

Сравнительный спектральный анализ современной популярной и классической музыки с применением методов цифровой обработки сигналов

Гиренко Александр Евгеньевич

Приамурский государственный университет им.Шолом-Алейхема

Студент

Аннотация

В статье представлен сравнительный спектральный анализ современной популярной и классической музыкальных композиций с использованием методов цифровой обработки сигналов (ЦОС). С применением библиотеки Librosa в среде Python выполнены вычисление кратковременного преобразования Фурье (STFT), извлечение мел-кепстральных коэффициентов (MFCC) и расчет спектральных дескрипторов для двух репрезентативных треков: произведения В.А. Моцарта в исполнении М.Ж. Пирес и современной композиции. Проведен количественный анализ ключевых акустических метрик: спектрального центроида, RMS-энергии, коэффициента пересечения нуля (ZCR) и динамического диапазона. Результаты демонстрируют статистически значимые различия в тембральной структуре: спектральный центроид современной музыки превышает аналогичный показатель классической композиции в 3,3 раза (3268 Гц против 984 Гц), что подтверждает гипотезу о влиянии феномена «loudness war» на звукоинженерную практику XXI века.

Ключевые слова: цифровая обработка сигналов, спектральный анализ, Librosa, MFCC, спектральный центроид, сравнительный анализ музыки, loudness war, тембральная структура.

Comparative spectral analysis of modern popular and classical music using digital signal processing methods

Girenko Aleksandr Evgen'evich

Sholom-Aleichem Priamursky State University

Student

Abstract

The article presents a comparative spectral analysis of modern popular and classical musical compositions using digital signal processing (DSP) methods. Using the Librosa library in the Python environment, short-time Fourier transform (STFT) computation, mel-frequency cepstral coefficients (MFCC) extraction, and spectral descriptors calculation were performed for two representative tracks: a composition by W.A. Mozart performed by M.J. Pires and a modern track. A quantitative analysis of key acoustic metrics was conducted: spectral centroid, RMS energy,

zero-crossing rate (ZCR), and dynamic range. The results demonstrate statistically significant differences in timbral structure: the spectral centroid of modern music exceeds that of the classical composition by 3.3 times (3268 Hz vs 984 Hz), which confirms the hypothesis about the influence of the "loudness war" phenomenon on 21st century sound engineering practice.

Keywords: digital signal processing, spectral analysis, Librosa, MFCC, spectral centroid, comparative music analysis, loudness war, timbral structure.

1 Введение

1.1 Актуальность

Развитие технологий цифровой обработки сигналов (ЦОС) открывает новые возможности для объективного количественного анализа музыкальных произведений. Если традиционное музыковедение опирается преимущественно на субъективные экспертные оценки и качественный анализ, то современные алгоритмы позволяют извлекать из аудиосигнала строгие числовые метрики, характеризующие тембральную структуру, динамический диапазон и гармоническое содержание композиции.

Особый научный интерес представляет сравнительный анализ современных популярных и классических музыкальных произведений. В звукоинженерной литературе активно обсуждается феномен «loudness war» («война за громкость») — тенденция к максимальному сжатию динамического диапазона в коммерческих записях последних десятилетий с целью достижения конкурентного преимущества при ротации на радио и в стриминговых сервисах. Данный процесс, по мнению ряда исследователей, привел к фундаментальным изменениям в спектральных характеристиках популярной музыки, которые могут быть объективно зафиксированы методами ЦОС.

Актуальность исследования обусловлена необходимостью демонстрации возможностей современных программных библиотек (в частности, Librosa) для решения прикладных задач музыкальной акустики и проверки гипотез о качественных различиях между музыкальными эпохами на строгой математической основе.

