Impact Factor (SJIF): 3.632

International Journal of Advance Research in Engineering, Science & Technology

e-ISSN: 2393-9877, p-ISSN: 2394-2444

Volume 4, Issue 3, March-2017

Survey for Classification of Music Instruments using MFCC and spectral Descriptors

Prof. Sarangi Mehta [1], Assistant Professor

Prof. Dipti Mathpal [2], Assistant Professor

[1] *Department of Information Technology, Sardar Patel College of Engineering, Anand, India*

[2] *Department of Computer Engineering, Sardar Patel College of Engineering, Anand, India*

Abstract ---Today music has been a soul medicine in this era and technology has provided much advancement to work in this area. Music Information Retrieval is one of the fields to provide more advancement in music. There are certain applications like automatic tagging, cataloguing of music for search engine, automatic musical transcription, genre classification which require audio classification. This paper concentrates on Instrument classification of different Indian instruments category. To classify instruments we apply feature extraction on audio. Till now too many features are used and experiments are made using different classifiers. This paper reviews different features and techniques to classify Indian music instruments.

Keywords: Music Information Retrieval, Music Instruments Classification, Feature Extraction, Audio Classification, Spectral Features

I. INTRODUCTION

Music Information Retrieval (MIR) involves classification of Indian music instruments which is an emerging research field. There are various applications which involve MIR like Automatic Music Transcription, Track Separation, Automatic Music Generation, Music Recommender System and Music Classification. For any music analysis signal processing is required to process audio. Basically music is classified into two categories. Monophonic and Polyphonic. Monophonic music is generated from single instrument where polyphonic music is generated from multiple instruments. There are many signal processing techniques and features used for classification like temporal features, perceptual features, energy features, and spectral features etc. Basically MFCC is used for instrument recognition. In this paper Section II gives details about instrument and its category, sound and its analysis which helps to understand instrument recognition. Section III gives idea about different properties to identify sound. Section IV contains features and its study. Section V contains literature survey for Instrument identification features over the last decade.[10][11]

II. INSTRUMENT CATEGORY

1.1 Categorization According to Physical Characteristics [13][14]

1) The *Tata Vadya* or Chordophones- Stringed instrument

The *Tata Vadya* is a category of instruments in which sound is produced by the vibration of a string or chord. These vibrations are caused by plucking or by bowing on the string which has been pulled taut. The length of the vibrating string or wire, the degree to which it has been tightened, determines the pitch of the note and also to some extent the duration of the sound.

All these string instruments were classified into four main categories

a) Plucked Instruments*: Tanpura, Swar Mandal*
b) Bowed Instruments*: Sarangi, Violin*
c) Stroking Instruments*: Sitar, Sarod*
d) Hammered Instruments: *Santoo*r

2) The *Sushira Vadya* or Aerophones- Wind instruments

In the Sushira Vadya group, sound is produced by blowing air into a hollow column. The pitch of the note is determined by controlling the air passage and the melody is played by using the fingers to open and close the in the instrument. Example *Flute, Shehnai, Harmonium.*

3) The *Avanaddha Vadya* or Membranophones- Percussion Instruments

In the Avanaddha Vadya category of instruments, sound is produced by striking the animal skin which has been stretched across an earthern or metal pot or a wooden barrel or frame.Example *Tabla, Dhol, Dholak* etc.

4) The *Ghana Vadya* or Idiophones- Solid Instruments

The earliest instruments invented by man are said to be the Ghana Vadya. This variety of instrument does not need special tuning prior to playing. In early times these instruments were the extension of the human body such as sticks, clappers, rods, etc. and were also closely related to objects of utility in daily life such as *pots, Manjira*, etc. They are principally rhythmic in function and are best suited as accompaniment to folk and tribal music and dance.

1.2 Categorization According to Note Texture

Based on notes (SA, RE, GA, MA, PA…) played together instrument can be categorized into two categories.

- **Monophonic:** An instrument which can play single note at a time is monophonic instrument. E.g. *Flute, Shehnai, Violin*

- **Polyphonic:** An instrument which can play multiple notes at a time is polyphonic instrument. E.g. *Harmonium, Keyboard.*

III. FEATURES OF INSTRUMENT SOUND

Any musical instrument is a sound and each sound has some general properties whether it is instrument sound, vocal sound, any object sound or any gender specific sound. Each sound can be basically categorized through some general properties.

- **Pitch**
  Pitch is related to the concept of the fundamental frequency of a sound; Higher the frequency higher the pitch of sound.

- **Amplitude**
  It is loudness of sound; the intensity is also defined as loudness.

