

УДК 004.42:004.032.26:519.876.5
DOI



CC BY 4.0

РАЗРАБОТКА АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ ДЛЯ АНАЛИЗА ГОРОДСКОГО ШУМА С ИСПОЛЬЗОВАНИЕМ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ

Бакаева О. А., Буткина А. А., Ледяйкин Н. В.

*Федеральное государственное бюджетное образовательное учреждение высшего образования
«Национальный исследовательский Мордовский государственный университет
имени Н. П. Огарёва», Саранск, Российская Федерация, e-mail: helga_rm@rambler.ru*

Для обеспечения комфортных условий жизни населения и сохранения здоровья окружающей среды становится все более важной задача измерения шума. Традиционные методы зачастую малоэффективны и дорогостоящи, поэтому перспективным подходом является использование беспилотных летательных аппаратов (дронов) для записи городского шума и его последующего анализа. Цель проводимого исследования – разработать автоматизированную систему для обработки аудиоданных, полученных с использованием дрона, и анализа зашумленности городских территорий. В статье предложен алгоритм исследования городского шума, состоящий из шести этапов. Одной из ключевых особенностей данного алгоритма является обработка полученных с помощью дрона аудиозаписей математическими методами (кратковременное преобразование Фурье, спектральная фильтрация) и специально обученной нейронной сетью с архитектурой phase-aware. Таким образом решается основная проблема исследования – очистка аудиофайлов городского шума от звуковых помех, вызванных работой винтов дрона и вибрацией его корпуса. Полученный чистый аудиосигнал поступает и анализируется в автоматизированной системе, разработанной на языке Python с использованием библиотек Tkinter, SciPy, Librosa, SoundFile, PyTorch и др. Конечным результатом работы системы является сформированный pdf-отчет с вычисленными характеристиками шума (RMS, средний и максимальный уровень шума и другими) и его разделением на составляющие. К тому же указано соответствие полученного уровня нормативным значениям.

Ключевые слова: автоматизированная система, шум, беспилотный летательный аппарат, преобразование Фурье, нейронная сеть, Python

DEVELOPMENT OF AN AUTOMATED SYSTEM FOR ANALYZING URBAN NOISE USING UNMANNED AERIAL VEHICLES

Bakaeva O. A., Butkina A. A., Ledyaykin N. V.

*Federal State Budgetary Educational Institution of Higher Education
“National Research Mordovia State University named after N. P. Ogarev”,
Saransk, Russian Federation, e-mail: helga_rm@rambler.ru*

To ensure comfortable living conditions for the population and preserve a healthy environment, noise measurement is becoming increasingly important. Traditional methods are often ineffective and expensive, so a promising approach is the use of unmanned aerial vehicles (drones) to record urban noise and subsequently analyze it. The objective of this study is to develop an automated system for processing drone audio data and analyzing urban noise levels. This article proposes a six-stage algorithm for studying urban noise. One of the key features of this algorithm is the processing of drone audio recordings using mathematical methods (short-time Fourier transform, spectral filtering) and a specially trained neural network with a phase-aware architecture. This solves the main research problem: removing noise interference from urban audio files caused by drone screws and drone body vibration during flight. The resulting clean audio signal is received and analyzed in an automated system developed in Python using Tkinter, SciPy, Librosa, SoundFile, PyTorch, and other libraries. The final output of the system is a generated pdf report with calculated noise characteristics (RMS, average and maximum noise level, and others) and its separation into components. The compliance of the obtained level with standard values is also indicated.

Keywords: automated system, noise, unmanned aerial vehicle, Fourier transform, neural network, Python

Введение

Воздействие шума на организм человека может оказывать негативное влияние, однако жители города непосредственно это могут не ощущать [1, 2]. Обычно для измерения шума в городах применяются стационарные или мобильные наземные средства его измерения, однако они являются дорогостоящими и не позволяют получить полную, объективную картину по зашумленности территории. Также их недостатком

является многоступенчатость процесса получения, измерения и анализа шума, что делает данный процесс длительным.

