

Поиск и идентификация музыкальных произведений

П. П. Кокорин, К. В. Фролов

Рассмотрены подходы компьютерной идентификации и поиска музыки. Предложен метод ассоциативного поиска фрагментов в музыкальном произведении. Показана практическая эффективность подхода.

Введение

В настоящее время в большинстве существующих поисковых систем применяется принцип поиска по ключевым словам (идентификаторам). Сервисы поиска аудио файлов, предоставляемые крупными поисковыми сайтами (Google, Yandex и др.), используют в качестве запроса имя файла, имя исполнителя, жанр и прочие данные, заданные в виде тегов. В своей работе эти сайты могут использовать и специально организованные тематические каталоги музыкальных произведений, негативным моментом применения которых является необходимость участия человека для составления и поддержания каталога в актуальном состоянии. Более подробно об основных принципах работы современных поисковых систем в сети Интернет можно ознакомиться в [1].

Принципиальная возможность построения ассоциативных идентификаторов связана с установлением связи между стандартно заданной древовидной структурой уникальных идентификаторов с полем семантико-смысовых понятий [2].

Понимание потребности в этом и вызывает последующее развитие от стандарта MPEG-1 к стандарту MPEG-7 [3]. Последовательность шагов от MPEG-1 к MPEG-7 есть переход от простого кодирования к организации возможности понимания – узнавания информационного сообщения. Основная цель применения MPEG-7 это поиск мультимедиа информации (так же как сейчас можно найти текст по какому-нибудь предложению). Работы над стандартом были начаты еще в середине прошлого десятилетия, но до сих пор каких-то значительных прорывов в этом направлении сделано не было.

Форматы хранения музыкальных данных

Существующее многообразие форматов музыкальных файлов можно разделить по классам решаемых задач:

- хранение цифрового представления сигнала (WAV, AU);*
- компрессия оцифрованного сигнала (MP3, OGG, WMA);*
- описание структуры сигнала (MIDI, VOC и др.)*

В формате MIDI звуковой поток представляется в виде последовательности звуковых событий, таких как начало или окончание каждого элементарного звука в последовательности, установка его параметров, т.е. звуковые сигналы, представляют собой последовательность идентифицированных элементарных звуков с известными параметрами (амплитуда, основная частота, спектр), возникающих в известные моменты времени и имеющих конечную длительность. В этом случае звуковой поток может быть описан как абстрактный текст, т.е. последовательность символов из фиксированного набора [4].

Подходы к поиску фрагментов музыкальных в произведениях

В сети Интернет существует несколько ресурсов, предлагающих пользователям найти по фрагменту интересующую их мелодию. Компания Melodis запустила сайт поиска музыки «Midomi». Сайт позволяет пользователям напевать даже без слов или наспистывать мелодии, а сервер находит названия песен. Мультимодальная адаптивная система распознавания (Multimodal Adaptive Recognition System) предназначена для анализа различных факторов в распознавании, таких как высота, изменение темпа, содержимое речи и расположение интервалов.

Другое решение, принадлежащее компании Phonehits, предназначено для автоматизированного распознавания музыки в реальном времени. Программа обеспечивает выдачу линка на музыку, соответствующую той, что в данный момент проигрывается по радио. Система идентифицирует практически любую мелодию в течение 12 с. Сервис был испытан компанией Phonehits в декабре 2004 г., тогда к системе подключили базу, насчитывающую около 7000 полифонических рингтонов. Были получены впечатляющие результаты – более 95% песен и более 80% композиций были безошибочно идентифицированы в реальном времени и связаны с соответствующим рингтоном. Эта технология распознавания музыки разработана в России компанией «Центр речевых технологий» (<www.speechpro.ru>), и называется Jingle Tracker.

На сайте <www.songtapper.com>, для поиска музыкального произведения, пользователям предлагается задать ритм искомого фрагмента с помощью клавиш клавиатуры. По полученной последовательности нажатий клавиш строится картина ритма фрагмента. Далее ритм фрагмента кодируется в строку, состоящую из символов: ‘d’ – неизменная длительность, ‘s’ – спад и ‘u’ – подъем. По полученной строке ищутся нестрогие совпадения в базе данных [5]. Испытания алгоритма показали, что для успешного поиска мелодии в базе, состоящей из 30 детских песен, необходимо сделать минимум 15 ударов по клавише. В таком случае результат был удовлетворительным, даже если пользователи не были музыкантами, и делали ошибки при задании ритма.