- **Timbre**
  It is the tonal quality of sound. Each sound has its own tone which is known as timbre. As an example voice of two persons are different due to different timbre of two persons.

- **Articulation**
  It is the envelope or shape of the signal which may help to identify an instrument. Some instruments directly play with the high amplitude which gives pick to the envelope in the start, which is called as attack and some play soft in starting and after some fragment it goes to the pick changing the shape of the envelope.

- **Tempo**
  It is the feature to identify rhythmic sound. The tempo is the speed at which a musical work is played, or expected to be played, by performers. The tempo is usually measured in beats per minute.

- **Diffusion**
  It is a direction of sound from where it comes, but basically for instrument identification this does not help to extract feature of sound.

These above explained properties are general for every sound. Each property is extracted by some spectral and harmonic features which give result for the identification of instruments. Basic understanding for properties is given to get to know about the sound and its definition and to extract how sound can be analyzed. Based on the basic analysis any speech or

sound produced from the instrument, object is classified. Timbre is difficult to identify programmatically as its frequency matching to the human's ear is similar to mel frequency and the feature like MFCC are required to get timbre value.

Below are the figures describing some of the properties for any sound.

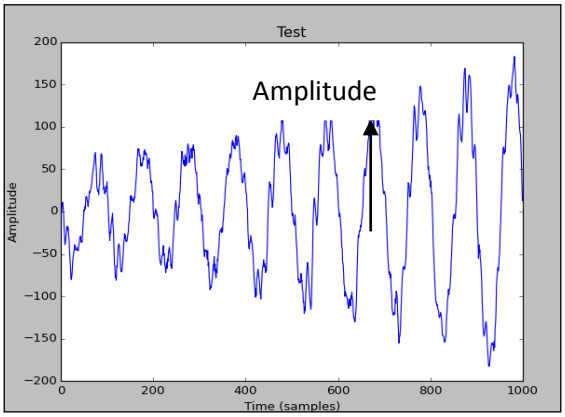

Figure 1. Understanding of sound properties (Amplitude and frequency)[18]

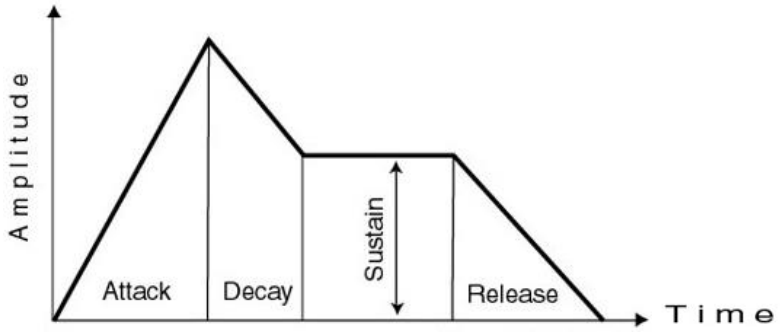

Figure 2. Understanding of sound properties (Articulation)[18]

IV. Music Features

V. LITERATURE SURVEY

Musical instruments recognition is an emerging field in signal processing. This technique has four steps: signal acquisition, feature extraction, dimensionality reduction and classification. After signal acquisition features are extracted from the audio signals and based on the features signal is classified into the category. From past many years basic feature used to classify the signal is MFCC (Mel Frequency Cepstral Coefficient) and others are combined to optimize classification result. Certain features are reduced using dimensionality reduction. Polyphonic music classification requires extraction of different instruments separation or audio is used to check the presence of certain instrument. For monophonic music this separation process is eliminated as it contains the single instrument audio [17]. Certain researchers estimate the pitch for instrument while focusing monophonic classification but it can be omitted also. Till now many classification algorithms are used and discussed in next section.

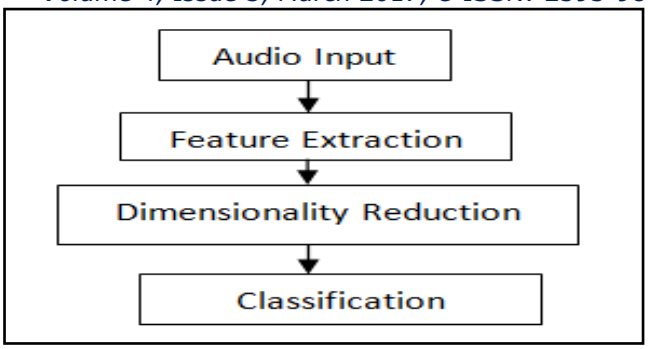

Figure 3. Block Diagram

El˙ zbieta Kubera et al.[1] recognized musical instrument from polyphonic audio and checked whether particular instrument is present or not. Total 10 instruments' mixes were taken from RWC, IOWA and MUMS. Random Forest was used as a classifier to check the instrument presence. Spectral features were obtained using Fourier Transform and other features like MFCC, ZCR (Zero Crossing Rate) are also used. Classifier was applied on two different training sets to check the accuracy. The classifier built for the training on single sounds and mixes gives quite low recall, but using real non-segmented recordings for training improves the recall significantly, no matter the size of the training set.