Для решения описанных проблем могут быть использованы беспилотные летательные аппараты (БПЛА). С их помощью сбор аудиоданных становится более мобильным и быстрым, предоставляя возможность охватывать разнообразные участки города, в том числе труднодоступные, в кратчайшие сроки. В [3] приведено начало авторского

исследования по оценке зашумленности городских территорий с помощью БПЛА, в котором также рассматривалась следующая проблема: наличие шума винтов дрона в аудиозаписи. Этот шум может сильно искажать полезную информацию, из-за чего невозможно проводить объективный анализ зашумленности территории. Поэтому предварительно необходимо очистить полученные аудиоданные от шума винтов и вибрации самого БПЛА [4]. В данной статье приводятся усовершенствованные методики отделения шума городской среды от шума дрона с использованием математических методов и современных нейросетевых технологий. После очистки звуковая дорожка должна быть проанализирована с точки зрения уровня зашумленности территории, выделения типов шумов [5], их составляющих и предоставления информации в виде отчетов с различными характеристиками. Для автоматизации данных процессов необходима разработка специального программного обеспечения, что будет являться результатом исследования.

Цель исследования – разработать автоматизированную систему для обработки аудиоданных, полученных с использованием БПЛА, и анализа зашумленности городских территорий.

Для достижения поставленной цели был разработан следующий алгоритм:

- 1) получение аудиофайла по БПЛА (шум улицы + шум от дрона);
- 2) обработка аудиофайла с помощью математических методов (шум улицы – грубая очистка);
- 3) обработка аудиофайла с помощью ИИ (шум улицы – тонкая очистка);
- 4) анализ «очищенного» шума: получение характеристик зашумленности;
- 5) формирование отчета о зашумленности;
- 6) формирование шумовой карты.

Материалы и методы исследования

Для обработки аудиосигналов использовались математические методы «грубой очистки» – кратковременное преобразование Фурье и спектральная фильтрация, а также «тонкая очистка» с помощью нейросети, обученной авторами.

Аудиозапись, полученная с использованием БПЛА, содержит в себе несколько компонент:

$$y[n] = s[n] + d[n] + v[n], \quad (1)$$

где $s[n]$ – городской шум, $d[n]$ – шум дрона, $v[n]$ – фоновый шум окружающей среды.

Для последующей обработки запись переводится в спектральную область с по-

мощью кратковременного преобразования Фурье [6]:

$$Y(t, f) = \sum_n y[n] \times w[n - tR] \times e^{-\frac{2i\pi fn}{N}}, \quad (2)$$

где $w[n]$ – окно [7], R – шаг окна, t – индекс временного кадра, f – частота.

Спектральная область используется, потому что в ней задача разделения упрощается, поскольку шум дрона имеет отчетливую гармоническую структуру, состоящую из интенсивных узкополосных пиков.

Гармоническая модель шума дрона будет выглядеть следующим образом:

$$D_{\text{harm}}(t, f) = \sum_{k=1}^K A_k(t) \times \delta(f - f_k(t)), \quad (3)$$

где $D_{\text{harm}}(t, f)$ – спектр гармонического шума дрона, $A_k(t)$ – амплитуда k -й гармоники, $f_k(t)$ – частота k -й гармоники, $\delta(f - f_k(t))$ – дельта-функция, моделирующая узкий спектральный пик, K – количество гармоник, которые необходимо подавить.

Формула (3) позволяет изолировать доминирующий узкополосный шум дрона (частоты пропеллеров) таким образом, чтобы обработка не внесла изменений в аудиозапись городского шума.

Следующий шаг – убрать широкополосный шум [8], который обычно связан с турбулентностью воздуха и вибрациями корпуса БПЛА. Для него оценивают спектральную плотность мощности [9, 10]:

$$\hat{\Phi}_{NN}(t, f) = E \left\{ \left| D_{\text{broad}}(t, f) \right|^2 \right\}. \quad (4)$$

После этого формируется спектральная маска [11]:

$$G(t, f) = \max \left(\frac{|Y(t, f)| - \hat{\Phi}_{NN}(t, f)}{|Y(t, f)|}, 0 \right), \quad (5)$$

где $|Y(t, f)|$ – амплитуда спектра текущего кадра.

