Impact Factor (SJIF): 5.301

International Journal of Advance Research in Engineering, Science & Technology

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

Volume 6, Issue 6, June-2019

Musical Instrument Detection

Prasad Dengle [1], Rohit Tamboli[2], Shailesh Chaudhary[3], Sanjiv Sonawane[4]

[1]*Computer Department, SAE KONDHWA*
[2]*Computer Department, SAE KONDHWA*
[3]*Computer Department, SAE KONDHWA*
Under the guidance of Prof.Shalini Wankhede

Abstract — Traditionally, musical instruments are categorized into four main categories or families: string, brass, woodwind, and per- cussion. For example, violin is a classic string instrument, and clarinet belong to the woodwind category, horn and trumpet are brass tools, and piano is usually classified as a drumming instrument. Sounds produced by these tools bear different acoustic qualities. A few features can be obtained from their sound envelopes, including attack (the time from silence to amplitude peak), sustain (the time length in preserving level amplitude), decay (the time the sound fades from sustain to silence), and issue (the time of the decay from the moment the device stops playing). To reach accurate classification of instruments, more complex features need to be extracted.

I. INTRODUCTION

Because of the difficulty of modeling instrument timbre, various feature arrangements have been proposed through acoustic study and pattern respect research. Our main intentions are to examine the performance of different feature systems and find a good feature combination for a robust instrument classifier. Here, we consider three different extraction resources, namely, perception-based features, MPEG-7-based features, and MFCC. The first two feature sets consist of temporal and spectral structures, whereas the last is based on spectral analysis.set for the description of harmonic instrument sounds. We consider in this paper only two lessons of timbral descriptors in the MPEG-7 framework: timbral spectral and timbral temporal. These contain seven feature descriptors: harmonic centroid (HC), harmonic deviation (HD), harmonic spread (HS), harmonic variation (HV), spectral centroid (SC), log attack time (LAT), and temporal centroid (TC). The first five belong to the timbral spectral feature scheme, while the last two belong to the timbral sequential scheme. Note that the SC feature value was obtained from the spectral analysis of the entire sample signal; thus, it is similar to but different from the CentroidM of the perception-based features. CentroidM was aggregated from the centroid feature extracted from short sections within a sound sample. Feature selection techniques are often necessary for optimizing the feature sets used in cataloging. This way, redundant features are detached from the classification course, and the dimensionality of the feature set is reduced to save computational cost and defy the “curse of dimensionality” that impedes the construction of good classifiers. To assess the excellence of a feature used for classification, a correlation-based method is often adopted. In general, a feature is good if it is relevant to the class concept but is not redundant given the presence of other relevant features. The core issue is modeling the association between two variables or structures. Based on information theory, a number of pointers can be developed to rank the features by their association to the class. Related features will yield a higher correlation.

III. Proposed System

Identifying the objects in the situation from the sound they produce is primary task of auditory system. The aim of Musical instrument respect is to classify the name and family of musical instrument from the sound they produce. The arithmetical pattern-recognition method for classification of some musical tool tones with few features based on log-lag correlogram

IV. System Architeture

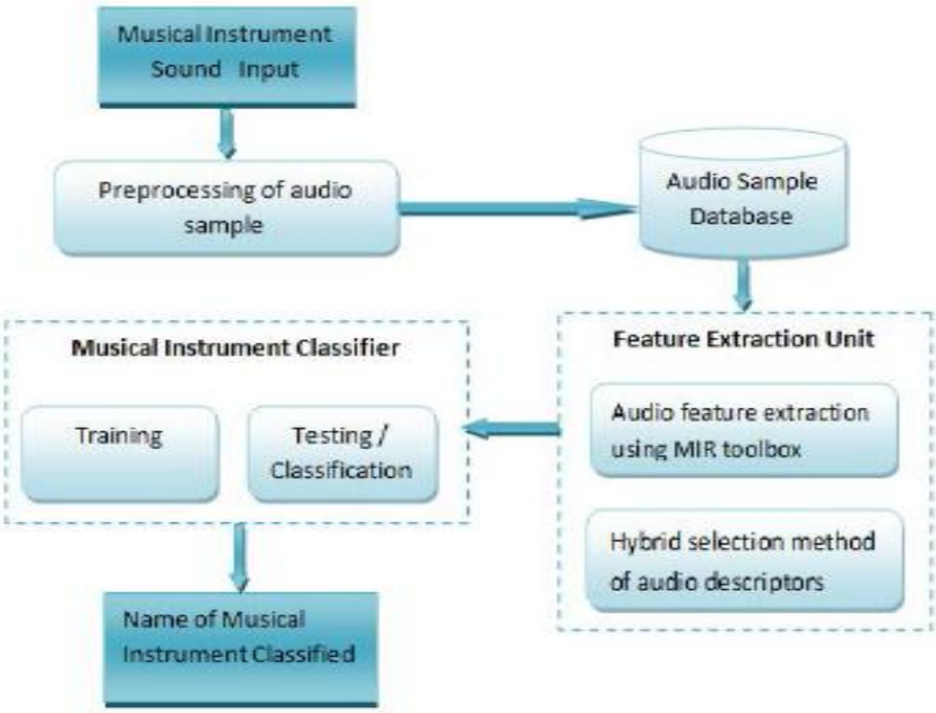

Fig 3.1: Block Diagram of Musical Instrument Recognition System

3.1 Explanation:

The system is design for the identification and classification of instruments. In order to classify musical instrument into particular class we need to find out significant information about the signal. This is referred as feature extraction process. The parameter extracted from music signal depends on type of instrument and it's playing style. This significant information will be used in categorization process to identify the right instrument. The system is divided in two parts: Training and Testing.