Kohshelan et. al [2] improved feature extraction technique for Indian musical instrument. They integrated basic feature with the feature called ZFE to compare accuracy of result. First experiment was done with MFCC, LPC (Linear Predictive Coding) and ZCR (Zero Crossing Rate).Then they combined each feature with ZFE(Zero Forcing Equalizer).The proposed technique showed better classification accuracy with noisy signal. They used dataset of string instrument "*Veena*" and its different forms. Classification techniques used were kNN, Bayesian Network and SVM. Live recording were used as dataset.

Dimitrios et.al [3] proposed another algorithm "Missing Feature Approach" for instrument recognition. Missing feature approach was used to recognize instrument from polyphonic audio without requiring track separation and pitch detection. They used 10 western instruments from RWC and MUMS datasets. MFCC and other spectral features are used to obtain feature vector. Performance was compared between GMM and proposed Algorithm. Oracle mask was used to observe spectral changes in a time domain of a signal. Proposed method outperformed the baseline.

Dominique et. al [4] proposed a method for better feature selection. Experiment was basically for non-western instruments. Proposed method offers practical solution for automatic ethnomusicological indexing of various instrument sounds. They proposed Inertia Ratio Maximization using Feature Space Projection (IRMFSP) for better feature selection from the large feature vector. Comparison was made with LDA (Linear Discriminant Analysis) and MI (Mutual Information). IOWA dataset was used for general instruments and CREM dataset was used for non-western instruments. Spectral, Perceptual and Harmonic descriptors were used for different datasets as features. As classification classes and selection of best descriptors are interrelated, proposed work gave better classification than traditional feature selection algorithm.

M. Kumari et. al [5] proposed their work for north Indian instruments. Experiment was made on 5 instruments like *Flute*, *Sitar*, *Dholak*, *Bhapang*, and *Mandar*. Spectral features and MFCC were used to classify instruments. For classification ANN was used to get the result. Autocorrelation coefficients were also used to classify and to check the accuracy of classification of result. The autocorrelation of a frame represents the distribution of the signal spectrum but in the time domain.

Satish Ramling Sankaye et.al [6] proposed their work for the category of *Sushir* and *Ghana vadya* category which is less concentrated till now. They used Linear Predictor Coefficient (LPC) features and Linear Discriminant Analysis (LDA) as the Classifiers. They used total 370 audio files of different instruments of category *Sushir* and *Ghana* including instruments like *Bansuri*, *Shehnai*, *Harmonium*, *Ghungaroo*, *Manjira*, *Triangle* and *Ghatam*. The experiment was carried out using 10-fold cross validation and 15 LPC Features. The highest mean accuracy of 94.44% was achieved by using standard Best First Decision Tree algorithm.

Aijit Ghoshal et. al [7] classified music as song (music with voice) and instrumental (music without voice) based on MFCC. Subsequently, instrumental signals and songs are classified based on instrument type and genres respectively. Hierarchical approach has been followed for such detailed categorization. Using two-stage process, instrumental signals are identified as one of the four types namely, string, woodwind, percussion or keyboard. Wavelet and MFCC based features are used for this purpose. For song classification, at first level signals are categorized as classical or non-classical

(popular) ones by capturing the MFCC pattern present in the high sub-band of wavelet decomposed signal. At second level, task of further classification of popular songs into various genres like Pop, Jazz, Bhangra (an Indian genre) based on amplitude variation pattern is adopted. RANSAC has been utilized as the classifier at all stages. Experimental result indicates the effectiveness of the proposed schemes.

