

Определение темпа музыкального произведения методом конкурирующих гипотез.

Курганский Д. А.
dmitry@widisoft.com
Москва, WIDISOFT

Ритмическая структура является одной из важнейших характеристик музыкального произведения. Музыкальное произведение представляет собой последовательность акустических событий, расположенных в узлах достаточно равномерной ритмической сетки (tatum grid) [1]. Важной составляющей задачи распознавания музыки является нахождение узлов этой сетки и определение их метрической роли (сильная или слабая доля).

Алгоритм, позволяющий определять положения музыкальных долей, может применяться не только для нахождения границ тактов, но и для определения точек вероятной смены гармонии [2], а также для предсказания положений последующих нот. В данной работе рассматривается алгоритм, определяющий положения узлов ритмической сетки и границ музыкальных долей. В работе [1] автором был рассмотрен алгоритм определения сетки татумов с помощью функции ошибки остатка, и продемонстрировано, что этот алгоритм применим только для случая квазистационарного темпа. В настоящей работе для решения поставленной задачи используется метод конкурирующих гипотез и показывается его применимость для случаев быстрой смены темпа и сложной ритмической фактуры.

Метод конкурирующих гипотез

В алгоритме производится последовательный анализ ряда, членами которого являются отсчеты времени, соответствующие началам нот. На основе этих отсчетов строятся локальные гипотезы G , содержащие начальное время t_1 , период p и штраф Z . Алгоритм строит последовательности локальных гипотез, создавая дочерние гипотезы от уже существующих. Наследование заключается в том, что начальная точка наследника совпадает с предсказанием родительской гипотезы, а штраф включает в себя штраф родительской гипотезы. Число наследников у одной гипотезы не ограничено, то есть формируется дерево гипотез. В случае, когда две гипотезы совпадают, выбирается имеющая меньший штраф. После анализа всего ряда формируются несколько цепочек гипотез, среди которых выбирается гипотеза с наименьшим штрафом.

Построение цепочки гипотез и присвоение штрафа

Имея в своем распоряжении всего два члена ряда, x_i и x_{i+1} , можно построить гипотезу $G(x_{i+1}, p_i, 0)$, где $p_i = x_{i+1} - x_i$. Гипотеза G предска-

зывает появление следующей ноты в точке $x_{i+2} = x_{i+1} + p_i$. Проверка гипотезы производится сравнением с истинным положением ноты. Предсказание гипотезы считается правильным, если предсказанная величина совпадает с началом ноты, с точностью до достаточно малой величины δp . Эта величина соответствует максимальному интервалу между началами нот, которые слышны как одновременные. В данной работе использовалось значение δp , равное 20 мс.

Если предсказание отличается от реальной ноты на величину, большую δp , то вычисляется ошибка предсказания, и гипотезе присваивается штраф за несовпадение $Z_{unmatch}$, зависящий от этой ошибки. В этом случае от точки x_{i+1} строится вторая гипотеза G_2 , с периодом p_{i2} , правильно предсказывающая поступившую ноту, которая также является наследником гипотезы G . Этой гипотезе присваивается штраф за изменение периода $Z_{irreg}(p_i/p_{i2})$.

Помимо описанных случаев, гипотеза может правильно предсказать не $i + 1$ ноту, а $i + 2$ или более позднюю. В этом случае гипотезе присваивается штраф за пропуск ноты Z_{skip} . Таким образом, штраф каждой гипотезы состоит из следующих слагаемых:

$$Z = Z_{inhert} + Z_{unmatch} + Z_{irreg} + Z_{skip}, \quad (1)$$

где Z_{inhert} — это величина штрафа, унаследованного от родительской гипотезы. Очевидно, что штраф может иметь нулевое значение, только если исходная последовательность времён является строго равномерной. Но и при переменном темпе выбор цепочки гипотез с наименьшим штрафом позволяет построить хорошее приближение сетки татумов из совокупности узлов всех гипотез.

Применение алгоритма

В настоящее время алгоритм реализован в двух версиях, осуществляющих нахождение узлов ритмической сетки в реальном времени, а также для случая предварительно записанной последовательности нот. Использование данных о громкости нот позволяет определить метрическую роль узлов сетки. Алгоритм демонстрирует хорошую устойчивость по отношению к изменениям темпа и сложным ритмическим картинам.

Качество работы алгоритма определяется в существенной мере подбором параметров штрафа, оптимальный набор которых зависит от музыкального жанра. Важно, что алгоритм показывает удовлетворительный результат, даже если исходная последовательность нот содержала существенное количество ошибок, например, была получена в результате автоматического распознавания музыкального произведения.

Учет глобальных характеристик, таких как средний темп и его дисперсия, соотношение общего числа нот и числа найденных узлов, и. т. д.,

позволит улучшить поведение алгоритма в ритмически сложных случаях.

Литература

- [1] *Курганский Д. А.* Анализ ритма и определение темпа музыкальных произведений // Доклады конференции ММРО-12, Москва: Макспресс, 2005. — С. 362–363.
- [2] *Masataka Goto, Yoichi Muraoka* Real-time beat tracking for drumless audio signals: Chord change detection for musical decisions // *Speech Communication*. — 1999. — Т. 1, № 27. — С. 311–335.