

ИСПОЛЬЗОВАНИЕ АЛГОРИТМА MFCC В ЗАДАЧЕ РАСПОЗНАВАНИЯ МУЗЫКАЛЬНЫХ КОМПОЗИЦИЙ

*Петровец В.Н., магистрант гр.256241, Мискевич П.Л., магистрант гр.256241,
Деменковец Д.В., аспирант*

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Деменковец Д.В. – старший преподаватель, аспирант

В данной работе описан алгоритм нахождения Mel-частотного кепстра и применения его в задаче идентификации индивидуальной музыкальной композиции. Также предложен алгоритм автоматизированного распознавания звуковых сигналов с использованием MFCC.

Распознавание музыки или возможность автоматически идентифицировать и классифицировать музыкальный контент в последние годы привлекли значительное внимание из-за распространения цифровой музыки и необходимости эффективного управления музыкой, рекомендаций и защиты авторских прав. Одним из ключевых методов, используемых в процессе распознавания музыки, является алгоритм кепстральных коэффициентов Mel-частоты (MFCC), который получил широкое распространение благодаря своей эффективности в захвате акустических характеристик музыкальных сигналов. Алгоритм MFCC основан на свойствах восприятия слуховой системы человека и применяется в различных задачах распознавания музыки, таких как классификация жанров, определение настроения и распознавание инструментов. Однако, несмотря на его популярность, все еще существует потребность во всесторонней оценке эффективности алгоритма MFCC в распознавании музыки, включая его сильные стороны, ограничения и потенциальные области для улучшения.

Исследования показали, что функции MFCC, извлеченные из аудиосигналов, могут эффективно представлять тембровые характеристики музыки, что имеет решающее значение для таких задач, как классификация жанров и определение настроения. MFCC также использовался в сочетании с другими методами, такими как скрытые марковские модели (HMM) и машины опорных векторов (SVM), для достижения высокой точности в задачах распознавания музыки. Однако следует также учитывать определенные ограничения алгоритма MFCC. Одним из основных ограничений является его чувствительность к шуму и изменчивости музыкальных сигналов, что может повлиять на точность и надежность системы распознавания. Кроме того, алгоритм MFCC может не захватывать музыкальные особенности более высокого уровня, такие как мелодическая или гармоническая информация, которые могут быть важны для определенных задач распознавания музыки. Кроме того, на производительность алгоритма MFCC может влиять выбор параметров, таких как количество кепстральных коэффициентов, размер набора фильтров Mel, а также длина кадра и перекрытие, используемые в процессе выделения признаков.

Особенность использования алгоритма MFCC для анализа музыкальных композиций заключается в термине Mel-частота. Mel – единица высоты звука, основанная на восприятии этого звука органами слуха человека. Амплитудно-частотная характеристика человеческого слухового аппарата нелинейна, как показывают исследования (рисунок 1).

