

Определение моментов начала нот (онсетов) при анализе музыкальных произведений

Андреевко С. А.

andreenk@widisoft.com

Москва, WIDISOFT

В настоящее время в мире проводятся интенсивные исследования, посвященные созданию автоматических систем анализа музыкальных записей. Такие системы начинают применяться для идентификации музыкальных композиций, контекстного музыкального поиска и контроля над соблюдением авторских прав. Чтобы создать эффективную систему анализа музыки, необходимо решить достаточно большое количество различных задач, таких как анализ ритма, тональности, гармонии, и т. д.

Данная работа посвящена поиску онсетов — моментов начала нот. Мы ограничимся инструментами с хорошо выраженной атакой, но достаточно сложным звуком, например, фортепиано. В записи одновременно с онсетом могут присутствовать частотные компоненты (основной тон и гармоники) уже звучащих нот, кроме того, частотные компоненты не являются непрерывными из-за биений, поэтому поиск онсетов является непростой задачей (Рис. 1).

В работах [1, 2] предлагается разделить спектр звукового сигнала на несколько полос и следить за уровнем и фазой сигнала в каждой полосе. Решение о наличии онсета предлагается принимать на основании вычисления расстояния между векторами амплитуд и фаз в последовательные моменты времени. Наши попытки повторить эту методику не дали ожидаемых результатов, и мы предлагаем другой способ поиска онсетов.

Нейросетевой подход

Сигнал разбивается банком фильтров на достаточно большое число полос, шириной в полутон ($2^{1/12}$ от центральной частоты), при этом центральные частоты соответствуют высотам нот в традиционной равномерно темперированной шкале. Вычисляются отсчеты интенсивности сигнала в каждой полосе через промежуток времени порядка 10 мс, таким способом строится сонограмма, аналогичная изображенной на Рис. 1. Поскольку нота представляет из себя сумму основного тона и гармоник, то в момент начала ноты появляется сразу несколько частотных компонент. Поэтому онсеты определяются в два этапа — сначала определяются те частотные полосы, в которых появились компоненты ноты, а потом определяется нота или ноты, способные породить такой набор частотных компонент.

Для решения этой задачи строится двухуровневая нейронная сеть. Первый уровень определяет появление частотной компоненты на каждой

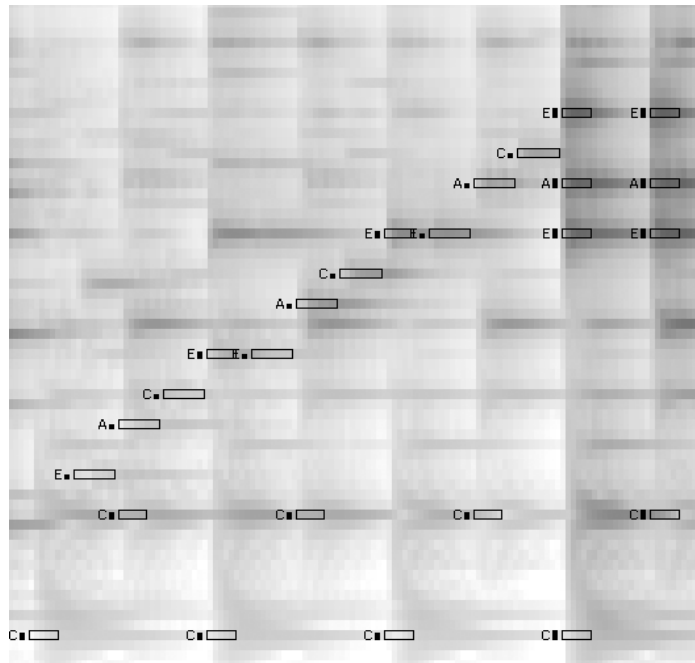


Рис. 1. Фрагмент сонограммы фортепианного произведения.

из полос частот. На вход этого уровня подаются близкие (и по времени, и по частоте) отсчёты сонограммы, т. е. значения из квадрата с центром в текущем моменте времени и полосе частот. Нейросеть обучается на эталонных записях, где положения онсетов известны, методом обратного распространения ошибки. При этом для каждой полосы частот (высоты ноты) строится собственная нейросеть.

Нейросеть второго уровня получает на вход результаты работы всех нейросетей первого уровня в текущий момент времени, и предназначена для определения собственно начал нот. Эта сеть анализирует появление частотных компонент в совокупности и принимает решение о появлении ноты. Данная сеть также позволяет скорректировать ошибки сетей первого уровня, поскольку детектирование появления только одной частотной компоненты с большой вероятностью не означает появления ноты. Это позволяет бороться с эффектом биений, когда возможно периодическое пропадание основного тона ноты. Нейросеть второго уровня также обучается методом обратного распространения ошибки по эталонным записям.

Использование такого механизма позволяет получить достаточно надежное и эффективное детектирование онсетов для фортепианных записей. Полученные таким образом онсетовые моменты могут быть использованы для системы определения ритма и темпа, а также для восстановления нотного текста произведения.

Литература

- [1] *Duxbury C, Sandler M, Davies M* A Hybrid Approach to Musical Note Onset Detection // Proc. of the 5th Int Conference on Digital Audio Effects (DAFx-02), Hamburg, 2002.
- [2] *Duxbury C, Bello J, Sandler M, Davies M* A Combined Phase and Amplitude Based Approach to Onset Detection for Audio Segmentation // Proc. of WIAMIS 2003, Queen Mary: University of London, 2003.