A Survey on Musical Feature Extraction and Classification Methods

Patel Prem[1], Prajapati Akshay[2], Gandhi Abhi[3], Pandya Mehul[4], Mrs. Mahajan Arpana[5], Dr. Sheshang Degadwala[6]

[1-5]U.G. Student, Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India
[6]Head of Department, Computer Engineering, Sigma Institute of Engineering, Vadodara, Gujarat, India

ABSTRACT

Identifying musical instrument is challenging task because of its multidimensional nature. Every particular instruments have its own characteristics and physical features like energy feature, rhythm feature, temporal feature, spectrum feature, harmony feature etc. In this paper toolboxes that are all publically available for extracting these features. & the perception of valance and arousal has been also discussed. This paper offers an overview of the set of upper features. Particular attention has been paid to design of a syntax that offers both simplicity of use & transparent addictiveness to a multiplicity of possible input also the same syntax can be used for analysis of signal audio files, batch files, series of audio segments multi banned signals. Also we have studied about the preprocessing and automatic speech recognition the preprocessing is done and voice speech is detected based on energy and zero crossing rates.

Keywords : Musical instruments, Musical Features, Rhythm, Temporal, Musical Classifier.

I. INTRODUCTION

Automatic Instrument recognition is a subtask of for the musical content identification. Also it is important for easy to access essential data from vast amount of data. It could be treated as the first step in developing annotation.

Musical instrument belongs to wide range of instruments with different characteristics that includes physical aspects different sound initiation process. So classifying musical instruments becomes a complex issue, because of multidimensional nature of musical instruments. Also now day there is rapidly growth in amount of digital media so, the need for effective data management is challenging.

Musical Instruments using SVM & MP classifiers helps to reduce manual work tool identify and classifier musical instruments according to their attributes. In first part this system takes the musical instrument as input and then identifies the attributes for musical instrument. Musical instruments with their attributes are inserted into dataset. In the second part SVM and MLP classifiers are applied upon dataset. These classifiers identify the correctly and incorrectly classified instances. In third part result generated by classifiers are analysed, analysis result shows that SVM gives better result than MLP classifier. Formal concept analysis generates conceptual structure for musical instruments stored in dataset.

II. RELATED WORKS

[1]Swati D. Patil, Tareek M. pattenwar et.al Musical Instruments Identification using SVM & MLP with formal concept Analysis. This system automatically classifies musical instruments. This system describes how intend and extend are related to each other. By using the concept analysis techniques. The limitation of the system is that it does not convert audio input to printable form.

[2]Olivier Lartillot, Petri Toiviainen et.al A Matlab Toolbox for musical features Extraction from Audio. In this Paper, we found an overview of the best of features related among others to timber, tonality, rhythm or form, that can be extracted with MIR toolbox. The toolbox includes functions for statistical analysis, segmentation and clustering. The limitation of the system is that MIR toolbox doesn't offer real time capabilities.

[3]Mayur R. Gamit, Prof. Kinnndle Dhameliya et.al English Digits Recognition using MFCC, LPC and Pearson's correlation. This paper presents the use of a Pearson's correlation for English digit recognition. The preprocessing is done and voiced speech is detected based on energy and zero crossing rates (ZCR). The limitation of the system is that the experimental result shows that by using the proposed MFCC and combination of both MFCC techniques and Pearson's correlation result are higher as compare to system.

[4]Yading Song, Simon Dixon, Marcus Pearce et.al Evaluation of musical features for emotion classification. In this paper the musical mood is linked to features based on rhythm, timber, spectrum and lyrics. We collect a ground truth data set of 2904 songs that have been tagged with the one of the four words "Happy", "Sad", "Angry", and "Relax". The limitation of the system is that this paper we only find moods and played music as songs.
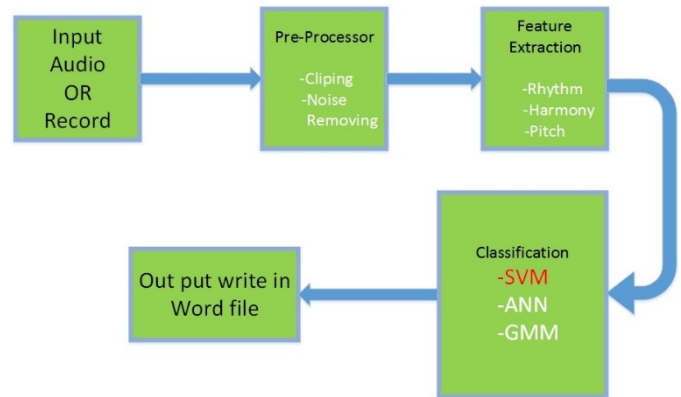
III. SYSTEM METHODOLOGY

FIG 1 SYSTEM METHODOLOGY

STEP 1: INPUT AUDIO

In this module we are giving an audio as an input to our system. The possible instruments in this module are Piano, Guitar, Electronic Keyboard, etc. This Musical Auto Notated systems are tightly coupled with the source of particular Melody Instrument, which is necessity of our software.

