

Sygnały akustyczne - Projekt zaliczeniowy

Lukasz Smalec
Nr albumu: 39433

Robert Piątek
Nr albumu: 39402

15 czerwca 2021

1 Założenia i realizacja projektu

Projekt zakładał napisanie skryptu w języku *Python*, pozwalający na dokonanie klasyfikacji utworu na podstawie jego gatunku. Jako metode klasyfikacji wykorzystano algorytm *SVM* zaimplementowany z wykorzystaniem biblioteki *sklearn*. Zrealizowany projekt pozwala na klasyfikację utworu jako jeden z czterech gatunków:

- Rock,
- Pop,
- Hiphop,
- Metal.

Do nauki i testowania klasyfikatora wykorzystano 30-sekundowe fragmenty utworów ze zbioru danych GTZAN Dataset.

2 Podział pracy

Łukasz Smalec

- Znalezienie i przygotowanie zbioru danych.
- Praca nad ekstraktorami cech.
- Przygotowanie statystyk działania skryptu.
- Zapis klasyfikatora i wyników do pliku.

Robert Piątek

- Znalezienie i przygotowanie zbioru uczącego.
- Praca nad ekstraktorami cech.
- Wstępne przetwarzanie danych.
- Klasyfikacja.

3 Wyniki

3.1 Klasyfikator binarny

Wstępnie zbudowano klasyfikator na 100 utworach z gatunku Rock oraz na mieszance 180 utworów o gatunkach innych niż Rock. Utwory zostały podzielone, 33% na zbiór testowy i 67% na zbiór uczący.

Tablica 1: Wyniki klasyfikacji binarnej (czy utwór należy do gatunku Rock, czy nie).

Metoda ekstrakcji cech	Poprawność klasyfikacji
MFCC	66.67%
FFT	78.49%
Spectral Centroid	65.59%

Wyniki przedstawione w tabeli 1 wskazują, że najlepszą metodą jest *FFT*, zaś *MFCC* i *Spectral Centroid* zwróciły podobne wyniki trochę gorsze od *FFT*. Należy jednak wspomnieć o tym, że *FFT* wykonuje się wielokrotnie dłużej od poprzednich metod.

3.2 Klasyfikator wieloklasowy — pierwsza wersja

Przy pierwszym podejściu w budowie klasyfikatora wieloklasowego wykorzystano 100 utworów z każdego z czterech wybranych gatunków. Utwory zostały podzielone tak samo jak wcześniej, 33% na zbiór testowy i 67% na zbiór uczący.

Tablica 2: Wyniki klasyfikacji wieloklasowej.

Metoda ekstrakcji cech	Czas wykonania	Poprawność klasyfikacji
MFCC	20.50s	66.23%
FFT	141.14s	75.76%
Spectral Centroid	36.52s	36.36%

Wyniki ukazane w tabeli 2 są zadowalające. Poziom poprawności klasyfikacji spadł jedynie dla ekstraktora *Spectral Centroid*. Co ciekawe, *FFT* utrzymało najlepszy wynik, jednak z czasem aż 7-krotnie dłuższym niż *MFCC*. Z tego powodu zdecydowano się na dalsze badania w celu podniesienia poziomu klasyfikacji *MFCC*.

3.3 Klasyfikator wieloklasowy — badania nad podniesieniem dokładności *MFCC*

W celu podniesienia poprawności klasyfikacji dla cech MFCC, badano wpływ zmiany parametrów na wyniki. Zdecydowano się zmienić ilość generowanych cech, gdyż ten parametr miał największe znaczenie.

Tablica 3: Wyniki badań w celu zwiększenia poprawności klasyfikacji na podstawie cech MFCC.

Ilość generowanych cech MFCC	Czas wykonania	Poprawność klasyfikacji
25 000	20.93s	66.67%
50 000	22.07s	68.94%
100 000	23.64s	72.73%
150 000	26.43s	71.97%

Dzięki badaniom przedstawionym w tabeli 3 udało się podnieść poprawność klasyfikacji do 72.73% przy 100 000 generowanych cech. Im więcej cech tym wyniki były lepsze, jednak próg 100 000 jest optymalny, gdyż po nim poprawność spada. Wynik jest bardzo zbliżony do poprawności *FFT* — 75.76%. Biorąc pod uwagę zdecydowanie krótszy czas wykonania przy cechach *MFCC*, wydaje się to lepsze rozwiązanie jeśli zależy na optymalności.

4 Szczegółowe wyniki klasyfikacji wieloklasowej

Tablica 4: Szczegółowe wyniki badań dla każdego gatunku.

Metoda	Gatunek	Poprawność klasyfikacji	Łączny czas wykonania
MFCC	Rock	81.25%	24.07s
	Hiphop	48.72%	
	Pop	87.10%	
	Metal	80.00%	
FFT	Rock	81.25%	145.00s
	Hiphop	61.54%	
	Pop	80.65%	
	Metal	83.33%	
Spectral Centroid	Rock	43.75%	30.10s
	Hiphop	12.82%	
	Pop	32.26%	
	Metal	63.33%	

Jak można zauważyć w tabeli 4 klasyfikator miał największy problem z gatunkiem Hiphop przy każdym rodzaju cech. Co ciekawe cechy *MFCC* poradziły sobie najlepiej z gatunkiem Pop, zaś *FFT* z gatunkiem Metal.

5 Podsumowanie i przyszłość pracy

Najlepszą cechą do klasyfikacji okazało się *FFT* biorąc pod uwagę poprawność klasyfikacji, jednak pod względem optymalizacji klasyfikacji znacznie lepszą cechą jest *MFCC*. Cecha *Spectral Centroid* okazała się najgorsza, co zostało szczególnie mocno wykazane przy klasyfikacji wieloklasowej.

Dalsze badania nad pracą mogłyby doprowadzić do ciekawych wyników. Można przetestować różne inne klasyfikatory, czy też spróbować innych cech do ekstrakcji z utworów. Warto też przeprowadzić szczegółowe badania co do parametrów pozostałych ekstraktorów cech.