

МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение
высшего образования

«САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г.ЧЕРНЫШЕВСКОГО»

Кафедра информатики и программирования

**РАЗРАБОТКА СИСТЕМЫ ДЛЯ РАСПОЗНАВАНИЯ
ПОСЛЕДОВАТЕЛЬНОСТИ АККОРДОВ В МУЗЫКАЛЬНЫХ
ПРОИЗВЕДЕНИЯХ**

АВТОРЕФЕРАТ МАГИСТЕРСКОЙ РАБОТЫ

студента 2 курса 273 группы

направления 02.04.03 Математическое обеспечение и администрирование
информационных систем

факультета компьютерных наук и информационных технологий

Никифорова Семена Алексеевича

Научный руководитель:

Доцент, к.э.н., доцент

_____ Кабанова Л. В.

подпись, дата

Зав. кафедрой:

к.ф.-м.н., доцент

_____ Огнева М. В.

подпись, дата

Саратов 2026

ВВЕДЕНИЕ

Актуальность темы. В настоящее время вся музыкальная индустрия интенсивно использует информационные технологии для хранения, создания, обработки музыки и любой информации о музыке. Благодаря большому количеству библиотек с аудиоматериалами, воспроизводящими звуки любых музыкальных инструментов, больше не нужно нанимать целый оркестр, чтобы обогатить звучание музыкального произведения.

Необходимость получения разнообразной информации о конкретной цифровой звукозаписи порождает множество задач, связанных с обработкой звука: идентификация композиции, нахождение разных версий одной композиции, определение заданной композиции в потоке звука, поиск похожих композиций, определение мелодии композиции для последующего воспроизведения на музыкальном инструменте и другие.

В сфере музыкального творчества многие высококвалифицированные специалисты и любители каждый день сталкиваются с задачей определения услышанных ими нот или аккордов. С развитием технологий и алгоритмов цифровой обработки сигналов появилась возможность ее автоматизации.

Автоматическое распознавание высоты звуков — это довольно трудоёмкий процесс цифровой обработки звуковых сигналов с последующим определением высотности нот, звучащих в каждый момент музыкального фрагмента. Степень трудоёмкости данного процесса зависит от ряда факторов, подлежащих обязательному учёту в ходе распознавания. Этими факторами является ритм, громкость, тембр звука и специфика инструмента, с помощью которого воспроизводится музыкальный фрагмент. Также осложняют задачу и некоторые составляющие процесса распознавания: разделение музыкального фрагмента на одинаковые части, в пределах которых не подразумевается смена высотности ноты или аккорда, применение весовой функции к каждому фрагменту записи, фильтрация шумов исходной записи. Инструмент, который будет распознавать аккорды может сильно упростить анализ неизвестной ранее музыкальной записи

звуковому инженеру или аранжировщику для дальнейшей работы с производением.

В ходе исследования были разработаны алгоритмы распознавания аккордов на основе скрытых марковских моделей и рекуррентных нейронных сетей долгой краткосрочной памяти, обеспечивающие автоматическое выделение и классификацию гармонических последовательностей по цифровой звукозаписи.

Функционирование алгоритма в изолированном режиме командной строки существенно ограничивает его практическую применимость. По этой причине было создано полнофункциональное веб-приложение с интуитивно понятным интерфейсом. Этот инструмент позволяет музыкантам не только загружать готовые аудиофайлы для мгновенного получения визуализации аккордов, но и задействовать режим реального времени для распознавания гармонии через микрофон устройства непосредственно в процессе исполнения. Благодаря таким возможностям педагоги могут наглядно демонстрировать ученикам гармонические структуры как на примере аудиозаписей, так и во время живой игры, а исследователи получают эффективное средство для обработки больших массивов музыкальных композиций. Данное приложение работает в любом браузере без необходимости установки специализированного программного обеспечения, что делает его доступным широкой аудитории.