STEP 2: PRE-PROCESSOR

Pre-Processor is the Module which is used for to increase the efficiency of subsequent feature extraction and classification stages and therefore to improve the overall recognition performance. In this stage we are removing the external Noise and Cropping.

STEP 3: FEATURE EXTRACTION

Feature Extraction is the module which describe the feature extraction methods for speech signals. In this module we will add the Audio-Features like Energy Features, Rhythm Features, Temporal Features, Spectrum Features, and Harmony Features.

A. Rhythm

Rhythm is the pattern of pulses/notes of varying strength. It is often described in terms of tempo, meter, or phrasing. A song with fast tempo is often perceived as having high arousal. Besides, flowing rhythm is usually associated with positive valence, whereas firm rhythm is associated with negative valence. Because

of their importance, many toolboxes can be utilized to extract rhythm features [8].

B. Temporal

Using the sound description toolbox to extract zero-crossing rate, temporal centroid and log attack time to capture the temporal quality of music. Zero-crossing rate [3], a measure of the signal noisiness, is computed by taking the mean and standard devia- tion of the number of signal values that cross the zero axis in each time window.

C. Spectrum

Spectrum features are features computed from the STFT of an audio signal. One can use Marsyas to extract the timbral texture features including MFCC, spectral centroid, spectral rolloff, spectral flux, spectral flatness measures (SFM), and spectral crest factors (SCF). These features are extracted for each frame and then by taking the mean and standard deviation for each second. The sequence of feature vectors is then collapsed into a single vector representing the entire signal by taking again the mean and standard deviation. Note that many of these features can also be extracted by the sound description toolbox [9].

D. Harmony

Harmony features are features computed from the sinusoidal harmonic modeling of the signal. A lot of natural sounds, especially musical ones, are harmonic each sound consists of a series of frequencies at a multiple ratio of the lowest frequency, called the fundamental frequency f=0.

STEP 4: CLASSIFICATION MODULE

Classification is a supervised machine learning algorithm which can be used for both Classification and regression Challenges. It is mostly used in classification problem. The different classifiers are SVM (Support Vector Machine), GMM (Gaussian Mixture Model), ANN (Artificial Neural Networks).

A. SVM

support vector machine are relatively new approach that has rapidly gained popularity because of the very fast and accurate result they have achieved in a wide variety of machine learning problems. This method is divide the n dimensional space represented data into two region hyper plane. This hyper plane always maximize the margin between the two regions(classes). The margin is defined by the longest distance between the example of the two classes and is computed based on the distance between the closest instances of both classes to the margin, which are called supporting vectors. support vector machine is also a supervised learning method for automatic pattern recognition.

B. KNN

The K-Nearest Neighbour(K-NN)classifier is an example based classifier; it means that the classifier needs training files for comparison rather then an precise category rep- resentation,like the category profiles used by existing other classifier. As such,there is no real training phase. When a new file needs to be categorized,the k most sim- ilar files(neighbour)are found and if a large adequate proportion of them have been assigned to a certain category,the new file is also assinged to this category, otherwise not. Additionally,finding the nearest neighbour can be accelerated using traditional indexing method. In k-NN classification, the output is a class membership. An object is classified by a majority vote of its neighbours, with the object being assigned to the class most common among its k nearest neighbours.

C. GMM

A Gaussian mixture model is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. One can think of mixture models as generalizing k-means clustering to incorporate information about the covariance

structure of the data as well as the centers of the latent Gaussian [6].

D. ANN

In machine learning and cognitive science, artificial neural networks (ANNs) are a family of models inspired by biological neural networks (the central nervous systems of animals, in particular the brain) and are used to estimate or approximate functions that can depend on a large number of inputs and are generally unknown. Artificial neural networks are generally presented as systems of interconnected "neurons" which exchange messages between each other. The connections have numeric weights that can be tuned based on experience, making neural nets adaptive to inputs and capable of learning [16].

STEP 5: OUTPUT

In this module finally our system recognizes the audio that we had given as input.