Если применить такую маску, широкополосная составляющая заметно уменьшается, а полезный сигнал остается:

$$Y_1(t, f) = G(t, f) \times Y(t, f). \quad (6)$$

Чтобы обработка была качественнее, дальше используем адаптивный фильтр Винера. Он аккуратно подавляет шум, не влияя на сигнал:

$$G_w(t, f) = \frac{\Phi_{SS}(t, f)}{\Phi_{SS}(t, f) + \Phi_{NN}(t, f)}, \quad (7)$$

где $\Phi_{SS}(t, f)$ – оценка мощности полезной части сигнала.

После этого остаточный шум становится меньше, и на выходе получается уже более чистый спектр:

$$Y_2(t, f) = G_w(t, f) \times Y_1(t, f). \quad (8)$$

Второй этап очищения аудиофайла – использование нейросетевой модели. Благодаря архитектуре phase-aware [12], аудиосигнал рассматривается как двумерная структура, в которой учитываются амплитуда и фазовые особенности. Поэтому нейронная сеть лучше восстанавливает детали, которые могли потеряться на предыдущих этапах обработки.

Нейросеть работает не с исходной волной, а с комплексным спектрограммным представлением [13]. На вход подаются

две компоненты спектра – действительная и мнимая части

$$X(t, f) = \text{Re}(X) + i\text{Im}(X),$$

а на выходе модель оценивает комплексную маску, которая доочищает сигнал $M(t, f)$.

Модель не восстанавливает сигнал «с нуля», а дорабатывает уже очищенный спектр, поэтому лучше сохраняет полезные детали и меньше искажает речь/фон:

$$\hat{X}(t, f) = X(t, f) \times M(t, f). \quad (9)$$

Обучение нейросети идет на парных данных: входное аудио после математической очистки и соответствующее ему аудио, очищенное от шума дрона.

Качество обучения контролируется составной функцией потерь, которая учитывает сразу несколько аспектов:

$$L = L_{RI} + 0.5 \times L_{mag} + 0.1 \times L_{phase} + 0.05 \times L_{energy}, \quad (10)$$

где $L_{RI} = \|\hat{X} - X\|$ – ошибка по действительной и мнимой частям, $L_{mag} = \left\| |\hat{X}| - |X| \right\|$ – ошибка по модулю спектра, $L_{phase} = 1 - \cos(\hat{X}, X)$ – штраф за фазовое расхождение, L_{energy} – ошибка по сглаженной энергии кадров.

После удаления шума БПЛА уровень очищенного сигнала удобно оценивать через среднеквадратичное значение (RMS) [14, 15], которое считается по формуле (11) и передает его воспринимаемую громкость:

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}. \quad (11)$$

Для удобства интерпретации этот абсолютный показатель переводят в логарифмический масштаб – в децибелы:

$$L_{dB} = 10 \log_{10}(RMS). \quad (12)$$

В процессе исследования для разработки автоматизированной системы анализа шума были использованы следующие библиотеки языка программирования Python, выбранного:

Tkinter – предоставляет базовые средства для создания интерфейса системы;

CustomTkinter – дает доступ к современному оформлению элементов интерфейса;

NumPy – реализует векторные вычисления, расчет спектральных характеристик, массивы;

SciPy – осуществляет фильтрацию, поиск пиков, билинейное преобразование фильтров;

Librosa – загружает аудиоданные, использует кратковременное преобразование Фурье (STFT) и обратное кратковременное преобразование Фурье (STFT);

SoundFile – сохраняет очищенные аудиофайлы в формате WAV;

PyTorch – реализует и запускает нейросеть, построенную на основе архитектуры Phase-Aware;

Matplotlib – осуществляет построение спектров, спектрограмм и карт зашумленности;

Pillow – загружает и отображает графические изображения в интерфейсе;

ReportLab – формирует pdf-отчеты.

Результаты исследования и их обсуждение

В результате проведенного исследования была разработана автоматизированная система, интерфейс главной страницы которой приведен на рис. 1.

Рабочий процесс обработки аудиоданных БПЛА включает три этапа: обработку аудиофайлов, их анализ и формирование отчета. После обработки аудиофайла можно просмотреть, какие этапы он прошел, нажав на кнопку «История обработки» (рис. 2).