1.2 Обзор исследований

В монографии А.Л. Михайловой и С.Ю. Щур рассматриваются возможности применения нейросетей в медиадизайне, анализируются современные подходы к использованию искусственного интеллекта в творческих индустриях [1]. Д.Ш.А. Дашкин исследует методы автоматического определения депрессии по голосу с использованием машинного обучения, описывая применение акустических признаков для классификации эмоциональных состояний [2]. Л. Виллемоес, Я. Клейса и П. Хеделин в патенте представляют звуковое кодирующее и декодирующее устройство, описывая методы обработки аудиосигналов с использованием спектрального анализа [3]. П.А. Садовников, А.Ю. Дроздов и Ю.Н. Фонин разрабатывают библиотеку параллельных функций цифровой обработки

сигналов для высокопроизводительных процессоров, описывая реализацию на языке C [4]. Ю. Нисидзава, С. Цудзи и К. Като в патенте предлагают устройство и способ спектрального анализа для производства и контроля качества стальной полосы, описывая применение оптических методов [5]. Ч. Цао проводит сравнительный анализ гармонических структур в классической и современной музыке, исследуя эволюцию музыкального языка и эстетики [6].

1.3 Цель исследования

Целью исследования является проведение сравнительного спектрального анализа современной популярной и классической музыкальных композиций с применением методов цифровой обработки сигналов и выявление количественных различий в их акустических характеристиках.

2 Материалы и методы

2.1 Описание набора данных

Для проведения исследования выбраны две репрезентативные музыкальные композиции:

1. Классическая музыка: В.А. Моцарт «Рондо Аллегretto» (К. 511) в исполнении Марии Жо Пирес, академическая запись, характеризующаяся минимальной постобработкой и сохранением естественной динамики фортепиано.
2. Современная популярная музыка: коммерческий трек, выпущенный в 2020-х годах, прошедший профессиональный мастеринг с применением компрессии и лимитирования.

Длительность классической композиции составляет 95 секунд, современной — 170 секунд. Формат файлов — MP3.

2.2 Инструментарий разработки

Для реализации исследования использовались следующие программные средства:

1. Язык программирования Python 3.10 — основная среда разработки.
2. Библиотека Librosa 0.10.0 — специализированный инструментарий для анализа музыки и аудио.
3. Библиотека NumPy 1.24.0 — для численных вычислений.
4. Библиотека Matplotlib 3.7.0 — для научной визуализации.

Выбор библиотеки Librosa обусловлен ее статусом де-факто стандарта в задачах анализа музыки: она предоставляет готовые реализации всех необходимых алгоритмов ЦОС и оптимизирована для работы с большими аудиофайлами.

2.3 Методология анализа

2.3.1 Кратковременное преобразование Фурье (STFT)

Для перехода из временной области в частотно-временную применено STFT с размером окна 2048 отсчетов и шагом 512. Это позволяет получить

двумерное представление сигнала, где по осям отложены время и частота, а интенсивность цвета соответствует амплитуде в децибелах.

2.3.2 Мел-кепстральные коэффициенты (MFCC)

MFCC аппроксимируют нелинейное восприятие частот человеческим слухом через мел-шкалу. Для каждого трека рассчитано 13 коэффициентов, которые формируют тембральный «отпечаток» композиции.

2.3.3 Спектральный центроид

Спектральный центроид характеризует «яркость» звука и вычисляется как взвешенное среднее частотного спектра:

$$SC = \frac{\sum_n f(n) \cdot x(n)}{\sum_n x(n)}$$

где $f(n)$ — частота n -го бина, $x(n)$ — ее амплитуда. Высокие значения центроида соответствуют «светлому», «яркому» звуку с преобладанием высоких частот.

2.3.4 Динамический диапазон

Динамический диапазон вычисляется как отношение максимальной амплитуды к 5-му перцентилю абсолютных значений сигнала, выраженное в децибелах:

$$DR = 20 \cdot \log_{10} \left(\frac{A_{max}}{A_{5\%}} \right)$$

Данный показатель характеризует степень компрессии аудиоматериала.