IV. RESULT ANALYSIS AND DISCUSSION

TABLE I
FEATURE EXTRACTION METHODS

No.	Feature Group			
	Energy	Temporal	Rhythm	Harmony
1	RMS, Pitch	MFCC, Spectral, Centroid, Entropy, Roughness	Strength, Regularity, Tempo, Envelop	---
2	RMS, Pitch	MFCC, Spectral, Centroid, Entropy, Roughness	Strength, Regularity, Tempo, Envelop	Spectrum, Chromagram, Wrapped chromagram Key, Keystrength, Mode

TABLE III
CLASSIFICATION METHODS ACCURACY

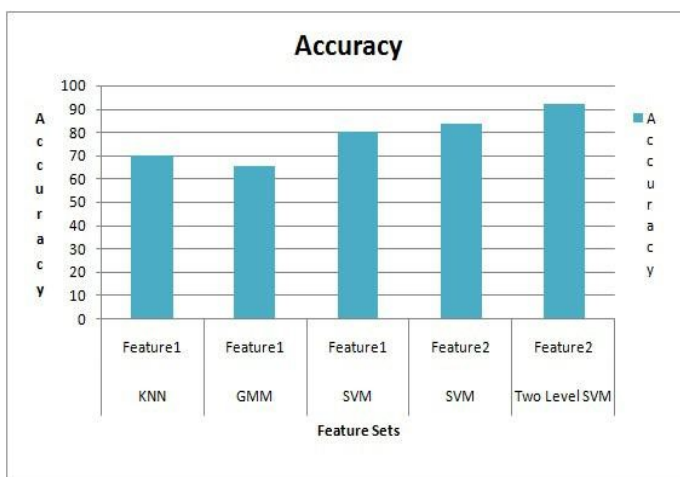| SR.NO	Features	Method	Accuracy
1	STFT+MFCC	K-NN	60%
2	STFT+MFCC	GMM	61%
3	MFCC+GMM	SVM	77.7%
4	MFCC+GMM	K-NN	70.4%
5	MFCC+ASE+OSC	SVM	72.4%

FIG 2 CLASSIFICATION METHODS ACCURACY

V. CONCLUSION

We have completed our project work based on using software engineering and system analysis and design approach. Work that we have done with preplanning scheduling related with time constrains and result oriented progress in project development. The proposed work is completed by making use of the literatures and other aids. The work deals with real-time algorithms and techniques for multiple Music Recognition of any tune. A new approach is presented where HOG features are incorporated with SVM, allowing the algorithm to detect multiple tune with greater efficiency. In future, more feature set with multiple classification methods can be added in our proposed work for better accuracy.

VI. REFERENCES

[1]. Swati D. Patil, Tareek M. Pattewar, PG Student, Department of Computer Engineering, SES's R.C. Patel Institute of Technology, Shirpur, Maharashtra, India, Assistant Professor, Department of Information Technology, SES's R.C. Patel Institute of Technology, Shirpur, Maharashtra, India; MUSICAL INSTRUMENT IDENTIFICATION USING SVM & MLP WITH FORMAL CONCEPT ANALYSIS.

[2]. Olivier Lartillot, Petri Toiviainen, University of Jyvaskyla Finland; A MATLAB TOOLBOX FOR MUSICAL FEATURES EXTRACTION FROM AUDIO.

[3]. Mayur R Gamit, Pro. Kinnal Dhameliya, M.Tech Research Student, Assistant Professor, C.G.P.I.T, Bardoli, India; ENGLISH DIGITS RECOGNITION USING MFCC, LPC AND PEARSON'S CORRELATION

[4]. Yading Song, Simon Dixon, Marcus Pearce, Centre for Digital Music, Queen Mary University of London; EVALUATION OF MUSICAL FEATURES FOR EMOTION CLASSIFICATION.

[5]. Babu Kaji Baniya and Choong Seon Hong, Department of Computer Engineering Kyung Hee University; MUSIC MOOD CLASSIFICATION USING REDUCED AUDIO FEATURES.

[6]. I. Nabney, Springer Advances In Pattern Recognition Series, chapter NETLAB: ALGORITHM FOR PATTERN RECOGNITION.

[7]. Xian Tang, Pacific – Asia Conference on Circuits, Communication and System, IEEE Computer Society 2009; HYBRID HIDDEN MARKOV MODEL AND ARTIFICIAL NEURAL NETWORK FOR AUTOMATIC SPEECH RECOGNITION.

[8]. Nidhi Desai, Prof. Kinnal Dhameliya, International Journal of Emerging Technology And Advance Engineering (IJETAE), Vol 3, Issue 12, December 2013; FEATURE EXTRACTION AND CLASSIFICATION TECHNIQUES FOR SPEECH RECOGNITION.

[9]. R. Panda, R. Malheiro, B.Rocha, A. Oliveira and R.P.Paiva, International Symposium on Computer Music Multidisciplinary Research 2013; MULTI - MODEL MUSIC EMOTION RECOGNITION.

[10]. Zengchang Qin, Wei Liu, Tao Wan, Inte1lligent Computing and Machine Learning Lab School of ASEE, Beihang University, Beijing, 100191, China, Case Western Reserve University, Cleveland, OH 44106, USA; A BAG OF TONES MODEL WITH MFCC FEATURES FOR MUSICAL GENRE CLASSIFICATION

.