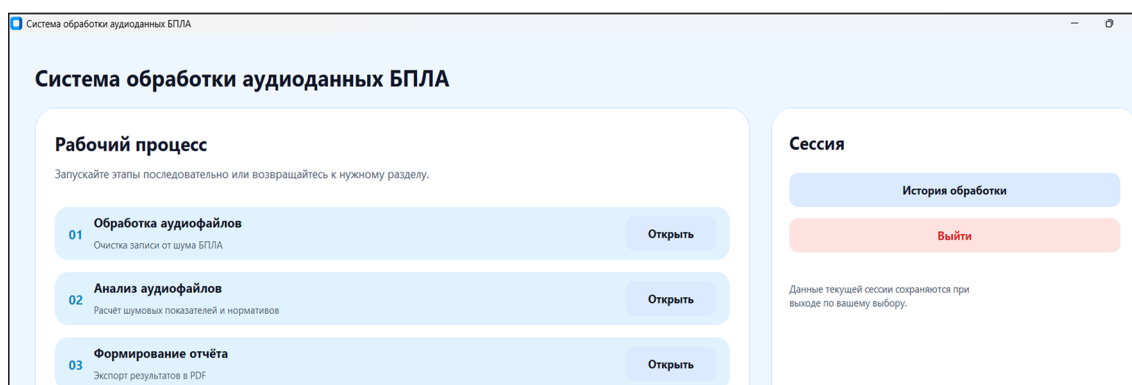


Рис. 1. Интерфейс главной страницы системы
Примечание: составлен авторами по результатам данного исследования

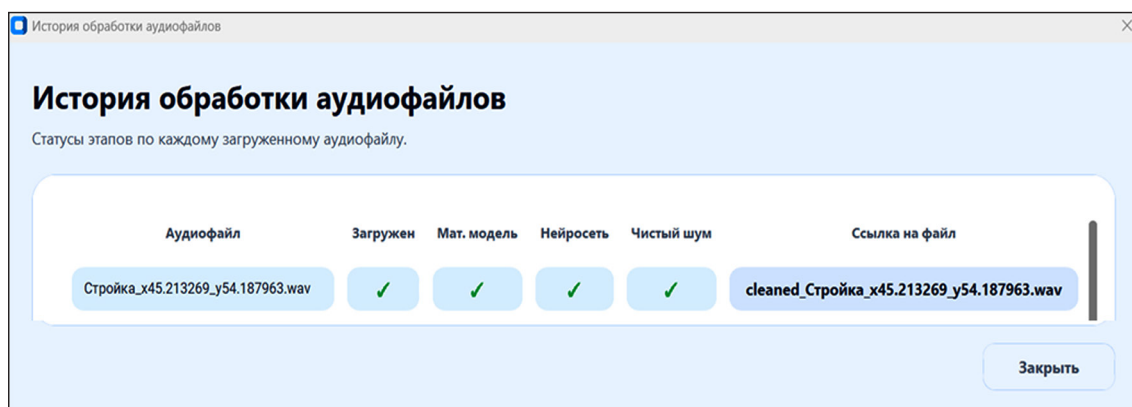


Рис. 2. Интерфейс окна «История обработки аудиофайлов»
Примечание: составлен авторами по результатам данного исследования