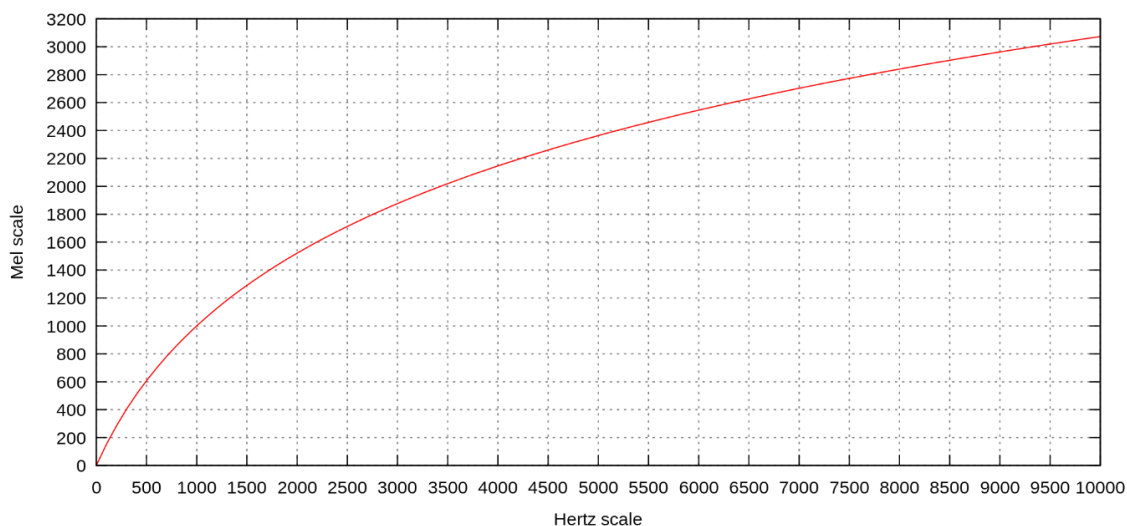


Рисунок 1 – Зависимость единицы измерения Mel-частот от Герц

Далее приведено описание алгоритма MFCC для участка музыкальной композиции:

1. Преобразование звукового кадра из временной в частотную область с помощью оконного преобразования Фурье.
2. Отображение частотных бинов на Mel-шкалу, используя треугольные окна с наложением
3. Расчет логарифмов отображенных значений Mel-диапазонов
4. Применение дискретного косинусного преобразования (DCT) к Mel-диапазонам.

Параметрами алгоритма являются: количество Mel-диапазонов (в литературе встречаются значения от 20 до 80), кол-во коэффициентов DCT (10-15), размер окна, коэффициент наложения окон.

Результат работы алгоритма MFCC – двумерная матрица, хранящая в себе векторы признаков для каждого окна исходного сигнала. Впоследствии набор данных векторов может использоваться для классификации исходной звуковой композиции.

Для использования алгоритма MFCC в процессе распознавания музыки необходимо выполнить следующий порядок действий:

1. Подготовка набора данных. Необходимо собрать разнообразный набор данных музыкальных сигналов, включающий различные жанры, настроения и инструменты, чтобы обеспечить всестороннюю оценку производительности алгоритма MFCC. Набор данных будет предварительно обработан для удаления любых артефактов (шума), нормализации уровней звука и разделен на наборы для обучения и тестирования.

2. Извлечение признаков. Алгоритм MFCC будет реализован для извлечения признаков из музыкальных сигналов в наборе данных. Различные конфигурации параметров, перечисленные при описании MFCC, должны быть протестированы для изучения их влияния на производительность алгоритма.

3. Распознавание музыки. Значения коэффициентов MFCC, извлеченные из аудиосигналов, используются в качестве входных данных для классификатора машинного обучения, такого как машина опорных векторов (SVM) или нейронная сеть, которая обучается на помеченном наборе данных музыкальных треков.

4. Оценка эффективности. Производительность алгоритма MFCC в задачах распознавания музыки должна оцениваться с использованием соответствующих показателей, таких как точность, воспроизводимость и оценка F1. Результаты должны быть проанализированы статистически, чтобы выявить любые существенные различия в производительности между алгоритмом MFCC и другими методами извлечения признаков.

MFCC широко используется в цифровой обработке сигналов для получения информации о звуковых композициях. Кроме того, данный алгоритм может быть использован в обработке изображений для аннулирования оттенков серого черно-белых изображений. Извлеченная из короткого окна, совокупность MFCC коэффициентов воспринимается как тембральный дескриптор звукового сигнала. В данной работе был предложен и описан подход для использования алгоритма MFCC в процессе анализа и индивидуального распознавания музыкальных композиций.

Список использованных источников:

1. Li, T.L.H. *Genre classification and the invariance of MFCC features to Key and Tempo* / T.L.H. Li, A.B. Chan // *International Conference on MultiMedia Modeling, Taipei, 2011*.
2. Prabakaran, D. *Speech Processing: MFCC Based Feature Extraction Techniques- An Investigation* / D. Prabakaran, S. Sriuppili // *Journal of Physics: Conference Series / IOP Publishing – London, 2020*.
3. Analytics Vidhya [Electronic resource] : *MFCC Technique for Speech Recognition*. – Mode of access: <https://www.analyticsvidhya.com/blog/2021/06/mfcc-technique-for-speech-recognition/>. – Date of access: 24.03.2023.
4. Gupta, S. *Feature extraction using MFCC* / S. Gupta [at et.] // *Signal & Image Processing : An International Journal (SIPIJ) / Indian institute of Information and Technology – Allahabad, 2013*.