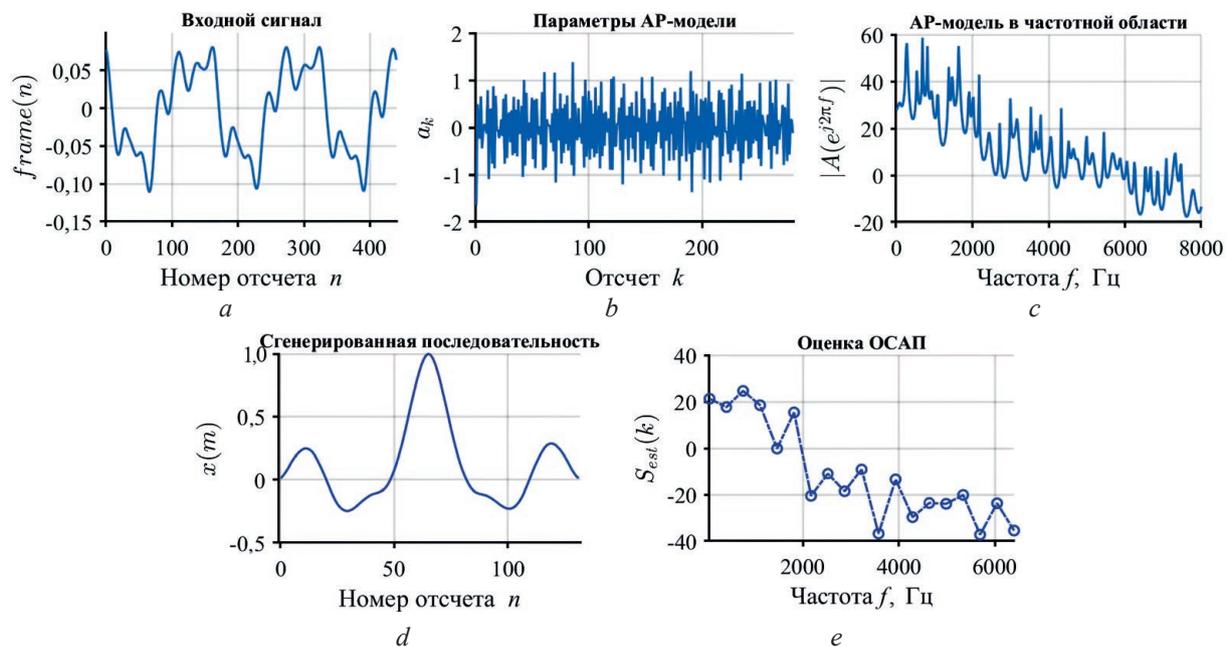


Рис. 8. Этапы получения оценки огибающей спектра авторегрессионного процесса:

a – один фрейм исходного сигнала; *b* – параметры АР-модели сигнала;

c – гармоническая структура сигнала; *d* – результат решения системы (6);

e – дискретное преобразование Фурье сгенерированной последовательности

Fig. 8. Stages of obtaining an estimate of the spectrum envelope of the autoregressive process:

a – one frame of the original signal; *b* – parameters of the AP signal model; *c* – harmonic structure of the signal;

d – the result of solving system (6); *e* – discrete Fourier transform of the generated sequence

В [26] проведен анализ эффективности метода вычисления дискретной оценки ОСАП путем сравнения точности и стабильности представления спектра тестового сигнала с помощью спектральных огибающих, полученных тремя способами на основе: ОСАП, ДПФ без применения окна, ДПФ с применением окна Ханна. Как показали результаты представленных в [26] экспериментов, спектральные огибающие, полученные с помощью метода на основе ОСАП, более точно, по сравнению с остальными методами, представляют спектр тестового сигнала с точки зрения минимальной квадратичной ошибки.

Результаты проведенных экспериментов по использованию признаков на основе ОСАП в системах машинного обучения говорят о том, что данный подход позволяет более эффективно извлекать полезную для классификации информацию из обучающих данных и ощутимо повышать точность классификации. Это особенно актуально в случаях, когда по каким-либо причинам доступен только небольшой объем обучающих данных. Применение сложных моделей классификации в таких случаях, например, на основе глубоких нейросетей, нецелесообразно вследствие явления переобучения. Повышение точности классификации в этих условиях возможно при использовании простых, но более адекватных речевому процессу моделей.