Fractional Fourier Transform Based Features for Musical Instrument Recognition Musical Instrument recognition system consists of three stages as,

i) preprocessing, ii) feature extraction, and iii) Classification. Initially, signal is given to pre-processing stage to make signals suitable for feature extraction. In pre-processing the silence part of the signal is removed. In preprocessing stage the signal is framed with 20 ms and windowed with hamming window. In feature extraction stage, acoustic features based on MFCC, Timbrel, Wavelet and FRFT have been extracted and applied to different classifiers.

3.1.1. **Training Phase**: In this phase set of known signal is used an input. The feature of known signal will be extracted using FRFT and this features placed in a matrix or vector format as a Reference Model which contain standard database for classification.

3.1.2. Testing Phase: In this phase an unknown test signal will be given as an input and using FRFT feature of the signal will be extracted. This feature will be compared with the reference features. By using classifier we are able to identify which feature matching amongst all feature. We are in position to identify instrument and family.

3.2 Block diagram consist of –

3.2.1. Pre-processing: In this stage noise present in signal, silence part from input signal is removed. Silence present in signal is removed with different methods like Zero Cross detection Rate (ZCR), short time energy distribution based method. After this Framing, windowing is done to remove end effects.

3.2.2. Feature extraction: Musical signals have different features. There are different feature extractions techniques are available to find exact feature of signal. We will use FRFT (Fractional Fourier Transform) to extract the features of signal. Features of signal will extract in this step.

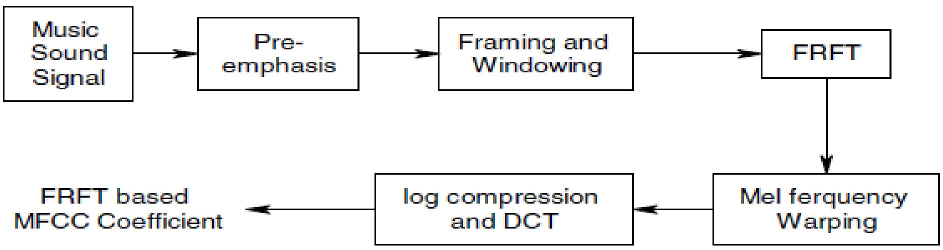

Fig 3.2: Feature Extraction

3.2.3. Reference model: In reference model, features are stored in matrix format for each instrument. Based on features vector unknown input signal is classified and recognized.

3.2.4. Classifier: It has input from reference model and extracted feature of unknown signal. Classifier is used to classify test signal on the basis of reference vector. We used KNN classifier for our system.

ALGORITHM:

CNN:

The Convolutional Neural Network gained popularity through its use with image data, and is currently the state of the art for detecting what an image is, or what is contained in the image. CNNs even play an integral role in tasks like automatically generating captions for images.

Haar Haar Cascades:

- We will see the basics using Haar Feature-based Cascade Classifiers

A Haar Cascade is basically a classifier which is used to detect the object for which it has been trained for, from the source. The Haar Cascade is trained by superimposing the positive image over a set of negative images. The training is generally done on a server and on various stages.

Haar-like features are digital image features used in object recognition. They owe their name to theirintuitive similarity with Haar wavelets and were used in the first real-time face detector. Historically, working with only image intensities (i.e., the RGB pixel values at each and every pixel of image) made the task of feature calculation computationally expensive. A publication by Papageorgiou et al. discussed working with an alternate feature set based on Haar wavelets instead of the usual image intensities. Viola and Jones[1] adapted the idea of using Haar wavelets and developed the so-called Haar-like features. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. For example, let us say we have an image database with human faces.

It is a common observation that among all faces the region of the eyes is darker than the region of the cheeks. Therefore a common Haar feature for face detection is a set of two adjacent rectangles that lie above the eye and the cheek region. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object (the face in this case).

CONCLUSION

We have described a system that can listen to a mixture of musical instruments and recognize them. The work started by reviewing Blind Source Separation and Musical Instrument Recognition. Features which make musical instrument distinguishable from each other are presented and discussed. The principle of classifier k-NN are described. Features are extracted from estimated sources and normalized to keep generality. The k-NN classifier is used to evaluate the testing data on this identification system. In order to make truly truthful estimations, more acoustic data would be needed.

REFERENCES

1. "Fractional Fourier Transform Based Features for Musical Instrument Recognition Using Machine Learning Techniques", by D.G.Bhalke – Prof. JSPM's RSCOE, Pune,India.

2. "Speech And Audio Processing" by Dr. Shaila D. Apte - Prof. JSPM's RSCOE, Pune India .

3. "Melody Extraction from Polyphonic Music Signal using STFT and Fanchirp Transform", by Sridevi S. H. & Prof. S. R. Gulhane , Dept. of E&TC, DYPCOE, Ambi, Talegoan, Pune, India.

4. "Automatic genre classification of Indian Tamil and western music using fractional MFCC" by Betsy Rajesh D. G. Bhalke Received: 14 January 2016 / Accepted: 6 June 2016_ Springer Science+Business Media New York 2016

5. "RWC Music Database: Music Genre Database and Musical Instrument Sound Database" by Masataka Goto-National Institute of Advanced Industrial Science and Technology (AIST) / "Information and Human Activity", PRESTO, JSTIT, AIST, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8568, Japan m.goto@aist.go.jp

6. "MUSICAL INSTRUMENT IDENTIFICATION IN CONTINUOUS RECORDINGS" by *Arie A. Livshin* Analysis/Synthesis Team Ircam, Paris,

7. "Melody extraction from polyphonic audio of western opera: A method based on detection of the singer's formant" by *Zhen Tang , University of Washington , Department of Electrical Engineering and Dawn A.A.* Black published in 15th international society for music information Retrieval Conference (ISMIR 2014).