Предлагаемый метод ассоциативного поиска

Успешное применение семиологического подхода для ассоциативного поиска по текстам (www.visualworld.ru) дает предпосылки для использования этого метода и для поиска музыкальных произведений заданных в виде абстрактного текста – MIDI формате.

Считается, что MIDI формат достаточен для аутентификации музыкального произведения, так как представляет музыкальный сигнал в виде совокупности параметров, таких как амплитуда, основная частота, спектр, время звучания и др., проявление которых и влияет на «узнаваемость» мелодии. Справедливость такого утверждения подтверждается рядом исследований [6, 7].

Так в работе [7] описывается нейросетевая модель музыкального лада, основанная на модели ассоциативной памяти Дж. Хопфилда. Из всех звуков, участвующих в образовании мелодии и приобретших характер опорных, с которых обычно начинается и оканчивается мелодия, в работе выделяют три наиболее устойчивых типа: тонику, медианту и доминанту.

На рисунке показано резонансное действие трех устойчивых ступеней мажорного и минорного ладов в виде импульсов с частотой ω , поступающих через слуховой нерв.

Из рисунка видно, что основное различие в резонансном воздействии мажорного и минорного лада сводится к воздействию медианты, а именно зависит от ее временного положения. Можно сделать вывод, что процесс восприятия музыки человеком непосредственно зависит от частот звуков, а также от характера их чередования, т.е. порядка следования и скорости смены, другими словами, процесс восприятия музыки обладает контекстной зависимостью.

Компьютерная идентификация музыкального произведения, заданного в виде абстрактного текста – MIDI формате, заключается в формировании на основе музыкального произведения его уникального идентификатора.

Элементы идентификатора (элементарные звуки – ноты) находятся в определенном порядке, с определенными переходами между ними, так как обладают контекстной семантикой.

Целиком уникальный идентификатор однозначно определяет музыкальное произведение, тогда как поиск совпадения только части идентификатора позволяет осуществлять ассоциативный поиск музыкальных произведений, т.е. нестрогий поиск музыкальных произведений схожих с искомым.

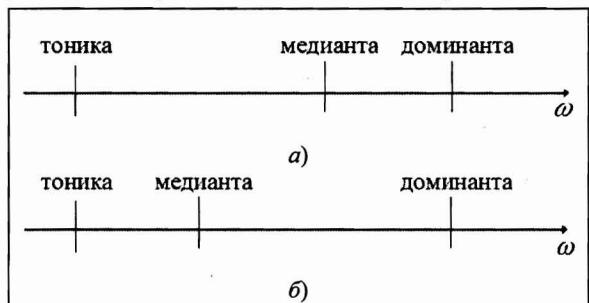


Схема резонансного воздействия устойчивых ступеней мажорного (а) и минорного (б) ладов

Построение уникального идентификатора музыкального произведения

Работа [4] посвящена исследованию общих закономерностей в музыкальных произведениях, заданных в виде последовательности MIDI-событий. Предложенная в ней математическая модель применяется для анализа музыкальных фрагментов с последующим синтезом новых, производных от нее произведений. С этим связана значительная сложность предложенной модели.

Для задачи идентификации и ассоциативного поиска столь точная модель не требуется. В связи с этим возможно использовать в качестве основы предложенную в [4] модель, исключив из нее блоки, отвечающие за синтез новых музыкальных произведений, и оставляя лишь части ответственные за анализ общих закономерностей в музыкальном фрагменте. Таким образом, существующая модель была расширена для решения задачи идентификации и поиска, а также был произведен выбор адекватных структур данных для её реализации.

В рамках полученной модели, музыкальное произведение представляется взвешенным ориентированным графом, в узлах которого расположены аккорды¹, а ориентированные дуги – переходы между аккордами с соответствующими вероятностями. Таким образом, процесс установления эквивалентности между двумя фрагментами музыкальных произведений, заключается, во-первых, в установлении взаимного соответствия множеств вершин графов с последующим нестрогим поиском подграфа в графе. Детальное описание алгоритма дается ниже.

Для идентификации фрагмента произведения, таким образом, необходима структура данных следующего размера:

$$N * T + M * P, \quad (1)$$

где N – количество различных аккордов в произведении; T – размер уникального идентификатора аккорда (список тонов); M – количество переходов между аккордами (дуг графа); P – размер данных, ассоциированных с каждым переходом (значение вероятности).