Цель магистерской работы – разработка системы для распознавания последовательности аккордов в музыкальных произведениях.

Поставленная цель определила **следующие задачи**:

1. изучение теоретических сведений о звуке и его свойствах, о музыкальном строе и составе аккордов;
2. анализ цифрового звука для распознавания аккордов;
3. постановка задачи распознавания аккордов;
4. исследование существующих методик и решений в области распознавания аккордов;

5. обзор и оценка существующих информационных систем распознавания аккордов в музыке;
6. анализ методов представления аудиосигналов для задач распознавания аккордов на основе хроматических признаков;
7. реализация процедуры сопоставления извлечённых признаков с шаблонами аккордов;
8. оценка точности и качества распознавания аккордов путём сравнения с эталонными аннотациями;
9. реализация модели распознавания аккордов на основе скрытых цепей Маркова;
10. подготовка и предобработка данных с целью расширения обучающей выборки;
11. проектирование и программная имплементация однонаправленных и двунаправленных моделей глубокого обучения на базе фреймворка PyTorch;
12. формирование требований к приложению по распознаванию аккордов, проектирование клиент-серверной модели;
13. определение и выбор технологического стека;
14. создание приложения с пользовательским интерфейсом для распознавания аккордов на основе загрузки аудиофайлов и онлайн распознавания через микрофон.

Методологические основы разработки системы для распознавания последовательности аккордов в музыкальных произведениях представлены в работах Мелешкина А. А. и Фирсовой С. А. [1], Анисимовой С. В. и Головатого А. И. [2], Martyna M. и Jacek M. [3], Wang S. [4], Глазырина Н. Ю. [5], Briot, J. P. [6].

Практическая значимость магистерской работы заключается в создании доступного и удобного инструмента автоматического распознавания аккордов, который облегчает анализ музыкальных записей для широкого круга пользователей: аранжировщикам, преподавателям, исследователям.

Разработанные алгоритмы на основе скрытых марковских моделей и рекуррентных нейронных сетей обеспечивают хорошую точность классификации гармонических последовательностей, а вебприложение с режимом загрузки файлов и распознаванием в реальном времени через микрофон позволяет музыкантам получать визуализацию аккордов без установки дополнительного ПО. Разработанное приложение сокращает затраты времени и труда при анализе гармонии, а также расширяет возможности интерактивного обучения музыке.

Структура и объём работы. Магистерская работа состоит из введения, 9 разделов, заключения, списка использованных источников и 7 приложений. Общий объём работы – 139 страниц, из них 81 страница – основное содержание, включая 22 рисунка, список использованных источников информации – 40 наименований.

КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ

Первый раздел «Методики распознавания аккордов» посвящен обзорному анализу подходов и алгоритмов, применяемых для автоматического определения гармонических структур по аудиозаписям. В разделе также рассматриваются преимущества и ограничения каждого класса методов в зависимости от музыкального контекста.

Второй раздел «Теоретические сведения о звуке и его свойствах» посвящен изложению базовых музыкально-акустических понятий, необходимых для формальной постановки задач распознавания гармонии и разработки алгоритмов. В разделе даются понятия интервала и его двух характеристик — количественной (число ступеней между звуками) и качественной (количество тонов и полутонов). Рассмотрено определение музыкального строя как системы соотношений нот и частот, а также дана формулировка аккорда как одновременного сочетания трёх и более звуков с описанием трезвучия и септаккорда.

В подразделе 2.1 «Анализ цифрового звука для распознавания аккордов» рассматриваются методы извлечения частотной информации и её

интерпретации для последующего распознавания гармонических структур. Описывается модель ноты как синусоидальной компоненты с определённой основной частотой и набором обертонов, формирующих тембр. Показано, как переход в частотную область через быстрое преобразование Фурье и формирование спектрограмм позволяет выделять частотные пики, соответствующие высотам звуков и их комбинациям.

