

УТВЕРЖДАЮ

Первый проректор – проректор по научной работе Московского физико-технического института (государственного университета), доктор технических наук, профессор



Горшков
Олег Анатольевич

«05» декабря 2014 г.

ОТЗЫВ

ведущей организации на диссертацию **Глазырина Николая Юрьевича** «Алгоритмическое распознавание аккордов в цифровом звуке», представленную на соискание ученой степени кандидата физико-математических наук по специальности 05.13.18 — математическое моделирование, численные методы и комплексы программ

Актуальность темы

Тема диссертации относится к области музыкального информационного поиска, получившей своё развитие в последние 20 лет. Автор ставит задачу автоматического определения последовательностей аккордов в музыкальных записях. Эта задача очень важна как способ упрощения подбора аккомпанемента для музыкантов, так и как метод автоматического анализа музыкальных композиций с целью поиска или кластеризации.

Для решения этой задачи необходимо построение сложной многостадийной системы анализа звука, включающей в себя множество алгоритмов, в том числе, определение спектрограммы, фильтрацию шума и выделение значащих особенностей спектрограммы, определение временной структуры записи, уменьшение размерности обрабатываемых данных, использование знаний о структуре звучания музыкальных инструментов и структуре аккордов, использование классификаторов для получения предположений относительно звучащего аккорда. Кроме того, необходимо попутно решать совсем другие вопросы, например, определение расстройки инструментов, играющих данную запись. Недостаточная проработка любой из этих стадий не позволит создать хорошую систему автоматического определения аккордов.

Все эти алгоритмы являются самостоятельными направлениями исследования, требующими совершенно разных подходов из области цифровой обработки сигналов, машинного обучения, музыкальной теории и т.д. Поэтому сложность поставленной

задачи обусловлена уже необходимостью в хорошей степени разобраться в этой огромной области, и успешно соединить вместе имеющиеся знания и запрограммировать требуемые алгоритмы, создав цельную систему анализа звукозаписей.

Научная задача и подход к решению

Диссертация посвящена решению задачи определения последовательности аккордов в звукозаписях. Формально, в этой задаче по оцифрованной звукозаписи необходимо определить его разметку — последовательность названий аккордов с указанием моментов начала и конца звучания каждого из них. Для этого поступающая на вход звукозапись преобразуется сначала в частотно-временное представление в виде спектрограммы, а затем в последовательность векторов признаков, для которой в дальнейшем решается задача классификации с целью определения последовательности звучащих аккордов.

В работе предложены два подхода к решению поставленной задачи. Первый основан на применении к спектрограмме специальным образом выбранных преобразований, направленных на выделение характерных черт в музыкальном сигнале. В этом подходе учитываются специфические для музыкальных звукозаписей особенности спектра звука, подавляются нежелательные шумы. После этого спектрограмма преобразуется в последовательность векторов признаков, к каждому из которых применяется простой алгоритм классификации для определения соответствующего вектору аккорда.

Второй подход основан на применении многослойной нейронной сети для преобразования спектрограммы в последовательность векторов. Для предварительной настройки весов используется послойное обучение без учителя, после чего сеть дообучается методом градиентного спуска на обучающей выборке. Последующая классификация полученных векторов признаков осуществляется аналогично первому подходу.

Основные результаты и их новизна

В рамках диссертационного исследования Н.Ю. Глазыриным были получены следующие результаты.

1. Разработан метод выделения в спектре звука компонент, соответствующих звучанию музыкальных инструментов. Метод успешно применён для получения представления звукозаписей в виде последовательности векторов признаков. Полученное представление позволяет использовать простой необучаемый классификатор для достижения сравнимой с более сложными аналогами точности распознавания последовательностей аккордов.
2. Предложен новый подход для преобразования спектрограммы звукозаписи в последовательность векторов признаков с использованием многослойной

нейронной сети – автоассоциатора. Полученные в результате векторы признаков позволяют добиться сравнимой с предыдущим методом точности распознавания аккордов, при этом сам подход имеет большой потенциал для развития.

3. Все предложенные методы вместе с набором вспомогательных алгоритмов реализованы в рамках единой системы анализа звукозаписей.
4. Проведено экспериментальное сравнение предложенных методов с существующими аналогами в рамках международных конкурсов MIREX Audio Chord Detection в 2012 и 2013 годах. Предложенный метод показал сравнимую с конкурентами точность распознавания аккордов при более высокой скорости обработки звукозаписей.