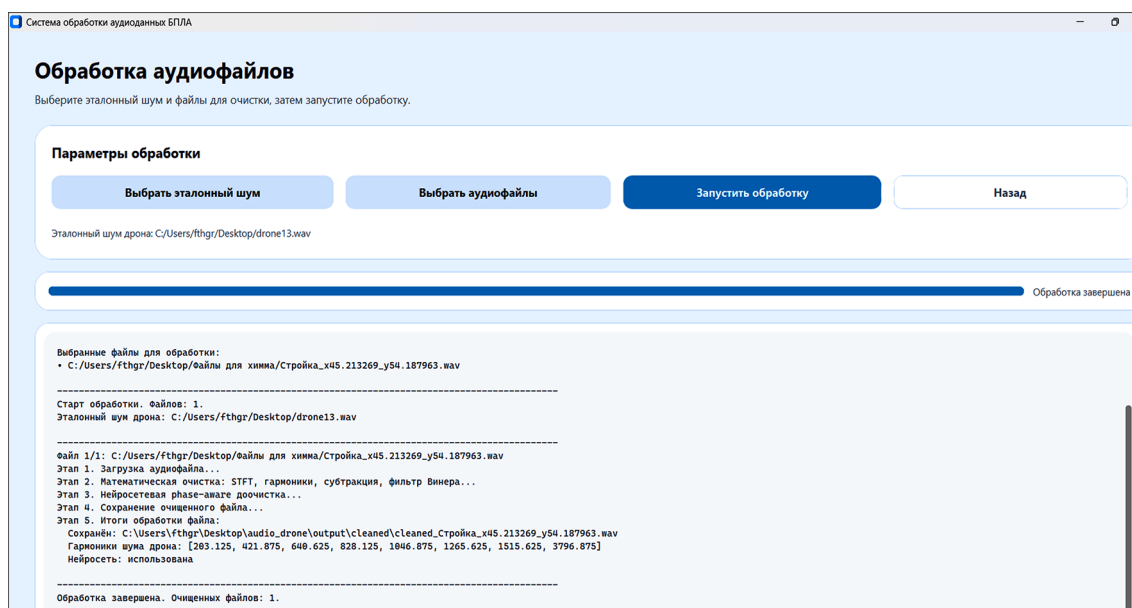


Рис. 3. Этап обработки аудиофайла
Примечание: составлен авторами по результатам данного исследования

Анализ файла 1/1: C:/Users/fthgr/Desktop/audio_drone/output/cleaned/cleaned_Стройка_х45.213269_y54.187963.wav
 Этап 1. Загрузка очищенного аудиофайла...
 Этап 2. Расчёт RMS, LAeq, LAFmax, LCpeak и частотных составляющих...
 Этап 3. Сохранение графиков спектра, спектрограммы и JSON-анализа...
 Файл: cleaned_Стройка_х45.213269_y54.187963.wav
 RMS: 0.051326
 LAeq: 66.78 дБА
 LAFmax: 81.44 дБА
 LCpeak: 95.78 дБ
 Доминирующая частота: 776.26 Гц
 Норматив: 55.0–70.0 дБА
 Соответствие: да
 Превышение верхней границы нормы не выявлено.
 Составляющие шумов:
 • Низкие частоты: 0.0%
 • Средние частоты: 65.5%
 • Высокие частоты: 34.5%
 Доминирующая составляющая: среднечастотная
 Комментарий: Преобладает среднечастотная составляющая: возможны речь, дорожная активность, сигналы и основной городской фон.

Рис. 4. Результат анализа аудиофайла
 Примечание: составлен авторами по результатам данного исследования

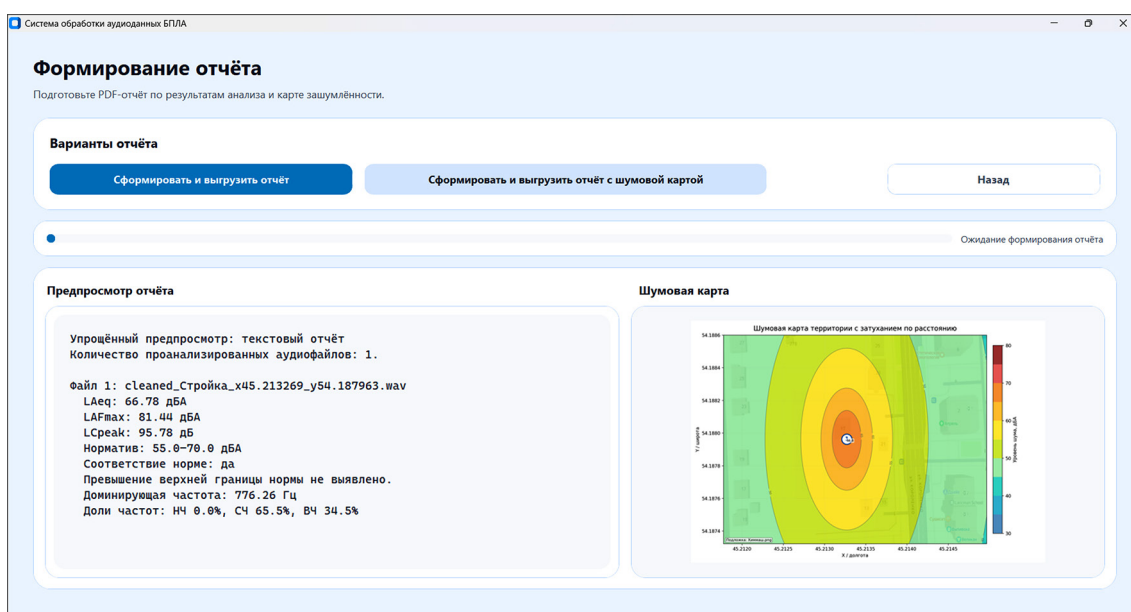


Рис. 5. Результат этапа «Формирование отчета»
 Примечание: составлен авторами по результатам данного исследования

