MUSIC GENRE CLASSIFICATION

Ayush Kumar[1], Bojja Krishna Sai Siva[2], G Sumin Reddy[3] , M.R.Rashmi[4]

[1]Ayush Kumar, Department of Computer Science, The National Institute of Engineering, Mysore, Karnataka , India
[2]Bojja Krishna Sai Siva, Department of Computer Science, The National Institute of Engineering, Mysore, Karnataka , India
[3]G Sumin Reddy, Department of Computer Science, The National Institute of Engineering, Mysore, Karnataka , India
[4]M.R.Rashmi, Assistant Professor, Department of Computer Science, The National Institute of Engineering, Mysore, Karnataka , India

---

Abstract – I*n this paper , we have put forth a music genre classification approach based on Mel Frequency Cepstral Coefficients (MFCC) . The Mel Frequency Cepstrum (MFC) encodes the power spectrum of a sound. It is calculated as the Fourier transform of the logarithm of the signal's spectrum. A modulation spectrogram is used corresponding to the collection of modulation spectra of Mel Frequency Cepstral Coefficients (MFCC) will be constructed. A spectrogram is a visual representation of the frequency content in a song. It shows the intensity of the frequencies on the y axis in the specified time intervals on the x axis; that is, the darker the color, the stronger the frequency is in the particular time window of the song. Each genre existing as of now is carefully designed with a particular number of prototype vectors which were designed with appropriate algorithms. An information fusion approach comprising of both feature and decision level fusion is involved for the appropriate outcome. Recently there was a development which saw a million song dataset being released which poised all the song features and its metadata. But , we find that our method comprising of the salient features listed above improves the accuracy for detecting genre of a music file .*

Key Words: Genre Classification, Mel Frequency Cepstral Coefficient, Mel Frequency Cepstrum, Modulation Spectrogram .

1.INTRODUCTION

The first thing that strikes one's mind with the heading is that what is a music genre? Music genre can be defined as a category or rather conventional category that recognises the characteristics or traits of sub-division of the music file belonging to a traditional or any conventional established music form

The term Music Genre Classification can be explained as categorising of music samples. A music genre classifier plays a vital role in adjudging song samples in a preliminary stage, for instance if a fresh song has been recorded it will help in categorising the song into its conventional category.

To determine the genre of a song it has to be distinguished by its unique audio features so that its contents can be analysed with respect to the produced wave signals. Another important aspect is the recognition of the instruments used in the song which is also known as the timbral characteristics. This plays a very vital role in specifying the music type based on the type of instrument helps us denote the traditional connect to the music.

In this paper, we have used the Mel Frequency Cepstrum Coefficient (MFCC) for encoding the power spectrum of the sound with the calculation of the Fourier transform of the logarithm of the signal's spectrum. Another important role player here is the spectrogram which helps us in the visual representation frequency content in a song. It shows the intensity of the frequencies on the y axis compared to the time interval on the x axis. This makes the job easier in predicting the potential genre of the music file thus fulfilling our search .

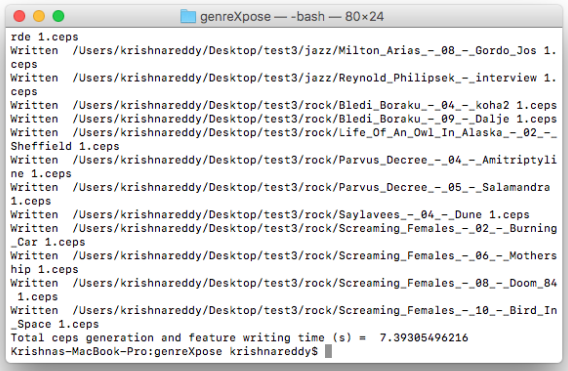
1.1 Proposed System

The Proposed system makes use of Mel Frequency Cepstral Coefficients (MFCC) and spectrogram, thus to illustrate it further the system is divided into two steps as follows :

i. Ceps construction phase

ii. Genre classification phase

i. *Ceps construction phase*- In this phase we use a python script which helps us to analyze and convert each file from the data-set in a representation that can be used by the classifier and be easily cached on to the disk. This little step prevents the classifier to convert the dataset each time the system is run.



ii. *Genre Classification Phase*- In this phase a dataset is used for feeding the data in the classifier, which creates a memory model within itself stated as regression model. This process is done by the Logistic Regression module of the scikit-learn library. The python script for this purpose is stated. Once the model has been created, we can use it to predict genres of other audio files. For efficient further use of the generated model, it is permanently serialized to the disk, and is de-serialized when it needs to be used again. This simple process improves performance greatly. As of now, the python script has to operate before any testing with unknown audio file can be performed. Once the script is run, it will save the generated model. Once the model has been successfully saved, the classification script need not be run again until some newly labeled training data is available.