Достоверность полученных результатов

Достоверность полученных результатов обусловлена многочисленными экспериментальными подтверждениями, а также проверкой в рамках международных конкурсов. Корректность экспериментов обусловлена использованием стандартных метрик и стандартных наборов данных, допускающим сравнение с аналогами.

Значимость полученных результатов для практики

Практическая значимость работы Н.Ю. Глазырина заключается в том, что разработанные методы позволяют с достаточно высокой скоростью и точностью преобразовывать звукозаписи в более удобное для дальнейшей обработки представление. Получаемые на выходе последовательности аккордов могут быть использованы музыкантами и музыковедами, а также системами для автоматического анализа и индексации музыки по содержанию. Кроме того, некоторые из предложенных преобразований могут быть использованы для решения задач, связанных с выделением других типов музыкальной информации, таких как информация о тональности, ритме, структуре композиции.

Полученные результаты могут использоваться в научных организациях и компаниях, занимающихся исследованиями и разработками в области обработки звукозаписей, создания музыки, обработки музыки в больших масштабах. Методы могут быть особенно интересны таким компаниям, как Яндекс, Google, Shazam, SoundHound, Spotify.

Замечания

1. Некоторые из использованных алгоритмов могут быть заменены на более очевидные. Например, непонятно игнорирование метода главных компонент для уменьшения объема информации при анализе спектральных векторов.

2. Результаты, изложенные в главе 4, выглядят недостаточно разработанными. Судя по представленным результатам, небольшие нейронные сети с тысячами параметров дают приблизительно такое же качество определения аккордов, как и огромные сети с сотнями тысяч параметров, и сравнимое с простыми эвристическими алгоритмами, свидетельствует либо о слишком больших потерях информации при подготовке данных, либо о некорректном использовании нейросетевых моделей. Это не означает недобросовестности автора, но в этом направлении есть большое пространство для дальнейшей работы.
3. Автор не всегда аккуратно относится к терминологии – например, термин «constant-Q transform», переводимый обычно как «фильтрация с постоянной добротностью», переводится им на протяжении всей диссертации собственным термином «преобразование постоянного качества» (стр. 8 и далее).
4. В диссертации встречается существенное количество дефектов текста - неточные формулировки (например, стр. 11, 16), оборванное предложение (стр. 55), ошибки в формулах (стр. 21, 23), в некоторых случаях отсутствуют необходимые пояснения (стр. 45, 46, 67).

Приведённые замечания не влияют на общую положительную оценку диссертации.

Заключительная оценка

Диссертация Н.Ю. Глазырина является законченной научно-исследовательской работой, имеющей высокую теоретическую и практическую ценность, а её тема является актуальной. Полученные результаты являются новыми и достоверными. Результаты служат развитию знаний в области музыкального информационного поиска, а методология включает в себя приёмы из областей спектрального анализа, цифровой обработки сигналов, машинного обучения, что обосновывает соответствие специальности 05.13.18 — математическое моделирование, численные методы и комплексы программ.

Полученные результаты опубликованы в научной печати и докладывались на российских и международных конференциях. Автореферат полностью и точно отражает содержание диссертации. Диссертационная работа «Алгоритмическое распознавание аккордов в цифровом звуке» полностью соответствует требованиям Положения о присуждении учёных степеней, предъявляемым к диссертациям на соискание степени кандидата физико-математических наук, а соискатель заслуживает присуждения учёной степени кандидата физико-математических наук по специальности 05.13.18 — математическое моделирование, численные методы и комплексы программ.

Отзыв на диссертацию составлен доктором физико-математических наук, профессором Буниной Еленой Игоревной и обсуждён на заседании кафедры анализа дан-

ных факультета инноваций и высоких технологий Московского физико-технического института (государственного университета) «15» ноября 2014 г., протокол № 4.

Заместитель заведующего кафедрой анализа данных
факультета инноваций и высоких технологий МФТИ,
доктор физ. – матем. наук, профессор

Бунина
Елена Игоревна

Почтовый адрес: 141700, Московская область, г. Долгопрудный, Институтский пер.,9

Телефон: 8 (495) 408 78 81

Адрес электронной почты: helenbunina@yandex.ru

Организация – место работы: федеральное государственное автономное образовательное учреждение высшего профессионального образования «Московский физико–технический институт (государственный университет)»