MUSICGENRECLASSIFICATION

MS.N.Harika (Asst.Professor), Computer Science and Engineering Department, VNITSW, Guntur, Andhrapradesh, India.

K. Lakshmi Prasanna, P.RajyaLakshmi ,Ch.Sireesha, M A Shabreentaj, IV BTech,Department Of Computer Science And Engineering,Vignan'sNirula Institute Of Technology And Science For Women,Peda Palakaluru,Guntur-522009,AndhraPradesh,India

ABSTRACT

Music genre is getting complex from time to time. As the size of digital media grows along with the amount of data, manual search of digital audio files according to its genre is considered impractical and inefficient; therefore a classification mechanism is needed to improve searching. Mel-Frequency cepstral coefficients (MFCC) and Covariance are a few of features that can be extracted from digital audio files to classify its genre. This research is conducted to classify music from digital audio (songs) into 10 genres: Blues, Classic, Hip Hop, Jazz, Pop, country, disco, metal, pop, reggae and Rock using above mentioned features, extracted from WAV audio files. Classification is performed several times using selected 3, 6, 9 and 10 genres respectively. The result shows that classification of 3 music genres (Country, Blues, Classic) has the highest accuracy (96.67%), followed by 6 genres (Country, Blues, Classic, Metal, Hip Hop, Jazz) with 70%, and 9 genres (Country, Blues, Classic, Metal, Hip Hop, Jazz, reggae, Pop) with 53.33% accuracy. Classification of all 10 music genres yields the lowest accuracy of 65.33%. The test results with the k-Nearest Neighbours algorithm to 120 songs for k = 3 accuracy reaches 22.5%, k = 5 accuracy reaches 22.5%, k = 7 accuracy reaching 26.7% and k = 9 accuracy reaches 26.7 %. Results showed that genre classification by matching the shortest distance through the centre of the class, yields better results than using the k-NN algorithm.

Keywords- Feature Extraction, Music, Hip, Classification, MFCC

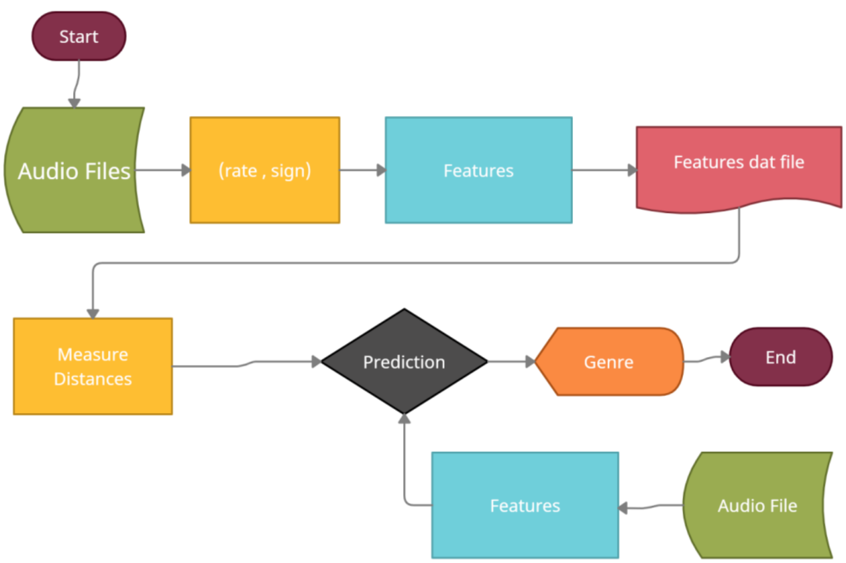
INTRODUCTION

Digital audio is now an integrated part of our lives. Many of us are not using analogue music players such as tape recorders anymore, and sounds generated in our everyday lives are now from digital sources. This switch into digital audio inevitably affects how we render music for listening. Much music now are available in digital format, which brings us the versatility of having a collection of many albums without having to worry about storing the physical media. This also benefits the radio stations in having a vast collection of music which may please more ears of their listeners. Traditionally, analogue music is digitised into a waveform (known as WAV files) of a certain bitrate for desired quality. Daily listeners are known to fancy having music in MP3 or other formats preferred for their file size, as the compression cuts many of the bytes but not so many details from a music file. However, many of the original digital copies of music are still stored in WAV format because up to now it still represents the highest fidelity compared to other formats available, and it is uncompressed. The problem of having such a wide range of collections, especially for radio stations, is how to classify the music according to its genre. Traditional ways of having CDs stacked into classified racks worked well also with files, but it was when the collections were not so many. Labelling the collection is also not a good option, since searching for music files of the same genre is an exhausting process. An automated process is in need to help with this problem. This research proposes that music genres can be classified by having a music piece sliced into certain lengths of time, extract distinguishing features from the slices which altogether formed a feature pattern for each music title, and classify the pattern. The patterns mentioned here are one or more patterns which are easily recognisable from the music's amplitude such as beat and flow, by having in mind that the most prominent pattern from a music genre is its beat and flow. Moreover, this research would also like to find whether those features are good enough to distinguish music genres. Titles having the same genre are supposed to have the same feature pattern. The idea is derived from feature classification in recognising speech and music – which basically detects a different flow between speech and music.