На рис. 3 представлен результат обработки файла, включающий пять этапов:

- 1) загрузка аудиофайла;
- 2) математическая очистка: STFT, гармоника, субтракция, фильтр Винера;
- 3) нейросетевая «тонкая» очистка с помощью архитектуры phase-aware;
- 4) сохранение очищенного файла в выбранную директорию;
- 5) отображение итогов обработки аудиофайла.

Второй этап «Анализ аудиофайлов» (рис. 4) отображает следующие полученные характеристики уровня звука:

– RMS – средняя мощность цифрового аудиосигнала (если RMS ближе к 1, запись очень громкая);

– LAeq – средний уровень шума за все время записи;

– LAFmax – максимальный уровень шума;

– LCpeak – максимальный мгновенный пик шума;

– доминирующая частота,

– норматив и соответствие того, что показатель находится в норме;

– превышение – показатель того, насколько шум выше допустимого уровня;

– составляющие шумов – распределение энергии аудиосигнала по частотным диапазонам и др.

Последним этапом является формирование отчета (рис. 5), который выгружается в pdf-формате.

Таким образом разработанная система выполняет полный цикл обработки аудио-файлов: от загрузки до формирования отчета с различными характеристиками шума.

Заключение

Разработанное ПО успешно прошло тестирование на реальных аудиозаписях городской среды и показало качественные результаты по «очистке» шума с помощью математических методов и использования обученной авторами нейронной сети.

Результаты анализа шума представлены в виде pdf-отчета, включающего средние и максимальные характеристики уровня шума, а также его разделение на типы шумов, что имеет важное практическое значение при экологическом мониторинге города и составлении его шумовой карты.