2. Model Generation

A dataset is used for training the classifier, which generates an in-memory regression model. This process is done by the Logistic Regression module of the scikit-learn library. The python script has been provided for this purpose. Once the model has been generated, we can use it to predict genres of other audio files. For efficient further

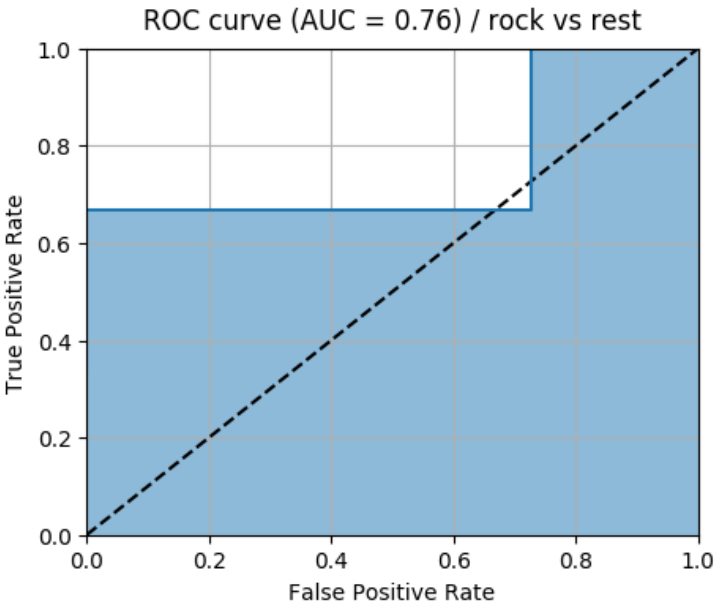
use of the generated model, it is permanently serialized to the disk, and is de-serialized when it needs to be used again. This simple process improves performance greatly. As of now, the python script must be run before any testing with unknown music can be done. Once the script is run, it will save the generated model. Once the model has been successfully saved, the classification script need not be run again until some newly labeled training data is available.

Few other steps in this process include the following:

i. *Testing*- A python script is used for checking up on new and fresh audio files and it helps in de-serializing the previously cached models. Thus, it labels the new files.

ii. *Output Interpreter*- All music files are classified and its trained model is saved to the disk. Also, graphs are generated which are saved in the directory.

iii. *ROC Curves*- The Receiver Operating Characteristic Curves are generated and saved which denotes the truthfulness of the defined genre after the music file is classified.



A few *Test Cases* have also been performed and can be listed as follows:

TC#	Description	Expected Input	Expected Output	Status of execution pass/fail
TC#1	Executing	.mp3 file	IOError: [Errno 2] No such file or directory: '/Users/XYZ/Desktop/test/1.mp3'	Fail
TC#2	Admin login	Admin enters the ID and password	The verification of ID and password takes place. If its correct then he can perform	Pass

			the assigned work.	
TC#3	Executing	.wav file	predicted genre = blues/rock/jazz	Pass
TC#4	Directory scan	/Users/XYZ/Desktop/dataset	IOError: [Errno 2] No such file or directory	Fail
TC#5	Runtime	Python tester.py	IOError: [Errno 2] Invalid numpy version	Fail

Table -1: Test Cases Table

3. CONCLUSIONS

This approach to classify a music genre can be termed as one which needs very few computations and does not even require much of any data. But, the result speak otherwise with respect to the data, it occurs to us that a large quantity of data needs to be provided and trained from its preliminary stage of the classification process.

Although this process is fast but a lot of optimization is required. Alas, we can say that in future more refinements to this case is possible.

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible whose constant guidance and encouragement crown all efforts. First and fore most we would like to thank our beloved principal Dr.G.L.Shekhar for being the patron and the beacon light for this project. We would like to express our sincere gratitude to our HOD Dr. H. D. PHANEENDRA, HOD, Department of CSE, NIE for his relentless support and encouragement. It gives us immense pleasure to thank our guide Mrs. M. R. RASHMI, Assistant Professor ,Department of CSE, NIE for their valuable suggestions and guidance during the process of the seminar and for having permitted us to pursue work on the subject.

REFERENCES

[1] Omar Diab, Anthony Manero, and Reid Watson. Musical Genre Tag Classification With Curated and Crowdsourced Datasets. Stanford University, Computer Science, 1 edition, 2012.

[2] N. Scaringella, G. Zoia, and D. Mlynek. Automatic genre classification of music content: a survey. IEEE Signal Process. Mag., 23(2):133–141, 2006.

[3] Sox.sourceforge.net. Sox - sound exchange — homepage, 2015.

[4] Bob L. Sturm. Classification accuracy is not enough. Journal of Intelligent Information Systems, 41(3):371–406, 2013.

[5] G. Tzanetakis and P. Cook. Musical genre classification of audio signals. IEEE Transactions on Speech and Audio Processing, 10(5):293–302, 2002.

[6] Jan W¨ulfing and Martin Riedmiller. Unsupervised learning of local features for music classification. In ISMIR, pages 139–144, 2012.

[7] Changsheng Xu, MC Maddage, Xi Shao, Fang Cao, and Qi Tian. Musical genre classification using support vector machines. In Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on, volume 5, pages V–429. IEEE, 2003.

[8] Music Genre Classification Using Modulation

[9] Spectral Features and Multiple Prototype Vectors Representation by Chang-Hsing Lee, Chih-Hsun Chou, Cheng-Chang Lien, and Jen-Cheng Fang Department of Computer Science and Information Engineering Chung Hua University, Hsinchu, Taiwan.