LITERATURE SURVEY

Several works and researches have been conducted on classifying genres of various data [1]. There were audio data classification using symbolic classification which transformed symbolic representation of audio signals into harmony using chord transcription algorithm [2].It willclassify music harmony into different pitch profiles [3]. A work focusing on document genres which automatically classified web pages according to their text genres has been researched. This research is known to emphasise genre annotation and granularity to improve accuracy of document classification [4]. Another research has been conducted to classify music genres using representation of what humans hear through temporal modulation of a strong music genre. The proposed classification explores the psycho-physiological aspect of hearing a musical record through temporal modulation and network-based classification. Reduced linear subspace learning used in this research has proven to be decisive for the research framework [5]. An approach using window of 256 samples in which each are smoothed, transformed using FFT and using 50 FFT vectors as classification features which then classified using Linear Vector Quantization (LVQ) networks have also been proposed to help radio stations and music TV stations in classifying their collection [6]. Tensor-based automated music genre classification has been proposed using mathematical models of Non-Negative Tensor Factorisation (NTF). From the research, it can be shown that the performance gains for the NTF classifier are not statistically significant against the SVM and MLP classifiers at 95% confidence level [7]. Another music genre classification are using musical surface feature calculation based on the Short Time Fourier Transform (STFT) that can be efficiently calculated using Fast Fourier Transform (FFT). The features are centroid, roll off, flux, zero crossings, and low energy. These features are added by rhythm features such as period, and amplitude [8].

PROPOSED SYSTEM



The first step of the process is data preparation of each audio file is converted into Mel-Frequency Cepstral Coefficients (MFCC) and Covariance . The features are pickles and dumped into a “dat” format.

Which is machine independent can be used wherever we like to use it. The frequency features of the audio streams are extracted. We consider Mel Frequency Cepstral Coefficients (MFCC) for dimensionality reduction.

Classification is to put objects according to the groups they belong [10]. A nearest centroid classifier or nearest prototype classifier is a classification model that assigns the label of a class of training samples whose centroid (mean) has the closest distance to the observed object or data. Distance can be calculated using Euclidean distance as of equation (4).

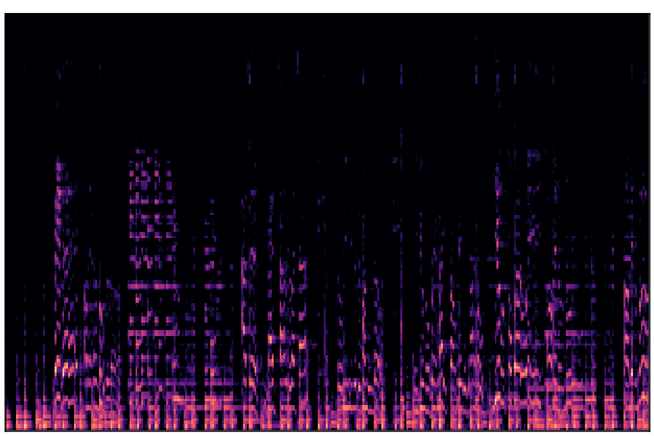
$d(p,q)=\sqrt{(p_1-q_1)^2+(p_2-q_2)^2+...+(p_n-q_n)^2}$ (4)

p is the observed data
q is a class centroid
p1…pn is features
1…n of the observed data
q1…qn is features
1…n of a class centroid

k-Nearest Neighbours (k-NN) also used distance to help classification, only this time the observed object's distance is calculated against its nearest neighbouring objects rather than the class centroid. The number of neighbours taken into consideration is specified as the k number. k=5 means that 5 nearest neighbours are taken into consideration in determining the observed object's class. Say there are two classes (A, B) and within the 5 nearest neighbours, 2 of them belong to class A and 3 of them belong to class B, then the majority class label (B) is assigned to the observed object. That is why the k number needs to be odd and bigger than 1 for easy results.