3 Результаты и обсуждение

3.1 Реализация программного модуля

Разработанный программный модуль реализует последовательность следующих этапов:

1. Интерактивный выбор аудиофайла через стандартный диалог операционной системы.
2. Загрузка и предобработка (конвертация в моно, нормализация частоты дискретизации).
3. Вычисление признаков: STFT, MFCC, спектрального центроида.
4. Построение индивидуальных графиков спектрального анализа (сетка 2×2).
5. Расчет и вывод ключевых метрик в консоль.

Фрагмент кода, реализующий вычисление спектрального центроида и построение спектрограммы, представлен на рисунке 1.

```

# 1. Спектральный центроид (яркость звука)
spectral_centroid = librosa.feature.spectral_centroid(y=y, sr=sr)[0]
mean_centroid = np.mean(spectral_centroid)
# 2. RMS энергия (воспринимаемая громкость)
rms = librosa.feature.rms(y=y)[0]
mean_rms = np.mean(rms)
# 3. Zero Crossing Rate (шумовость / наличие перкуссии)
zcr = librosa.feature.zero_crossing_rate(y=y)[0]
mean_zcr = np.mean(zcr)
# 4. Спектральный контраст (разница между пиками и впадинами спектра)
contrast = librosa.feature.spectral_contrast(y=y, sr=sr)
mean_contrast = np.mean(contrast, axis=1)
# 5. Динамический диапазон (дБ)
max_amp = np.max(np.abs(y))
percentile_5 = np.percentile(np.abs(y), q=5)
dynamic_range_db = 20 * np.log10(max_amp / (percentile_5 + 1e-10))
# Построение графиков
fig, axes = plt.subplots(nrows=2, ncols=2, figsize=(14, 10))
fig.suptitle(t=f'Спектральный анализ: {track_name}', fontsize=16, fontweight='bold')
# График 1: Waveform (Временная область)
axes[0, 0].plot(np.linspace(start=0, duration, len(y)), y, color='#3498db', linewidth=0.3, alpha=0.7)
axes[0, 0].set_title('Временная область (Waveform)')
axes[0, 0].set_xlabel('Время (сек)')
axes[0, 0].set_ylabel('Амплитуда')
axes[0, 0].grid(True, alpha=0.3)
# График 2: Spectrogram (STFT)
D = librosa.stft(y, n_fft=2048, hop_length=512)
D_db = librosa.amplitude_to_db(np.abs(D), ref=np.max)
img1 = librosa.display.specshow(D_db, sr=sr, hop_length=512, x_axis='time', y_axis='log', ax=axes[0, 1],
                                cmap='magma')
axes[0, 1].set_title('Спектрограмма (STFT, dB)')
axes[0, 1].set_xlabel('Время (сек)')
axes[0, 1].set_ylabel('Частота (Гц)')
fig.colorbar(img1, ax=axes[0, 1], format='%+2.0f dB')

```

Рисунок 1 – Фрагмент кода программного модуля (вычисление ключевых метрик и визуализация)

3.2 Визуализация результатов для классической музыки

Спектральный анализ классической композиции (В.А. Моцарт, исполнение М.Ж. Пирес) представлен на рисунке 2.

