

Content-based and User Customized Music Recommendation System

Vitor Galassi Luquezi*, Tiago Fernandes Tavares

Abstract

In this work was developed a music recommendation system that automatically suggests tracks for an users's playlist by finding tracks in a personal music collection that have similar timbral characteristics to those of the playlist's tracks, using machine learning and expectation maximization algorithms. First, each track of the collection goes through a feature extraction process, which uses signal processing techniques to extract low level psychoacoustic inspired features from a track file and join them as a feature vector. Then, using a Gaussian Mixture Model, it is possible to calculate an optimal gaussian model that would generate the playlist tracks' feature vectors. Finally, this model is used for verifying the likelihood of each database tracks' feature vector by finding which ones have the highest chance of being generated by the model, and using them as a recommendation for the user's playlist.

Key words:

Digital signal processing, recommendation system, automatic audio signal classification.

Introduction

Recommendation systems are widely used nowadays in a variety of contexts, such as virtual stores, music or film streaming services, social networks, news platforms. The goal of this work is the development of a simple music recommendation system that could be used for organizing a personal collection of songs or to provide a user a better music listening experience by suggesting music that are similar to a playlist previously made by him.

The approach used to accomplish this objective was based on the automatic musical genre classification of audio files once proposed by George Tzanetakis and Perry Cook¹. Similarity between audio signals can be found by comparing (using machine learning algorithms) the feature vector that represents each signal. The features used are related with the overall timbral texture of the track (as they were psychoacoustically inspired), considering well known features of the literature, like mean an variance of the spectral centroid, spectral rolloff, spectral flux, spectral flatness, energy, Mel-frequency cepstral coefficients and zero-crossing rate over a time frame, and their first and second derivatives as well.

Results and Discussion

First, it is relevant to present some information related to the tracks database used for this current work. The database was taken from personal collections of ordinary users and consists of 500 songs, being 100 songs of each genre: rock, pop, jazz, electronic and folk. It is known that music genre classification is a complex and subjective task, but for evaluation purposes, it was necessary to label each track based on their most common classification found over internet sources, and often without regard to which musical subgenres they belong to (e.g. a pop track with rock traditional characteristics).

For the evaluation of the recommendation system, a playlist model has been proposed considering all the 100 tracks of a single genre a time, and splitting them into five groups, being four of them used for "train" and one for "test", and then using another one for "test" in each iteration. Therefore, for each of the five iterations, 80 tracks were used for obtaining a GMM model of the system and the 420 remaining tracks for testing (20 being the "answer"

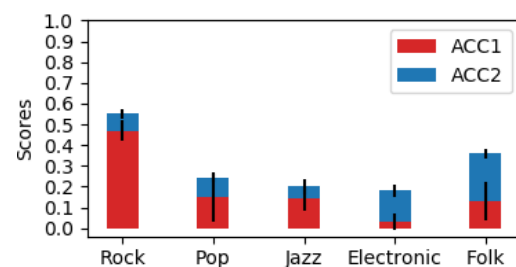
- chosen genre - and the remaining 400 of the database - other genres).

The duty of the system lies on "guessing" (i.e. recommending) the 20 answer tracks between the 420 group of the rest of the database, thus obtaining two evaluation metrics based on common ones found on literature: *Accuracy 1* (ACC1) and *Accuracy 2* (ACC2). These metrics were defined as below:

Accuracy 1: # of tracks correctly suggested (of the 20 tracks in the "answer group", over the total recommended tracks.

Accuracy 2: # of tracks that would be a reasonable recommendation (and not included on the "answer group"), over the total recommended tracks.

Image 1. Results of the recommendation system



Conclusions

The content-based recommendation system had a worst-case accuracy of 20%, whereas a random recommendation is expected to yield 2.5% accuracy. Interestingly, the accuracy for each genre is different. Further investigation on this matter is a possible direction for future work.

Acknowledgement

This scientific research is part of the PIBIC 2016/2017, sponsored by CNPq and SAE UNICAMP.

¹ Tzanetakis, G.; Cook, P.; Musical Genre Classification of Audio Signals. *IEEE Transactions on Speech and Audio Processing* **2002**, 293-302.