Priyanka S. Jadhav [8] proposed a system which deals with recognition of musical instruments from monophonic audios. The music related features are extracted from audio samples by using timbral features extractors as well as MFCC feature extraction method. Three different classifiers namely K-Nearest Neighbors (K-NN), Support Vector Machine (SVM) and Binary Tree (BT) are used to identify musical instrument from a sound sample. The system gives maximum percentage accuracies of 92.00% and 88.00% for combinations of MFCC with BT classifier and Timbral ADs with BT classifier respectively; for five musical instruments. MFCC with K-NN classifier and Timbral ADs with BT classifier give maximum percentage accuracies of 77.00% and 84.00% respectively; for ten musical instruments. For fifteen musical instruments;MFCC with K-NN classifier and Timbral ADs with BT classifier give maximum percentage accuracies of 75.33% and 73.33% respectively. By studying all results one can conclude that the proposed system gives higher accuracy for MFCC if K-NN classifier is used and for Timbral ADs if BT classifier is used.

Table 1. Survey Content [7][8]

Year	Features	Reduction Technique	Classifier
2003	Cepstral coefficient, constant Q transforms frequency spectrum, RMS amplitude envelope, spectral centroid and vibrato	PCA	kNN
2005	Peak kurtosis, Modulation features, Temporal feature	PCA, LDA	Bayes decision rule
2007	Temporal , Energy, Spectral, Harmonic, Perceptual	LDA	kNN
2008	ZCR, Spectral Features, MFCC, Harmonic Deviation, Spread, Variation, Long attack Time	PCA, Isomap	K-NN, RBF & SVM
2009	Spectral,Temporal Features, MFCC		SVM
2010	OverCs, SparCs	-	SVM
2013	Pitch histogram and Pitch-frequency scaled spectrum	-	single hidden layer feed-forward NN
2015	STRF	-	SVM

V. CONCLUSION

Music Information Retrieval and Digital signal processing are both emerging field. Till now most of the work is on concentrating some of the categories of music instruments and western instruments. All four categories still need to be focused to get better result. As per the survey different features are used with different classifiers but most of the research work includes one of the western instruments and as the conclusion pure Indian classical instruments with different set of features and large dataset will always get attention of music and technology seeking researchers.

VI. REFERENCES

[1]Elz˙bietaKubera, and Alicja A. Wieczorkowska, ”Mining Audio Data for Multiple Instrument Recognition in Classical Music” in Springer International Publishing Switzerland 2013.
[2]Kohshelan and Noorhaniza Wahid , “Improvement of Audio Feature Extraction Techniques in Traditional Indian Musical Instrument “,in Springer International Publishing Switzerland 2014.
[3]Dimitrios Giannoulis, and Anssi Klapuri “Musical Instrument Recognition in Polyphonic Audio Using Missing Feature Approach “in IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 21, NO. 9, SEPTEMBER 2013

[4]Dominique Fourer, Jean-Luc Rouas, Pierre Hanna, Matthias Robine"Automatic Timbre Classification of Ethnomusicological Audio Recordings" in 15th International Society for Music Information Retrieval Conference (ISMIR 2014)

[5]M. Kumari, P. Kumar and S. S. Solanki " Classification Of North Indian Musical Instruments using Spectral Features" in GESJ: Computer Science and Telecommunications 2010|No.6(29)

[6] Satish Ramling Sankaye , S.C. Mehrotra, Perspective Identification of Indian Musical Instrument Ghan and Sushir Vadya U.S. Tandon in ijcsit, 2015

[7]Arijit Ghosal, Rudrasis Chakraborty, Bibhas Chandra Dhara,and Sanjoy Kumar Saha, "Music Classification based on MFCC Variants and Amplitude Variation Pattern: A Hierarchical Approach", in International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 5, No. 1, March, 2012

[8]Priyanka S. Jadhav, "Classification of Musical Instruments sounds by Using MFCC and Timbral Audio Descriptors",in International Journal on Recent and Innovation Trends in Computing and Communication ,Volume: 3 Issue: 7

[9]Manish Kumar Saini, Vipin Bhargava" Signal Processing Techniques for Musical Instrument Recognition", in ELSEVIER 2014

[10]Kailash Patil and Mounya Elhilali, Biomimetic spectro-temporal features for music instrument recognition in isolated notes and solo phrases, in EURASIP 2015

[11] Ferdinand Fuhrmann," Automatic musical instrument recognition from polyphonic music audio signals "in 2012

[12]MUSTAFA YANKAYIŞ FEATURE EXTRACTION MEL FREQUENCY CEPSTRAL COEFFICIENTS (MFCC)" a thesis report

[13] Music Book "Sangeet Saurabh"

[14]https://www.swarganga.org/articles/details.php?id=6

[15]https://ccrma.stanford.edu/workshops/music-information-retrieval-mir

[16]https://en.wikipedia.org/wiki/Music_information_retrieval

[17]http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/

[18]https://www.google.co.in/search?q=Envelop+Graph