Алгоритм деструктуризации музыкальной последовательности заключается в следующих шагах:

Шаг 1. Предварительная обработка музыкального фрагмента. Этот этап может включать в себя общие задачи обработки сигналов: очистка от шумов, нормализация и пр.

Шаг 2. Распознавание звукового сигнала, разделение его на типовые составляющие звуки. Фактически этот этап осуществляет преобразование звукового сигнала в символьную форму (WAV → MIDI). Стоит отметить сложность полного такого преобразования. В нашем случае будет достаточно выделения тонов (нот) в сигнале и их длительностей.

Шаг 3. Формирование списка тонов произведения.

Шаг 4. Выделение семантически значимых фрагментов (аккордов) в описании сигнала. Аккорд содержит список звучащих в нем тонов. Новый аккорд начинается в случае, если изменяется состав списка его тонов, т.е. если в текущем аккорде тон прекращает звучать или добавляется новый тон.

Шаг 5. Построение графа переходов между аккордами.

- а) вычисление частотности появления аккордов и вероятности перехода между ними;
- б) вычисление периодичности появления аккордов.

Шаг 6. Поиск более крупных повторяющихся элементов, состоящих из нескольких аккордов.

Шаг 7. Перестроение графа аккордов с учетом новых периодов и вероятностей переходов.

Таким образом, на выходе алгоритма получаем граф переходов между аккордами. Этот график в дальнейшем используется для ассоциативного поиска.

С точки зрения системного анализа, предложенный алгоритм является фильтром, выделяющим из хаотического потока входных данных контекстно-упорядоченную информацию путем декомпозиции с последующим синтезом [8].

В общем случае поиск может осуществляться как по фрагменту оригинального произведения, так и по фрагменту, исполненному пользователем (например, напетому, наспистанному и пр.).

Алгоритм ассоциативного поиска состоит в выполнении тех же самых шагов для поискового фрагмента, как и в алгоритме деструктуризации (1–7). Далее осуществляется нечеткий поиск полу-

¹Аккордом назовем совокупность одновременно звучащих тонов (нот).

енного графа переходов аккордов в графе всего музыкального произведения. При этом учитываются значения вероятностей переходов, а также состав самого аккорда. Здесь необходимо отметить, что в общем случае списки тонов поискового фрагмента и исходного произведения могут различаться. Основная задача состоит в установлении однозначного соответствия между аккордами искомого фрагмента и исходного произведения, так как их состав может быть различным. Выражаясь в терминах семиологической теории – не важен отдельный тон или аккорд, важны их взаимосвязи.

Процедуру установления факта эквивалентности двух музыкальных произведений, по аналогии идентификацией неразличимости Лейбница [8], можно представить следующим образом. Пусть A – список аккордов искомого фрагмента, а аккорды исходного произведения C . Тогда B – аккорд, для которого устанавливается соответствие с аккордом из C . В этом случае, если B раскрывается через элементы X из C (другими словами, если аккорд из C связан с аккордами X , подобным же образом, как B связан с аккордами в A), то считается, что найдено взаимное соответствие для B . С учетом вновь полученных данных эта процедура повторяется для оставшихся аккордов из A . В итоге, если удалось раскрыть аккорды искомого фрагмента, через аккорды исходного произведения, и искомый граф является подграфом в исходном графе, то считается, что искомый фрагмент найден.

Реализация и испытания метода

Непосредственная реализация *шага 1* и *шага 2* описанного алгоритма выходит за рамки данной работы. Разработанная программа работает с базой музыкальных произведений в MIDI формате. База состояла из 50 популярных музыкальных произведений, как классической, так и современной музыки.

Поисковый фрагмент задавался в виде MIDI последовательности, которая выбиралась вручную для каждого музыкального произведения из базы. Ручной выбор поискового фрагмента обусловлен желанием приблизить испытание к реальным условиям работы, когда пользователь выбирает значимый на его взгляд фрагмент произведения.

Испытания предложенного алгоритма проводились в три этапа:

Этап 1. Идентификация музыкального произведения по фрагменту – этот этап включал в себя последовательный выбор всех музыкальных произведений из базы, и для выбранного произведения проводились попытки его идентификации по фрагменту. Вычислялась точность идентификации произведений, а также минимальная длина поискового фрагмента необходимая для успешной идентификации.