Заключение

В ближайшие годы ожидается стремительный рост объемов рынка технологий обработки изображений и речи. В этой связи актуальными являются научные исследования и разработки, а также подготовка инженеров и магистров по данным направлениям. Особое внимание необходимо уделить интеграции технологий цифровой обработки изображений, речевых сигналов и машинного обучения.

Список литературы

1. GII. Global Information [Электронный ресурс]. Режим доступа: <https://www.giiresearch.com/>. Дата доступа: 29.01.2024.
2. Fortune Business Insights Information [Электронный ресурс]. Режим доступа: <https://www.fortunebusinessinsights.com/>. Дата доступа: 29.01.2024.
3. Blum, H. A Transformation for Extracting New Descriptors of Shape / H. Blum. Cambridge: MIT Press, MA, 1967. P. 362–380.
4. Saha, P. K. A Survey on Skeletonization Algorithms and Their Applications / P. K. Saha, G. Borgefors, G. Sanniti di Baja // Pattern Recognition Letter. 2016. Vol. 76. P. 3–12. DOI: 10.1016/j.patrec.2015.04.006.
5. Ogniewicz, R. L. Hierarchic Voronoi Skeletons / R. L. Ogniewicz, O. Kübler // Pattern Recognition. 1995. Vol. 28. P. 343–359.
6. Leymarie, F. Simulating the Grassfire Transform Using an Active Contour Model / F. Leymarie, M. D. Levine // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1992. Vol. 14. P. 56–75.
7. Saha, P. K. Skeletonization: Theory, Methods, and Applications / P. K. Saha, G. Borgefors, G. Sanniti di Baja // Academic Press. London. 2017.
8. Bai, X. Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution / X. Bai, L. J. Latecki, W. Y. Liu // IEEE Trans. Pattern Anal. Mach. Intell. 2007. Vol. 29, No 3. P. 449–462. DOI: 10.1109/TPAMI.2007.59.
9. Skeleton Pruning as Trade-Off Between Skeleton Simplicity and Reconstruction Error / S. Wei [et al.] // Sci. China Inf. Sci. 2013. Vol. 56. P. 1–14. DOI: 10.1007/s11432-012-4715-3.
10. A Skeleton Pruning Algorithm Based on Information Fusion / H. Liu [et al.] // Pattern Recognit. Lett. 2013. Vol. 34, No 10. P. 1138–1145. DOI: 10.1016/j.patrec.2013.03.013.
11. A Skeleton Pruning Method Based on Saliency Sorting / G. Siyu [et al.] // 14th IEEE International Conference on Electronic Measurement & Instruments, Changsha, China. 2019. P. 593–599. DOI: 10.1109/ICEMI46757.2019.9101710.
12. Latecki, L. J. Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution / L. J. Latecki, R. Lakämper // Comput. Vis. Image Underst. 1999. Vol. 73, No 3. P. 441–454. DOI: 10.1006/cviu.1998.0738.
13. Latecki, L. J. Polygon Evolution by Vertex Deletion / L. J. Latecki, R. Lakämper // Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer, Berlin, Heidelberg. 1999. Vol. 1682. P. 398–409. DOI: 10.1007/3-540-48236-9_35.
14. Latecki, L. J. Shape Similarity Measure Based on Correspondence of Visual Parts / L. J. Latecki, R. Lakämper // IEEE Trans. Pattern Anal. Mach. Intell. 2000. Vol. 22, No 10. P. 1185–1190. DOI: 10.1109/34.879802.

