Music Information Retrieval Using Random Forest Classifier

by

Chibueze Ukachi

A **Semester Project** submitted

in the

School of Electrical Engineering

Signal Processing Laboratory LTS2

Assisted by: Michaël Defferrard

Supervised by: Professor Pierre Vandergheynst

June 2017

Acknowledgements

Thank you to Professor Pierre Vandergheynst for access to the Signal Processing Laboratory LTS2 and its facilities. Also, thank you to the many artists who released their Music License-Free and to the Free Music Archive team for collating all these songs into a single easily accessible platform.

Above all, thank you to Michaël Defferrard (Doctoral Assistant at EPFL) for your persistent support with this project. My working pattern can sometimes be more like an alternating current than a direct current, so I thank you for being really patient through the lows and extremely supportive through the highs. It has been my greatest academic pleasure and honour to work with you.

Abstract

This project covers the implementation of Music Genre Classifcation using the Free Music Archive (FMA) Dataset. The FMA is a free, open and readily available Dataset. It was chosen because it has a sizeable song hub of over 80,000 songs, all of which are available under the non-restrictive Creative Commons Licence.

We will be implementing the Random Forest Classifier as the Base Case Machine Learning algorithm for Genre Classification because each tree is created using a random subset of the training data and also a random subset of the input features. It is therefore less susceptible to misclassifcation of individual songs.

We propose the extraction of Time and Frequency Domain Features such as Zero Crossing Rate and Spectral Centroid, as well as the computation of different statistical moments to predict the Genres of individual songs.

List of Figures

Contents

1 Introduction

Music Information Retrieval (MIR) is an interdisciplinary field that ranges from Musicology to Machine Learning. The goal of MIR is to computationally extract information that can be used to classify, modify and even generate music. This has led to growing interests in the field in recent times as the computation power has increased whilst becoming more affordable to the masses.

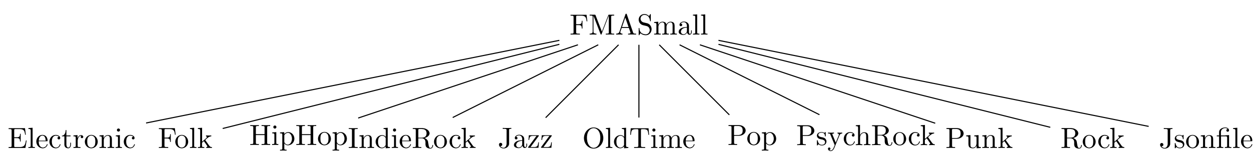
The aim of this project is to computationally extract features of a song that carries information which are useful for detecting different Song Genres. We use the 4000 songs out of 80,000 songs available from the License-Free Free Music Archive Dataset. FMA releases numerous metadata such song title and plays counts. Nonetheless, we restrict the usage of metadata, and only retrieve the Genre of individual songs, so that we can directly see the effects of changing or using different feature combinations on the Genre Classification.

2 Free Music Archive Dataset

The Free Music Archive is a free and open Dataset released under the non-restrictive Creative Commons Licence. Although it has a sizeable collection of 80,000 songs, this project uses a subset of 4000 songs that have a balanced genre distribution.

2.1 File Structure

The FMASmall folder holds 3.8 gigabytes comprising of 4000 songs. The tree diagram below shows the file structure for each Genre Folder and the metadata is stored in the JSON file.



2.2 File Distribution

The data set is a genre-balanced with each music genre containing 400 songs. This is shown clearly in Figure 1 below.

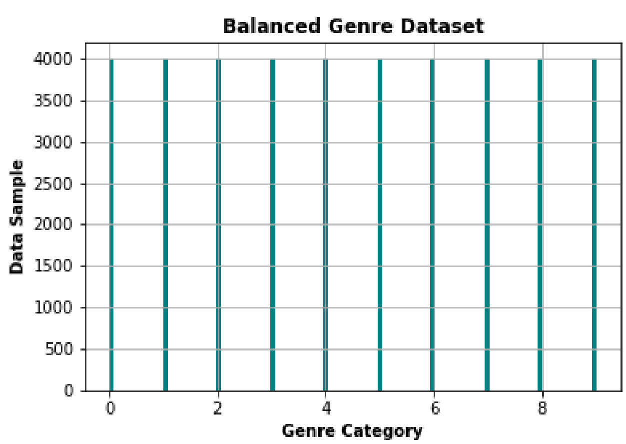


Figure 1: The distribution of genres in the fma small dataset

3 Music Information Retrieval and Genre Recognition

4 Methodology

4.1 Song Retrieval

The song extraction process was performed so the each song could be located completely independent of the users operating system.

- Each song was located via its full path in order to guarantee one-to-one mapping and accurately retrieve songs. This system was also used to retrieve the genre of each song as provided by the Json meta-data file.

- Each song was stripped to 30 second slices so that all 4,000 songs had uniform length. The 4,000 songs of 30seconds were further split into 40,000 songs of 3 seconds. This provided an easy way to increase the dataset which further increases the accuracy of machine learning algorithms. The rationale was that each 3 second clip could be viewed as an individual song and any relevant features detected on the smaller clip should be useful for classifying the genre of the song. This classification would be performed via a voting system where the most frequent genre in each group of ten 3 second clips that constituted a song would be classed as the genre of that song.

4.2 Feature Selection

For this project we retrieved two Time and Frequency Domain Features and computed some statistical moments such as mean, variance, skew and kurtosis.

4.2.1 Time Domain Features

- Zero Crossing Rate
- Onsets

4.2.2 Frequency Domain Features

- Mel Frequency Cepstral Coefficients - Split the song into smaller frames. The assumption is that the signal is statistically stationary (moments like mean, variance are constant over time) during the frame .(s) - Take the power spectrum of each frame to identify its frequency contents - Mel Filterbanks are used to ascern the difference between two closely spaced frequencies. These are the filterbank energies. - Take the logarithm of the filterbank energies as the human ear hears sound logarithmically rather than linearly - Calculate the Cepstral Coefficients by taking the Discrete Cosine Transform (DCT) of the log filter bank energies. - Calculate statistical moments such as mean, variance, skew and kurtosis on each song segment
- Spectral Centroid

4.3 Machine Learning

brief overview of algorithms used

5 Experiments

The different approaches that I tried.

6 Results

All tables, figures and explanations

7 Limitations and Future Improvements

Drawbacks of the strategy and things that can be done better

8 Conclusion

Summary and overview. Also, what I think was the most effective strategy
Complete bibliography *ref1*[1]

[1]http://links to refenrences

9 Bibliography

Here is the info [1]

References

[1] D. Adams. *The Hitchhiker's Guide to the Galaxy.* San Val, 1995.