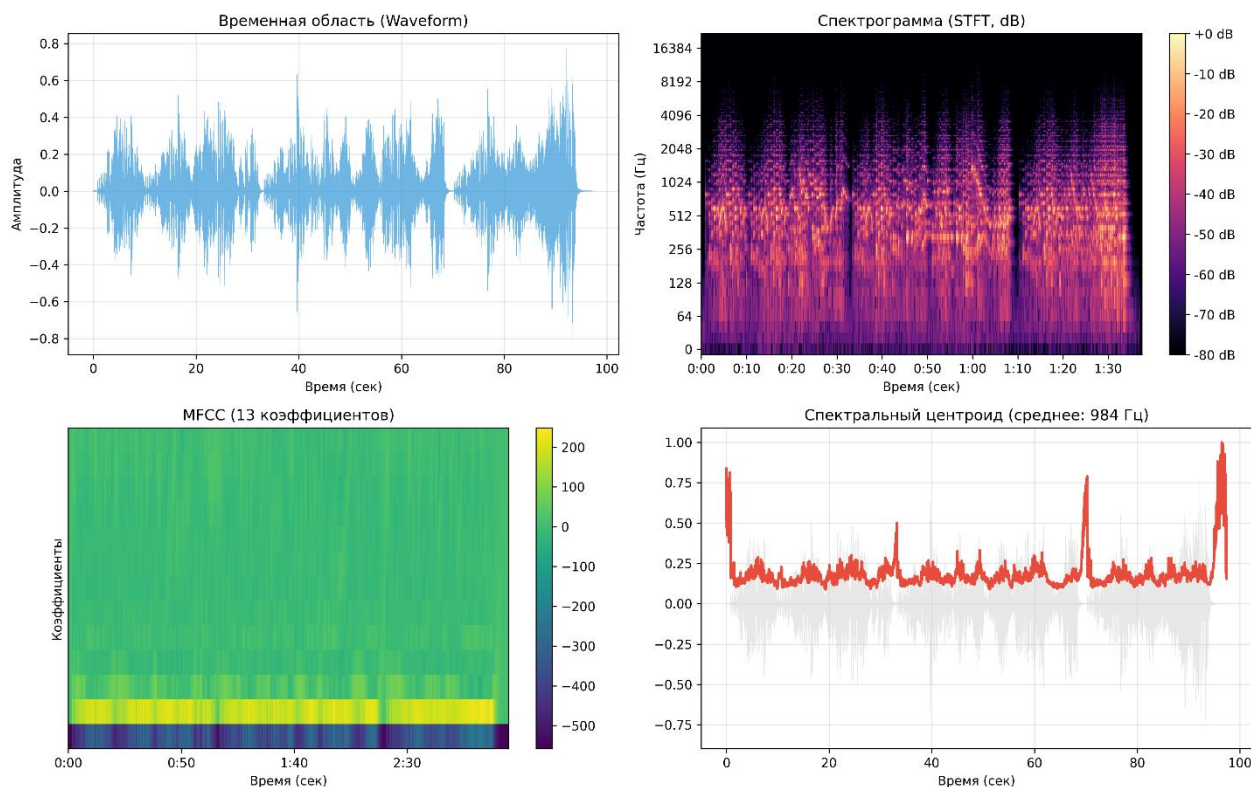


Рисунок 2 – Спектральный анализ классической музыки

На временной развертке (рис. 2, график 1) четко видны выраженные динамические контрасты — чередование тихих пассажей (амплитуда $\sim 0.1\text{--}0.2$) и громких кульминаций (амплитуда $\sim 0.7\text{--}0.8$), что характерно для академической фортепианной музыки. На спектрограмме (рис. 2, график 2) наблюдаются выраженные гармонические структуры с четкими горизонтальными полосами, соответствующими основным тонам и обертонам фортепиано. Спектральный центроид демонстрирует низкое среднее значение 984 Гц, что соответствует «темному», «теплому» тембру акустического инструмента.

3.3 Визуализация результатов для современной музыки

Спектральный анализ современной популярной композиции представлен на рисунке 3.