15. Ma, J. A Novel Fast Iterative Parallel Thinning Algorithm / J. Ma, X. H. Ren, V. Yu. Tsviatkou // Proceedings of the 2020 4th International Conference on Vision, Image and Signal Processing, New York, NY, USA. 2020. Article 7. P. 1–5. DOI: 10.1145/3448823.3448836.
16. Choi, W. Extraction of the Euclidean Skeleton Based on a Connectivity Criterion / W. Choi, K. Lam, W. Siu // Pattern Recognition. 2003. Vol. 36. P. 721–729. DOI: 10.1016/S0031-3203(02)00098-5.
17. Novel Fully Parallel Skeletonization Algorithm / J. Ma [et al.] // Pattern Analysis and Applications. 2021. P. 1–20. DOI: 10.1007/s10044-021-01039-y.
18. Noise-Against Skeleton Extraction Framework and Application on Hand Gesture Recognition / J. Ma [et al.] // IEEE Access. 2023. Vol. 11. P. 9547–9559. DOI: 10.1109/ACCESS.2023.3240313.
19. A One-Pass Thinning Algorithm and Its Parallel Implementation / R. T. Chin [et al.] // Computer Vision, Graphics and Image Processing. Vol. 40. P. 30–40. DOI: 10.1016/0734-189X(87)90054-5.
20. Ma, J. Hand Gesture Recognition Based on Skeletal Image Properties / J. Ma, V. Y. Tsviatkou, A. A. Boriskevich // Open Semantic Technologies for Intelligent Systems: Collec. of Sci. Pap. Minsk, 2023. Iss. 7. P. 247–256.
21. Петровский, Н. А. Двумерные неразделимые банки фильтров на основе кватернионов / Н. А. Петровский, Е. В. Рыбенков, А. А. Петровский // Algorithms, Architectures, Arrangements, and Applications (SPA): тр. Междунар. науч.-техн. конф. Signal Processing, 2018. С. 120–125.
22. Рыбенков, Е. В. 2D неразделимая целочисленная реализация параунитарных банков фильтров на основе блочно-лестничной структуры / Е. В. Рыбенков, Н. А. Петровский // 27th European Signal Processing Conference (EUSIPCO): тр. Междунар. науч.-техн. конф., 2019. С. 1494–1499.
23. Рыбенков, Е. В. Проектирование неразделимых многомерных банков фильтров в алгебре кватернионов / Е. В. Рыбенков, Н. А. Петровский // Цифровая обработка сигналов и ее применение (DSPA'2019): тр. 21-й Междунар. конф., 2019., Т. 2. С. 368–373.
24. Малоразмерные спектральные признаки для машинного обучения в задачах анализа и классификации голосового сигнала / Д. С. Лихачёв [и др.] // Информатика. 2023. Т. 20, № 1. С. 102–112. DOI: 10.37661/1816-0301-2023-20-1-102-112.
25. Генеративный метод получения спектральных огибающих в контексте задачи анализа состояния голосовой функции человека / Д. С. Лихачёв [и др.] // Медэлектроника-2022. Средства медицинской электроники и новые медицинские технологии: сб. науч. ст. XI Междунар. науч.-тех. конф., г. Минск, 8–9 декабря 2022 г. Минск: Белор. гос. ун-т информ. и радиоэл., 2022. С. 281–283.
26. Комбинированный метод отбора информативных признаков для выявления речевых патологий по голосу / Д. С. Лихачёв [и др.] // Доклады БГУИР. 2023. Т. 21, № 4. С. 110–117. DOI: 10.35596/1729-7648-2023-21-4-110-117.