Этап 2. Поиск музыкального произведения в базе по поисковому фрагменту – этап состоял в выборе поискового фрагмента из случайно выбранного из базы музыкального произведения, и в поиске произведения по данному фрагменту. В результате подсчитывался процент успешно найденных произведений, и вычислялась минимальная длина поискового фрагмента.

Этап 3. Поиск музыкальных произведений «похожих» на поисковый фрагмент является логическим продолжением предыдущего. Здесь, как и во втором этапе анализировался результат поиска музыкальных произведений по фрагменту. Оценивалась адекватность произведений, найденных методом ассоциативного поиска, поисковому запросу. Необходимо отметить субъективность оценки результата, так как степень схожести двух музыкальных произведений оценивалась человеком-экспертом.

Для вычисления минимальной длины поискового фрагмента на этапах 1 и 2 осуществлялась следующая процедура: в качестве поискового фрагмента выбиралось целиком искомое произведение, далее длина поискового фрагмента уменьшалась до тех пор, пока была возможна идентификация.

В результате испытаний метода на этапе 1 была получена точность идентификации порядка 5%, а минимальная длина поисковой последовательности составила около 10% от длины целого произведения. На втором этапе испытаний была достигнута 80% точность поиска для поисковых фрагментов, а минимальная длина поисковой последовательности составила 15%. Результаты третьего этапа испытаний были оценены экспертами как удовлетворительные.

Заключение

В настоящей работе предложен семиологический подход ассоциативного поиска музыкальных произведений. Для создания полностью адекватной поставленной задаче программы еще предстоит реализовать 1 и 2 шаги предложенного алгоритма, выходящие за рамки данной работы – переход от звуковой формы к текстуальному описанию сигнала. Тем не менее, реализованная функциональность программы позволяет продемонстрировать работоспособность метода. Результаты проведенных испытаний подтверждают теоретические положения эффективности данного подхода.

ЛИТЕРАТУРА

1. Александров В.В., Андреева Н.А., Кулешов С.В. Методы построения информационно-логистических систем. – СПб, 2005.
2. Зайцева А. А. Ассоциативная идентификация // Труды СПИИРАН, т. 1. – СПб: СПИИРАН, 2001.
3. Neil Day, Jose M. Martinez. Introduction to MPEG-7. – International Organisation for Standardisation. ISO/IEC JTC1/SC29/WG11 №3751. La Baule: October 2000.
4. Чубирёв С. В. Исследование математических моделей, разработка алгоритмов и интерфейса программного комплекса обработки звуковых фрагментов в формате MIDI // Диссертация на соискание ученой степени кандидата технических наук Специальность 05.13.18. СПб. – 2006 РГПУ им. Герцена.
5. Peters G., Anthony C., Schwartz M. Song Search and Retrieval by Tapping School of Computing Science, Faculty of Business Administration, and Cognitive Science Program. – Simon Fraser University 8888 University Drive, Burnaby, BC, Canada.
6. Yadgari S. D. Self-similar Synthesis. On the Border Between Sound and Music. – The Media and Sciences Section, School of Architecture and Planning in partial fulfillment of the requirements for the degree of Master of Science at the MIT. September 1992.
7. Евин И. А. Синергетика мозга и синергетика искусства. – Изд-во: Институт компьютерных исследований, Научно-издательский центр «Регулярная и хаотическая динамика», 2005.
8. Александров В. В., Арсентьев А. В. Информация и развивающиеся структуры. – Л.: ЛИИАН, 1984.

Поступила 6 ноября 2007 г.

Computer Searching and Identifying for Music

P. P. Kokorin, K. V. Frolov

In this paper the method for computer identification and search for music are considered. The semiological approach of associative search for music is presented.

The successful practice of semiological approach for associative text search is a basis for application of that method for associative music search. From this view musical signal is presented in MIDI format as abstract text (the sequence of MIDI events). It is considered that MIDI representation of musical signal is sufficient for music authentication hence it represents signal as a sequence of parameters: amplitude, frequency, spectrum, duration which are directly affect on music recognition.

Developed method of music search implies extraction of authentic identifier of musical fragment. This identifier is used for associative search for music composition by its fragment or search for music similar to given fragment.

Despite limitations for input data formats, proposed method of associative music search shows effective results which are in good agreement with theoretical assumptions.