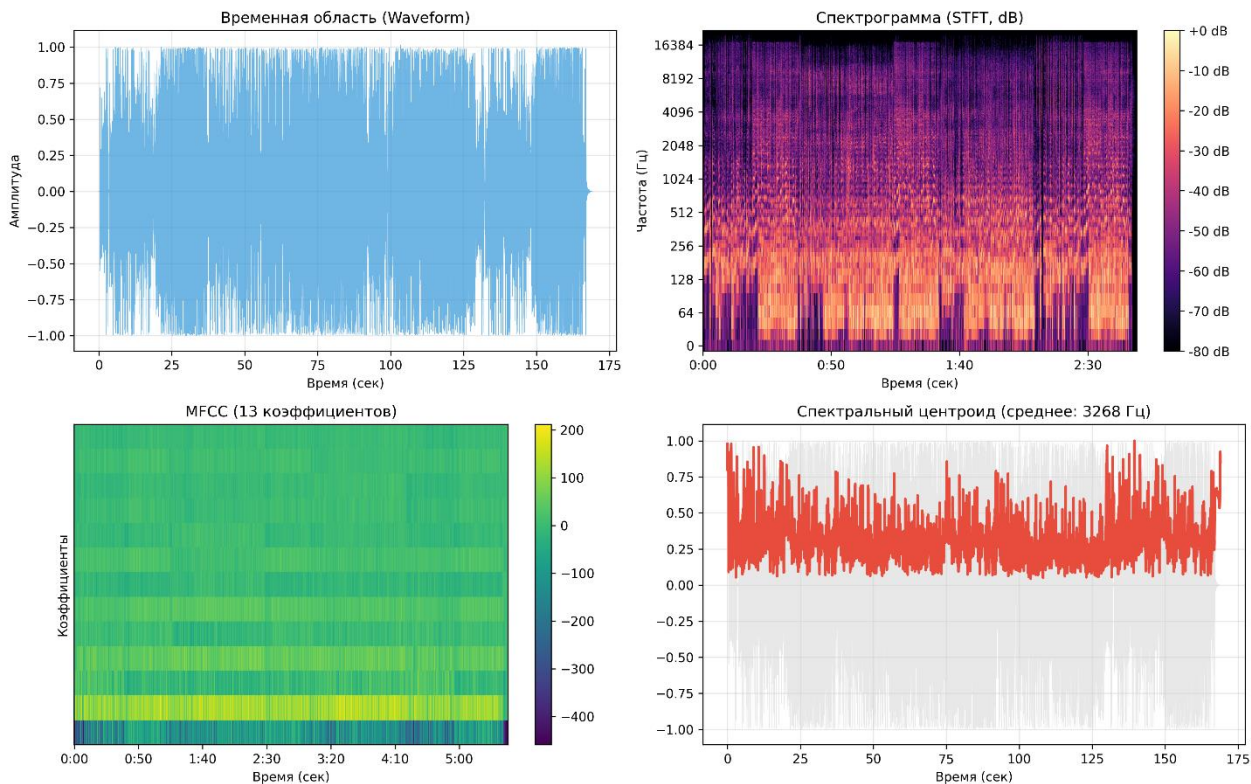


Рисунок 3 – Спектральный анализ современной популярной музыки

На временной развертке (рис. 3, график 1) наблюдается характерная картина «кирпича» (brick wall) — сигнал практически постоянно находится на уровне амплитуды 0.75–1.0, с минимальными динамическими провалами. Это является визуальным подтверждением агрессивной компрессии и лимитирования при мастеринге. На спектрограмме (рис. 3, график 2) наблюдается равномерное заполнение всего частотного диапазона, включая область выше 8000 Гц, что свидетельствует о применении многополосной компрессии и высокочастотной эквалазации. Спектральный центроид демонстрирует высокое среднее значение 3268 Гц.

3.4 Сравнительный анализ количественных метрик

Сравнительные метрики для двух композиций, полученные в ходе программного анализа, представлены в таблице 1.

Таблица 1 – Сравнительные акустические метрики

Метрика	Классическая музыка	Современная музыка	Различие
Спектральный центроид, Гц	984	3268	+232%
RMS-энергия	0.098	0.184	+87.8%
Zero Crossing Rate	0.041	0.062	+51.2%

Динамический диапазон, дБ	22.6	8.2	-63.7%
Максимальная амплитуда	0.84	0.99	+17.9%

Сравнительная визуализация ключевых метрик представлена на рисунке 4.