References

1. *GII. Global Information*. Available: <https://www.giiresearch.com/> (Accessed 29 January 2024).
2. *Fortune Business Insights Information*. Available: <https://www.fortunebusinessinsights.com/> (Accessed 29 January 2024).
3. Blum H. (1967). *A Transformation for Extracting New Descriptors of Shape*. Cambridge, MIT Press, MA. 362–380.
4. Saha P. K., Borgefors G., Sanniti di Baja G. (2016) A Survey on Skeletonization Algorithms and Their Applications. *Pattern Recognition Letter*. 76, 3–12. DOI: 10.1016/j.patrec.2015.04.006.
5. Ogniewicz R. L., Kübler O. (1995) Hierarchic Voronoi Skeletons. *Pattern Recognition*. 28, 343–359.
6. Leymarie F., Levine M. D. (1992) Simulating the Grassfire Transform Using an Active Contour Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 14, 56–75.
7. Saha P. K., Borgefors G., Sanniti di Baja G. (2017) *Skeletonization: Theory, Methods, and Applications*. Academic Press. London.
8. Bai X., Latecki L. J., Liu W. Y. (2007) Skeleton Pruning by Contour Partitioning with Discrete Curve Evolution. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (3), 449–462. DOI: 10.1109/TPAMI.2007.59.
9. Wei S., Xiang B. A. I., Xingwei Y., Jan L. L. (2013) Skeleton Pruning as Trade-Off Between Skeleton Simplicity and Reconstruction Error. *Sci. China Inf. Sci.* 56, 1–14. DOI: 10.1007/s11432-012-4715-3.
10. Liu H., Wu Z. H., Zhang X., Hsu D. F. (2013) A Skeleton Pruning Algorithm Based on Information Fusion. *Pattern Recognit. Lett.* 34 (10), 1138–1145. DOI: 10.1016/j.patrec.2013.03.013.
11. Siyu G., Pingping H., Zhigang L., He W., Min L. (2019) A Skeleton Pruning Method Based on Saliency Sorting. *14th IEEE International Conference on Electronic Measurement & Instruments*. Changsha, China. 593–599. DOI: 10.1109/ICEMI46757.2019.9101710.
12. Latecki L. J., Lakämper R. (1999) Convexity Rule for Shape Decomposition Based on Discrete Contour Evolution. *Comput. Vis. Image Underst.* 73 (3), 441–454. DOI: 10.1006/cviu.1998.0738.

13. Latecki L. J., Lakämper R. (1999) Polygon Evolution by Vertex Deletion. *Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, Springer, Berlin, Heidelberg. 1682, 398–409. DOI: 10.1007/3-540-48236-9_35.
14. Latecki L. J., Lakämper R. (1999) Shape Similarity Measure Based on Correspondence of Visual Parts. *IEEE Trans. Pattern Anal. Mach. Intell.* 22 (10), 1185–1190. DOI: 10.1109/34.879802.
15. Ma J., Ren X. H., Tsviatkou V. Yu. (2020) A Novel Fast Iterative Parallel Thinning Algorithm. *Proceedings of the 2020 4th International Conference on Vision, Image and Signal Processing, New York, NY, USA*. Article 7, 1–5. DOI: 10.1145/3448823.3448836.
16. Choi W., Lam K., Siu W. (2003) Extraction of the Euclidean Skeleton Based on a Connectivity Criterion. *Pattern Recognition*. 36, 721–729. DOI: 10.1016/S0031-3203(02)00098-5.
17. Ma J., Ren X., Tsviatkou V. Yu., Kanapelka V. K. (2021) A Novel Fully Parallel Skeletonization Algorithm. *Pattern Analysis and Applications*. 1–20. DOI: 10.1007/s10044-021-01039-y.
18. Ma J., Ren X., Li H., Li W., Tsviatkou V. Y., Boriskevich A. A. (2023) Noise-Against Skeleton Extraction Framework and Application on Hand Gesture Recognition. *IEEE Access*. 11, 9547–9559. DOI: 10.1109/ACCESS.2023.3240313.
19. Chin R. T., Wan H. K., Stover D., Iverson R. (1987) A One-Pass Thinning Algorithm and Its Parallel Implementation. *Computer Vision, Graphics and Image Processing*. 40, 30–40. DOI: 10.1016/0734-189X(87)90054-5.
20. Ma J., Tsviatkou V. Y., Boriskevich A. A. (2023) Hand Gesture Recognition Based on Skeletal Image Properties. *Open Semantic Technologies for Intelligent Systems: Collection of Scientific Papers*. (7), 247–256.
21. Petrovsky N. A., Rybenkov E. V., Petrovsky A. A. (2018) Two-Dimensional Non-Separable Quaternionic Paraunitary Filter Banks. *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*. 120–125.
22. Rybenkov E. V., Petrovsky N. A. (2019) 2D Non-Separable Integer Implementation of Paraunitary Filter Bank Based on the Quaternionic Multiplier Block-Lifting Structure. *27th European Signal Processing Conference (EUSIPCO)*. 1494–1499.
23. Rybenkov E. V., Petrovsky N. A. (2019) Design of Non-Separable Multidimensional Filter Banks in Quaternionic Algebra. *Digital Signal Processing and its Applications: Proc. of 21 International Conference (DSPA'2019)*. 2, 368–373.
24. Likhachov D. S., Vashkevich M. I., Petrovsky N. A., Azarov E. S. (2023) Small-Size Spectral Features for Machine Learning in Voice Signal Analysis and Classification Tasks. *Informatics*. 20 (1), 102–112. DOI: 10.37661/1816-0301-2023-20-1-102-112 (in Russian).
25. Likhachov D. S., Vashkevich M. I., Petrovsky N. A., Azarov I. S. (2022) Generative Method of Obtaining Spectral Envelopes for Speech Signal Analysis and Processing Tasks. *Medelectronics-2022. Means of Medical Electronics and New Medical Technologies: Collection of Scientific Articles XI Intern. Scientific and Technical Conference, Minsk, Dec. 8–9*. Minsk, Belarusian State University of Informatics and Radioelectronics. 281–283 (in Russian).
26. Likhachov D. S., Vashkevich M. I., Petrovsky N. A., Azarov E. S. (2023) Combined Method for Informative Feature Selection for Speech Pathology Detection. *Doklady BGUIR*. 21 (4), 110–117. DOI: 10.35596/1729-7648-2023-21-4-110-117 (in Russian).