Третий раздел «Постановка задачи» посвящен формализации задачи автоматического определения последовательности аккордов в цифровой звуковой записи и описанию свойств входных сигналов, влияющих на решение этой задачи. Задача сформулирована математически следующим образом. Пусть задан непрерывный звуковой сигнал $x(t)$, и множество возможных названий аккордов Y . Необходимо для каждого момента времени указать аккорд y , звучащий в этот момент.

Приводится стандартная модель перехода к цифровому представлению через дискретизацию и квантование. Для этого с некоторой частотой ν раз в секунду измеряется амплитуда функции (дискретизация), после чего каждое полученное значение заменяется на — ближайшее из заданного множества возможных значений амплитуды (квантование). В соответствии с классической теоремой Котельникова, если спектр сигнала ограничен частотой $\nu/2$, то исходный сигнал может быть восстановлен однозначно и без потерь по измеренным значениям.

Разбираются важные ограничения цифровых звукозаписей, влияющие на качество распознавания: наличие шумов, инструментов с неопределенной высотой звучания, ритма, одновременное звучание нескольких инструментов.

Четвертый раздел «Существующие информационные системы распознавания музыки и аккордов» посвящен обзору и сопоставительному анализу популярных систем, среди которых приложения Shazam, Chord Tracker, Chord AI и ChordMini. После анализа существующих аналогов были выделены критерии разрабатываемого приложения. Необходимо обеспечить универсальную доступность, которая позволит работать в любом

современном браузере без необходимости установки дополнительного программного обеспечения. Синхронизированная визуализация временной шкалы, гарантирующей постоянную видимость текущего аккорда, которая отсутствует у многих существующих приложений. В отличие от CLI-утилит присутствие визуализации последовательности аккордов обеспечивает профессиональный пользовательский опыт с интерактивными элементами управления.

Пятый раздел «Распознавание аккордов на основе шаблонов» описывает реализацию классического подхода к автоматическому определению гармонических структур, основанному на сегментации аудиозаписи и сопоставлении признаков с заранее заданными эталонными паттернами аккордов. Задача сформулирована следующим образом, необходимо разбить музыкальную запись на временные фрагменты, задать начало и конец каждого аккорда и присвоить каждому фрагменту метку аккорда из заданного множества.

Подраздел 5.1 «Хроматическое представление признаков» раскрывает способ перевода аудиозаписи в хроматическую последовательность. На основе равномерно темперированной шкалы хроматические значения соответствуют следующему набору {C, C#, D, D#, E, F, F#, G, G#, A, A#, B}. Данный список отождествляется с множеством от 0 до 11. Хроматическая характеристика звучащего аккорда может быть выражена как 12-мерный вектор .

Описаны три различных метода вычисления хроматических характеристик: на основе оконного преобразования Фурье, на основе фильтра с бесконечной импульсной характеристикой и на основе преобразования ConstantQ . Введено понятие хромограммы как временной последовательности 12мерных хромав екторов, показывающих интенсивность каждого полутона во времени. Описана функция вычисления хромограммы из аудиофайла. Основным методом выбрано оконное преобразование Фурье как универсальное, быстрое и стабильное средство

получения спектральных характеристик.

Подраздел 5.2 «Мера сходства и процедура сопоставления» описывает второй ключевой этап системы распознавания аккордов на основе шаблонов — сопоставление каждого хроматического вектора признаков с меткой аккорда из заданного множества. В качестве множества аккордов выбраны 24 шаблона — 12 мажорных и 12 минорных трезвучий. Шаблоны заранее вычислены, они представляет собой хроматический вектор соответствующего аккорда. Процедура сопоставления заключается в поиске для каждого кадра такой метки аккорда, которая обеспечивает максимальное сходство между его шаблоном и вектором признаков.

Сравнение результатов автоматического распознавания с эталонной аннотацией показывает согласованность для большинства кадров, что подтверждает работоспособность предложенного метода.