RESULTS AND DISCUSSION

Although the k-NN classification using k=7 and k=9 bring the same classification accuracy rate, according to the confusion matrices, they are actually sum up from different numbers, e.g. there are 24 music files of genre C (classical music) which correctly classified using k=7, while there are only 23 files classified correctly using k=9. In the other hand, there were 2 music files of class I (pop music) which classified erroneously as blues using k=7, but there are more (3 files) using k=9. Overall, the maximum classification accuracy rate is 26.7% using either k=7 or k=9.



CONCLUSION

Experiments conducted in this research brings the conclusion that the features selected to classify digital music files of the WAV format by its sample properties, which are Zero Crossing Rate (ZCR), Average Energy (E) and Silence Ratio (SR) respectively, were able to distinguish files according to their music genres, albeit having accuracy as low as 33.3% with Nearest Centroid Classifier (NCC) and 22.5% with k-Nearest Neighbours (k-NN). The best class which is correctly classified is classical music, which yields a 100% classification accuracy rate for NCC and 56.3% for k-NN. This may have been caused by the absence of beat from most classical music, therefore any classical music used as test subject will have the shortest distance to the class centroid, using either method. Beat is also the problem for many classes since nowadays genres are seldom easily recognised by the human mind, and music which is a mix between genres does exist. The selection of many genres (Ballad, Blues, Classical, Harmony, Hip Hop, Jazz, Keroncong, Latin, Pop, Electronic, Reggae and Rock) themselves may also have been the cause of low accuracy, since the original class data may scatter tightly over the space causing confusion on test data, but classification for just 3 classes (Ballad, Blues, and Classic which centroids are notably near to each other) shown a good result of 96,7% classification accuracy rate. Generally speaking, the accuracy rate of k-NN for the same tests conducted are lower than NCC because in tight space, a test data is able to determine which centroid is nearer using NCC and associate itself with it, regardless the population surrounding it; while using k-NN, population of neighbours and their association are taken into consideration, in which for classes nearly spread, this may lead to confusion. Moreover, features selected, although proven to be able to classify to some extent, may also not be quite enough to determine a genre, and need other additional features.

REFERENCES

[1] G. Lu, Multimedia Database Management Systems. London: Artech House Inc., 1999, pp. 107-115. [2] Z.N. Li, and M.S. Drew, Fundamentals of Multimedia. Upper Saddle River, New Jersey: Pearson, Prentice Hall, 2004, p. 137.

[3] C. Perez-Sancho, and D. Rizo, “Genre Classification of Music by Tonal Harmony”, 2008, http://eprints.pascalnetwork.org/archive/00005171/01/mml08_cperez.pdf, accessed 5 July 2013.

[4] M. Santini, “Common Criteria for Genre Classification : Annotation and Granularity”, 2007, http://www.itri.brighton.ac.uk/~Marina.Santini/#Download, accessed 5 July 2013.

[5] Y. PanagakisPanagakis, C. Kotropoulos, and G.R. Arce, “Music Genre Classification Via Sparse Representations Of Auditory Temporal Modulations”, 2004 http://www.ece.udel.edu/about/documents/Music_Genre_Classification_v ia_Sparse_Representation_of_Auditory_Temporal_Modulations_EUSIP CO2009.pdf, accessed 18 July 2013.

[6] M. Talupur, S. Nath, and H. Yang, “Classification Of Music Genre”, 2004, http://www.cs.cmu.edu/~yh/files/GCfA.pdf, accessed 9 August 2013.

[7] E. Benetos, and C. Kotropoulos, “A Tensor-Based Approach For Automatic Music Genre Classification”, 2008, http://poseidon.csd.auth.gr/papers/PUBLISHED/CONFERENCE/pdf/Ben etos08a.pdf, accessed 9 August 2013.

[8] G. Tzanetakis, G. Esel, and P. Cook, “Automatic Musical Genre Classification Of Audio Signals”, 2001, http://ismir2001.ismir.net/pdf/tzanetakis.pdf, accessed 9 August 2013.

[9] Microsoft Corporation, “New Multimedia Data Types and Data Techniques”, 1994, archived by http://wwwmmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/Docs/RIFFNEW. pdf, accessed 5 July 2013 [10] R.O. Duda, P.E. Hart, and D.G. Strock, Pattern Classification, Second Edition. New York: John Wiley dan Sons, Inc., 2000, p. 12