Вклад авторов

Лихачёв Д. С. разработал генеративный метод спектрального оценивания речевых сигналов на базе авторегрессионной модели, провел экспериментальные исследования.

Ма Ц. разработал алгоритм скелетизации изображений, провел экспериментальные исследования по распознаванию жестов рук.

Петровский Н. А. разработал структуру параунитарных банков фильтров на основе алгебры кватернионов, алгоритм их синтеза, провел экспериментальные исследования.

Азаров И. С., Цветков В. Ю. определили направления и цели исследований, осуществили постановку задач и редактирование статьи, выполнили интерпретацию полученных результатов.

Authors' contribution

Likhachov D. S. developed a generative method for spectral estimation of speech signals based on an autoregressive model, conducted experimental studies.

Ma J. developed an image skeletonization algorithm and conducted experimental studies on hand gesture recognition.

Petrovsky N. A. developed the structure of paraunitary filter banks based on quaternion algebra, an algorithm for their synthesis, and conducted experimental studies.

Azarov E. S., Tsviatkou V. Yu. determined the directions and goals of the research, formulated the research problems and edited the article, and interpreted the results obtained.

Сведения об авторах

Лихачёв Д. С., канд. техн. наук, доц., доц. каф. электронных вычислительных средств, Белорусский государственный университет информатики и радиоэлектроники (БГУИР)

Ма Ц., ассист. каф. инфокоммуникационных технологий, БГУИР

Петровский Н. А., канд. техн. наук, доц., доц. каф. электронных вычислительных средств, БГУИР

Азаров И. С., д-р техн. наук, проф., зав. каф. электронных вычислительных средств, БГУИР

Цветков В. Ю., д-р техн. наук, проф., зав. каф. инфокоммуникационных технологий, БГУИР

Адрес для корреспонденции

220013, Республика Беларусь,
г. Минск, ул. П. Бровки, 6
Белорусский государственный университет
информатики и радиоэлектроники
Тел.: +375 17 293-84-08
E-mail: vtsvet@bsuir.by
Цветков Виктор Юрьевич

Information about the authors

Likhachov D. S., Cand. of Sci., Associate Professor, Associate Professor at Computer Engineering Department, Belarusian State University of Informatics and Radioelectronics (BSUIR)

Ma J., Assistant at the Department of Infocommunication Technologies, BSUIR

Petrovsky N. A., Cand. of Sci., Associate Professor, Associate Professor at Computer Engineering Department, BSUIR

Azarov I. S., Dr. of Sci. (Tech.), Associate Professor, Head of Computer Engineering Department, BSUIR

Tsviatkou V. Yu., Dr. of Sci. (Tech.), Professor, Head of the Department of Infocommunication Technologies, BSUIR

Address for correspondence

220013, Republic of Belarus,
Minsk, P. Brovki St., 6
Belarusian State University
of Informatics and Radioelectronics
Tel.: +375 17 293-84-08
E-mail: vtsvet@bsuir.by
Tsviatkou Viktor Yur'evich