Шестой раздел «Оценка распознавания аккордов» посвящен методологии и результатам количественной оценки качества автоматического распознавания аккордов через сравнение предсказанных меток с эталонной аннотацией.

Предложен следующий подход к оценке. Для сценария распознавания аккордов необходимо определить набор элементов как \mathcal{E} . Метка, не относящаяся к аккорду, не учитывается. Тогда положительные или релевантные элементы можно определить следующим образом: \mathcal{E}_+ . И также элементы, которые оценены как положительные или полученные элементы, можно представить, как \mathcal{E}_+ . С учетом этих понятий элемент называется истинно положительным (TP), если метка верна, то есть \mathcal{E}_+ . В противном случае называется ложноположительным результатом (FP), а \mathcal{E}_- — ложноотрицательный результат (FN). Все остальные элементы в \mathcal{E} называются истинно отрицательными. С помощью этих понятий можно определить стандартную точность (P), полноту (R) и F-меру (F): P, R, F .

По вычисленным метрикам оценки качества шаблонный метод показал следующие результаты $P = 0.757, R = 0.768, F = 0.763$.

Седьмой раздел «Скрытая марковская модель» описывает применение скрытой марковской модели для учета временных зависимостей между аккордами и улучшения распознавания по сравнению с чистым шаблонным подходом. Для построения матрицы вероятностей переходов между аккордами был использован датасет Chords and Lyrics Dataset, содержащий большое количество песен с разметкой последовательностей аккордов.

Распознаватель аккордов на основе скрытой марковской модели превосходит шаблонный подход. Улучшения достигаются за счёт модели переходов, которая вводит контекстно-зависимое сглаживание.

Восьмой раздел «Рекуррентная нейронная сеть» описывает применение глубокого обучения, в частности рекуррентных нейронных сетей (RNN) и их варианта с долгой краткосрочной памятью (LSTM), для распознавания аккордов. Для обучения и тестирования использовался набор данных Isophonics, содержащий 180 композиций группы The Beatles. За счёт аугментации методом питч-шифтинга набор был увеличен в 5 раз для повышения обобщающей способности моделей.

В ходе предобработки аудиосигналы преобразовывались в спектрограммы, где наилучшие результаты показало константное Q-преобразование. Реализация нейросетей выполнена во фреймворке PyTorch на базе архитектуры LSTM, учитывающей контекст аккордовых последовательностей. По итогам тестирования наивысшую точность Accuracy = 73.061% показала двунаправленная модель BiLSTM.

Девятый раздел «Разработка веб-приложения» посвящен проектированию клиент-серверной архитектуры для системы автоматического распознавания аккордов, разделяющей вычислительную часть и интерфейс пользователя. Описан сценарий обработки аудиофайла, включающий извлечение признаков, формирование JSON-результата и визуализацию гармонической структуры.

Подраздел 9.1 «Выбор технологического стека» обосновывает выбор

Python и FastAPI для серверной части благодаря их высокой производительности, асинхронности и средствам обработки аудиоданных. Для клиентской части были выбраны Vue.js и TypeScript, которые обеспечивают модульность и современную компонентную разработку.

Подраздел 9.2 «Реализация системы распознавания аккордов» посвящен проектированию и программной реализации серверных эндпоинтов, включая POST-запросы для пакетного анализа (HMM и LSTM) и специализированный протокол WebSocket для потоковой обработки сигналов. Разработана и внедрена система многоуровневой валидации загружаемых аудиофайлов по MIME-типу, размеру, расширению и сигнатуре. Описывается разработка реактивного адаптивного интерфейса одностраничного приложения. Созданы логические блоки для пакетного режима, включая drag-and-drop селектор и синхронизированный таймлайн-визуализатор, и полноценный модуль онлайн-распознавания с микрофона.