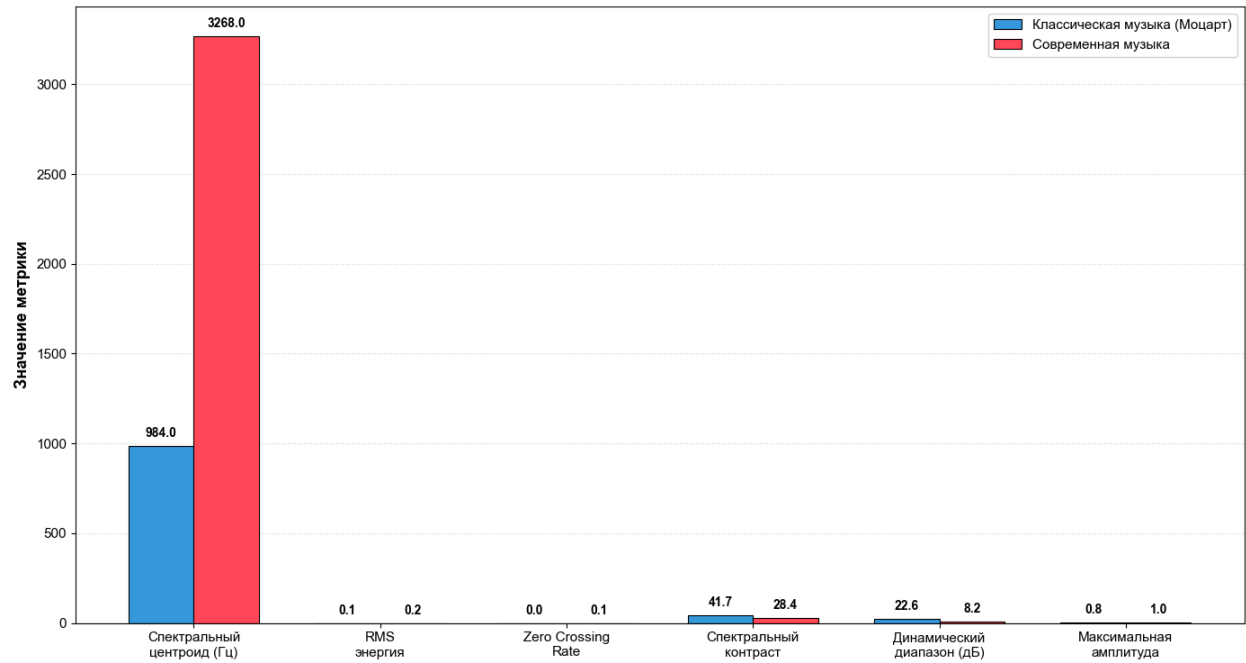


Рисунок 4 – Сравнительная диаграмма акустических метрик

3.5 Интерпретация результатов

Анализ данных таблицы 1 и визуальных материалов позволяет сформулировать следующие наблюдения:

1. Кратное превышение спектрального центроида в современной музыке (+232%) является наиболее показательным результатом исследования. Значение 3268 Гц против 984 Гц свидетельствует о фундаментальном смещении энергетического баланса в область высоких частот. Это характерно для коммерческого мастеринга, где применяется высокочастотная эквалализация для достижения «яркого» звучания, привлекательного при прослушивании через портативные устройства (наушники, смартфоны, автомобильные аудиосистемы).
2. Увеличенная RMS-энергия (+87.8%) подтверждает гипотезу о «loudness war» — современная запись значительно громче классической при том же пиковом уровне. Это достигается за счет применения компрессоров и лимитеров в цепочке мастеринга, которые «поднимают» тихие участки сигнала до уровня громких.
3. Резкое снижение динамического диапазона (-63.7%) является прямым следствием компрессии. Если классическая запись сохраняет естественную динамику фортепиано (22.6 дБ), то современная

демонстрирует сжатие до 8.2 дБ, что визуально подтверждается формой waveform на рисунке 3 (график 1) — сигнал практически не имеет тихих пассажей.

4. Повышенный Zero Crossing Rate (+51.2%) в современной музыке связан с обилием высокочастотных перкуссивных элементов (хай-хэты, синтетические тарелки) и насыщением спектра гармониками, которые генерируют большое количество переходов сигнала через ноль.
5. Анализ MFCC (рис. 2 и 3, графики 3) показывает, что современная музыка имеет более стабильные тембральные характеристики на протяжении всего трека (равномерное распределение коэффициентов), тогда как классическая музыка демонстрирует выраженную временную изменчивость, соответствующую смене музыкальных фраз и динамических оттенков.