Список литературы

1. Тимофеева Н. В., Ключниченко В. Н. Картографирование шумового загрязнения урбанизированных территорий как инструмент формирования комфортной среды обитания человека // Вестник Балтийского федерального университета им. И. Канта. Сер.: Естественные и медицинские науки. 2024. № 2. С. 85–100. URL: <https://cyberleninka.ru/article/n/kartografirovanie-shumovogo-zagryazneniya-urbanizirovannyh-territoriy-kak-instrument-formirovaniya-komfortnoy-sredy-obitaniya/viewer> (дата обращения: 09.06.2026). DOI: 10.5922/vestniknat-2024-2-6.
2. Ручкина Е. М., Шаталова К. С. Шумовое загрязнение: влияние шума на организм человека // Вестник Калужского университета. 2023. № 4 (61). С. 105–108. URL: <https://www.elibrary.ru/item.asp?id=62778924> (дата обращения: 09.06.2026).
3. Бакаева О. А., Буткина А. А., Ледайкин Н. В., Серегина А. А. Разработка программного обеспечения для оценки зашумленности городских территорий с использованием беспилотных летательных аппаратов // Современные наукоемкие технологии. 2025. № 8. С. 8–13. URL: <https://top-technologies.ru/ru/article/view?id=40458> (дата обращения: 09.06.2026). DOI: 10.17513/snt.40458.
4. Гончаренко Б. И., Кузьменков В. Ю., Котов А. Н. Экспериментальное исследование особенностей формирования спектра шумов беспилотного летательного аппарата // Noise Theory and Practice. 2020. № 4 (22). С. 49–59. URL: <https://cyberleninka.ru/article/n/eksperimentalnoe-issledovanie-osobennostey-formirovaniya-spektra-shumov-bespilotnogo-letatel'nogo-apparata> (дата обращения: 09.06.2026).
5. Ледайкин Н. В., Серегина А. А., Бакаева О. А. Исследование зашумленности городской среды с использованием беспилотных летательных аппаратов // Юность и знания – гарантия успеха – 2025: сборник научных статей 12-й Международной молодежной научной конференции (г. Курск, 25–26 сентября 2025 г.). Курск: ЗАО «Университетская книга», 2025. Т. 3. С. 61–65. EDN: FVBPWH.
6. Романовская Ю. М., Ильюшин Е. А. Методы аугментации аудиосигнала // International Journal of Open Information Technologies. 2023. № 3. С. 12–18. URL: <https://cyberleninka.ru/article/n/metody-augmentatsii-audio-signala> (дата обращения: 09.06.2026).
7. Меерсон В. Э., Хорольский О. С. Применение весовых окон для оценки спектральной плотности // Современные аспекты моделирования систем и процессов: материалы Всероссийской научно-практической конференции (г. Воронеж, 8 февраля 2021 г.). Воронеж: ВГЛУ им. Г. Ф. Морозова, 2021. С. 71–77. DOI: 10.34220/MAMSP_71-77. EDN: AWMQXT.
8. Хайло Н. С. Робастное обнаружение и оценивание широкополосных сигналов // Доклады Академии наук высшей школы Российской Федерации. 2024. № 1 (62). С. 90–99. URL: <https://elibrary.ru/item.asp?id=65623631> (дата обращения: 09.06.2026). DOI: 10.17212/1727-2769-2024-1-90-99.
9. Ефимов Е. Н., Шевгунов Т. Я., Кузнецов Ю. В. Оценка циклической спектральной плотности мощности инфокоммуникационных сигналов // Труды МАИ. 2017. № 97. URL: <https://trudymai.ru/published.php?ID=87294> (дата обращения: 09.06.2026).
10. Якимов В. Н., Машков А. В. Вычислительно-эффективное математическое обеспечение для оценки спектральной плотности мощности периодограммным и коррелограммным методами // Цифровая обработка сигналов и ее применение DSPA – 2024: доклады XXVI Международной конференции (г. Москва, 27–29 марта 2024 г.). М.: РНТОРЭС имени А. С. Попова, 2024. С. 54–58. URL: <https://elibrary.ru/item.asp?id=69216619> (дата обращения: 09.06.2026).
11. Хорев А. А., Дворянкин С. В., Козлачков С. Б., Василевская Н. В. Анализ предельных возможностей методов шумопонижения и реконструкции речевых сигналов, маскируемых различными типами помех // Вопросы кибербезопасности. 2024. № 1 (59). URL: <https://elibrary.ru/item.asp?id=59932757> (дата обращения: 09.06.2026). DOI: 10.21681/2311-3456-2024-1-89-100.
12. Peer T. and Gerkmann T. Phase-aware deep speech enhancement: It's all about the frame length // JASA Express Lett. 2022. Vol. 2. Is. 10. № 104802. URL: <https://pubs.aip.org/asa/jel/article/2/10/104802/2845278/Phase-aware-deep-speech-enhancement-It-s-all-about> (дата обращения: 09.06.2026). DOI: 10.1121/10.0014875.
13. Майер Д. В., Стаценко Л. Г. Оптимизация параметров предобработки аудиосигналов для нейросетевой классификации шумов // Noise Theory and Practice. 2025. Т. 11. № 4. С. 184–198. URL: <https://elibrary.ru/item.asp?id=89063099> (дата обращения: 09.06.2026).
14. Яковлев А. А., Кузин Д. А., Свердлов Е. Классификация звуков окружающей среды с использованием моделей машинного обучения // Журнал СФУ. Техника и технологии. 2025. Т. 18. № 5. С. 688–707. URL: <https://elibrary.ru/item.asp?id=82749929> (дата обращения: 09.06.2026).
15. Белов С. В., Катунин С. С., Кучин И. Ю., Руденко М. Ф. Тестирование эффективности гибридной методики шумоподавления в речевом сигнале для системы видеоконференций // Вестник АГТУ. Серия: Управление, вычислительная техника и информатика. 2025. № 1. С. 93–102. URL: <https://www.elibrary.ru/item.asp?id=81024289> (дата обращения: 11.06.2026). DOI: 10.24143/2072-9502-2025-1-93-102.

Конфликт интересов: Авторы заявляют об отсутствии конфликта интересов.

Conflict of interest: The authors declare that there is no conflict of interest.

Финансирование: Авторы заявляют об отсутствии внешнего финансирования.

Financing: The research was performed without external funding.