ЗАКЛЮЧЕНИЕ

В ходе выполнения выпускной квалификационной работы была разработана система автоматического распознавания музыкальных аккордов, основанная на нескольких подходах к анализу аудиосигналов. В частности, была реализована модель на основе скрытой марковской модели, которая позволила учитывать контекстную информацию за счёт вероятностей переходов между аккордами и тем самым повысить устойчивость распознавания. Наряду с этим была исследована и применена модель рекуррентной нейронной сети долгой краткосрочной памяти, способная выявлять временные зависимости в музыкальном сигнале и эффективнее обрабатывать последовательности аккордов.

Проведённое сравнение показало, что использование HMM и LSTM позволяет уменьшить количество случайных и кратковременных ошибок распознавания по сравнению с шаблонными методами, особенно в музыкальных фрагментах со сложной гармонической структурой и неоднозначностями. Полученные результаты подтверждают, что применение

методов машинного обучения является перспективным направлением для повышения точности автоматического распознавания аккордов.

Кроме того, в рамках работы было разработано полнофункциональное веб-приложение, которое обеспечивает загрузку аудиофайлов и онлайн-распознавание аккордов в реальном времени. Это позволило объединить алгоритмы обработки аудиосигналов с удобным интерактивным интерфейсом и создать завершённый инструмент для практического использования. Приложение обеспечивает полный цикл работы, начиная от загрузки аудиоматериала до анализа гармонической структуры композиции и отображения результатов пользователю.

Таким образом, цель выпускной квалификационной работы достигнута. Разработанная система может быть использована как основа для дальнейшего совершенствования методов автоматического распознавания аккордов, а также для расширения функциональности веб-приложения за счёт повышения точности моделей и добавления новых режимов анализа.

По тематике магистерской работы была опубликована статья «Разработка веб-приложения для автоматического распознавания музыкальных аккордов» в научном журнале Интернаука [40]. Кроме того, материалы и основные результаты были представлены на студенческой научной конференции факультета компьютерных наук и информационных технологий.

Отдельные части магистерской работы были опубликованы:

1. Никифоров, С.А., Кабанова, Л.В. РАЗРАБОТКА ВЕБ-ПРИЛОЖЕНИЯ ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ МУЗЫКАЛЬНЫХ АККОРДОВ // Интернаука: электрон. научн. журн. 2026. № 15(426).

Основные источники информации:

1. Мелешкин, А. А., Фирсова, С. А. РАЗРАБОТКА СИСТЕМЫ ДЛЯ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ АККОРДОВОЙ ПОСЛЕДОВАТЕЛЬНОСТИ В ЦИФРОВЫХ АУДИОФАЙЛАХ / А. А. Мелешкин, С. А. Фирсова // Огарёв-Online. – 2021. – №12 (165).

2. Анисимова, С. В., Головатый, А. И. АЛГОРИТМЫ РАСПОЗНАВАНИЯ МУЗЫКАЛЬНЫХ ОТРЫВКОВ, НОТ И АККОРДОВ // Вестник Полоцкого государственного университета. Серия С. Фундаментальные науки. – 2021. – №4.
3. Martyna, M., Jacek, M. Training chord recognition models on artificially generated audio: Training chord recognition models on artificially generated audio. // Neural Computing & Applications. – 2026. – P. 22–24.
4. Wang, S. Intelligent Music Chord Recognition and Evaluation Based on Convolution and Attention / S. Wang, L. Xiaobing, Q. Zhou, Y. Tie, Y. Gao, X. Zhang // IEEE International Conference on Multimedia and Expo Workshops (ICMEW), – 2024. – P. 1–6.
5. Глазырин, Н. Ю. О задаче распознавания аккордов в цифровых звукозаписях / Н. Ю. Глазырин // Известия Иркутского государственного университета, – 2013. – №2.
6. Briot, J. P. From artificial neural networks to deep learning for music generation: history, concepts and trends // Neural Computing and Applications. – 2021. – P. 39–65.
7. Mycka, J., Mańdziuk J. Classification of piano performers with deep learning models // Journal of Computational Science. – 2026. – P. 73–79.