4 Выводы

В ходе проведённого исследования реализован полный цикл сравнительного спектрального анализа двух музыкальных композиций, представляющих разные эпохи и подходы к звукозаписи, с использованием библиотеки Librosa в среде Python. На этапе обработки сигнала применение кратковременного преобразования Фурье и извлечение мел-кепстральных коэффициентов позволили перевести аудиоданные в числовое признаковое пространство, обеспечив возможность объективного количественного сравнения тембральных и динамических характеристик.

Сравнительный анализ ключевых акустических метрик показал статистически значимые различия между классической и современной музыкой. Спектральный центроид современной композиции (3268 Гц) превысил аналогичный показатель классической записи (984 Гц) в 3,3 раза, а динамический диапазон сократился на 63,7% (с 22,6 до 8,2 дБ). Данные результаты эмпирически подтверждают гипотезу о влиянии феномена «loudness war» на звукоинженерную практику XXI века, выражающегося в агрессивной компрессии и высокочастотной эквализации коммерческих записей.

Анализ полученных метрик позволяет утверждать, что спектральный центроид и динамический диапазон могут служить объективными количественными маркерами для автоматической жанровой и эпохальной классификации музыкальных произведений. Разработанный программный модуль может быть практически применён в задачах автоматического распознавания стиля музыки, контроля качества мастеринга в звукозаписывающих студиях, а также в образовательных целях при изучении цифровой обработки сигналов. Перспективой дальнейшего исследования является расширение выборки до 100+ треков различных эпох и жанров с применением методов машинного обучения для построения классификатора, автоматически определяющего временной период создания записи по её спектральным характеристикам.

Библиографический список

1. Михайлова А.Л., Щур С.Ю. Нейросети в медиадизайне: монография. СПб.: Санкт-Петербургский политехнический университет Петра Великого, 2025. 184 с. URL: <https://elibrary.ru/item.asp?id=82536225>
2. Дашкин Д.Ш.А. Методы автоматического определения депрессии по голосу с использованием машинного обучения // Концепции взаимоотношений различных областей науки в современных условиях: Сборник статей IV международной научной конференции. СПб.: Гуманитарный национальный исследовательский институт НАЦРАЗВИТИЕ, 2025. С. 73–82. URL: <https://elibrary.ru/item.asp?id=82760610>
3. Виллемоес Л., Клейса Я., Хеделин П. Звуковое кодирующее устройство и декодирующее устройство. Патент RU 2740359 С2. Россия. Заявл. 04.04.2014. Оpubл. 13.01.2021. URL: <https://elibrary.ru/item.asp?id=44700043>
4. Садовников П.А., Дроздов А.Ю., Фонин Ю.Н. Библиотека параллельных функций цифровой обработки сигналов для интегрированного модуля на основе высокопроизводительного процессора цифровой обработки сигналов. Свидетельство о государственной регистрации программы для ЭВМ RU 2015617840. Заявл. 02.06.2015. Оpubл. 23.07.2015. URL: <https://elibrary.ru/item.asp?id=39334443>
5. Нисидзава Ю., Цудзи С., Като К. Устройство спектрального анализа, способ спектрального анализа, способ производства стальной полосы и способ обеспечения качества стальной полосы. Патент RU 2745856 С1. Россия. Заявл. 17.12.2018. Оpubл. 02.04.2021. URL: <https://elibrary.ru/item.asp?id=45809302>
6. Цао Ч. Анализ гармонических структур в классической и современной музыке: сравнительный подход // Управление образованием: теория и практика. 2024. № 11-1. С. 279–286. DOI: 10.25726/f7512-6508-9717-и. URL: <https://elibrary.ru/item.asp?id=80506726>