International Journal of Latest Engineering Research and Applications (IJLERA) ISSN: 2455-7137

Volume – 02, Issue – 05, May – 2017, PP – 11-16

A Review: A Singer Identification and Analysis of Musical Performance

Kumari Rambha Ranjan[1], Kartik Mahto[2]
[1](*Dept. of Electronics and Communication, Birla Institute of Technology, Mesra, Ranchi, India)*
[2](*Asst. Professor, Dept. of Electronics and Communication, Birla Institute of Technology, Mesra, Ranchi, India)*

Abstract: This paper aims to review/survey a singer identification and classification of musical performance and also classified by extracting audio features from real sample sounds. This study provides an overview of various feature and model-based approaches developed in past for robust singer voice identification. The importance and discussion of some standard methods applied for robust singer voice verification tasks have been highlighted. The main focus is to summarily introduce popular state-of-the-art techniques adopted for enhancing singer voice verification performance in noisy conditions.
Keywords: singer identification, MFCC, ANN, SVM, feature extraction etc.

I. Introduction

Automatic singer voice identification is a subtask of for the musical content identification, automatic singer identification is important for easy to access essential data from vast amount of data. It could be treated as the first step in developing automatic musical transcription and multimedia database annotation. The use of real world recordings should be encouraged in voice recognition [3].Now days there is rapidly growth in amount of digital media, so the need for an effective data management is challenging task.

In this context, we present a study about automatically identifying musical instruments from music using classifiers. Information related musical instrumentation is among the most important semantic concepts that humans use for communicate musical meaning [4][5]. Musical instrument belongs to wide range of devices with different characteristics that includes physical aspects, different sound initiation process. So classifying musical instrument becomes a complex issue, because of multidimensional nature of musical instruments. Hence the need of automated system is arises to classify musical instruments automatically[6][7] Musical instrument identification using SVM and MLP classifiers helps to reduce manual work to identify and classify musical instruments according to their attributes.

II. Litrature Survey

Vipul Arora and Laxmidhar Behera[1] presented Probabilistic latent component analysis (PLCA) is a popular tool for decomposing the spectra of polyphonic music for identifying the constituting musical instruments. Supervised PLCA reconstructs the observed polyphonic spectra using instrument specific spectral parts. However, the distance metrics between two instrument classes in the spectral space may not be the same as that in the acoustic space. This paper proposes a novel geometrical approach to PLCA by modifying the metrics in a way so as to enhance the inter-class distances, while reducing the intra-class distances. Hence, the likelihood function of PLCA is made to give more weight to reconstructing the features which discriminate the sounds of different instruments better. Experiments on binary instrument classification give encouraging results with the proposed approach.

Swati D. Patil, Tareek M. Pattewar [2]Identifying musical instrument is challenging task because of its multidimensional nature. Our aim is to develop the system which automatically classifies musical instruments. This system describes how intend and extend are related to each other. By using the concept analysis techniques recognition process can be made less dependent on human supervision. System is further evaluated with SVM and MLP classifiers. Analysis result for SVM for correct classification is greater than MLP. This system can be used in future for ontology generation of musical instruments.

EmmanouilBenetos, Margarita Kotti, and Constantine Kotropoulos[3] proposed Large Scale Musical Instrument Identification which addressed automatic musical instrument identification using a variety of classifiers. Where training is performed for each audio class individually using classifier which is based on non-negative matrix factorization (NMF)techniques. For testing recording is projected onto several training matrices, which have been Gram-Schmidt orthogonal zed.

Tetsuro Kitahara, iMasataka[4] In this paper, present a method using. an FO-dependent multivariate normal distribution of which mean is represented by a function of fundamental frequency (FO). This FO-dependent mean function represents the pitch dependency of each feature, while the FO-normalized covariance represents the non-pitch dependency. Musical instrument sounds are first analyzed by the FO-dependent multivariate normal distribution, and then identified by using the discriminant function based on the Bayes decision rule. Experimental results of identifying 6,247 solo tones of 19 musical instruments by 10-fold cross validation showed that the proposed method improved the recognition rate at individual-instrument level from 75.73% to 79.73%, and the recognition rate at category level from 88.20% to 90.65%.

A.Azarlooet.al.[5]Proposed Automatic Musical Instrument Recognition Using K-NN and MLP Neural Networks. The author used seven different musical instruments to be played simultaneously from solos to quartets. The data of this system is 296 feature vectors that used in audio signal classification by MLP neural networks and K-NN algorithm. Finally, MLP achieved as the best neural network in musical instrument recognition.

Pamomaol Jincahitru [6] A system which tries to identify the musical instruments playing concurrently in a mixture is investigated in this paper. The features used in classification are derived from the Independent Subspace Analysis (ISA)which somewhat decomposes each source, and the mixture, into its statistically "independent" components. Without re-grouping or actually separating the sources, they offer physiologically motivated classification of instruments, assuming the decomposition is robust to the mixing process. The system is evaluated on two-tonal instrument mixtures from **a** set of five instruments and a phrase of real song from CD.

Juan Pablo Bello, Laurent Daudet [7]The goal of this paper is to review,categorize, and compare some of the most commonly used techniques for onset detection, and to present possible enhancements. They discuss methods based on the use of explicitly predefined signal features: the signal's amplitude envelope, spectral magnitudes and phases, time-frequency representations; and methods based on probabilistic signal models: model-based change point detection, surprise signals, etc. Using a choice of test cases, we provide some guidelines for choosing the appropriate method for a given application.

Jayme Garcia Arnal Barbedo George[8] A new approach to instrument identification based on individual partials is presented. It makes identification possible even when the concurrently played instrument sounds have a high degree of spectral overlapping. A pairwise comparison scheme which emphasizes the specific differences between each pair of instruments is used for classification. Finally, the proposed method only requires a single note from each instrument to perform the classification. If more than one partial is available the resulting multiple classification decisions can be summarized to further improves instrument identification for the whole signal. Encouraging classification results have been obtained in the identification of four instruments (saxophone, piano, violin and guitar).

Mari Okamura, Masanori Takehara[9]we propose a musical instrument identification method based on sparse representation. Compared with conventional methods, our proposed approach has an advantage that sound source separation is not required and sound overlapping can be treated as a combination of features obtained from musical instruments contained in the polyphonic sounds. At first, we evaluated the effectiveness of the proposed method for monophonic sounds, then the average accuracy rate of 91.9% was obtained and it was roughly as same as SVM. And in the case of using the spectrum divided by sub-band as the features, the accuracy of the proposed method higher than of SVM. It is found that the proposed method is thus effective. Secondly, we evaluated the performance for polyphonic sounds mixed with two instrument sounds.

Vipul Arora and Laxmidhar Behera[10] In this paper, we proposed a novel scheme for modeling the acoustic space, where similar sounding instruments are mutually closer in distance and the dissimilar instruments are farther apart. The main novelty of the work is to use Riemannian metrics for the PLCA framework which enhance thedifferences among the instruments for better discrimination.

The improvements in the classification accuracy shown by the experimental results are encouraging. Further investigations are needed to find better methods to estimate the discriminating directions in the feature space. The proposed binary classification approach can be used as a supplement to the general multi-instrument PLCA so as to enhance the distinction between similar sounding instruments, which are otherwise difficult to distinguish by the later alone. Also, this work can easily be extended to multi-instrument classification.

III. Methodology Used

A. SVM

SVM is strong classification algorithm because it is simple in structure and it requires less number of features. Fig.1 describes the Framework of SVM. SVM currently considered the most efficient family of algorithm in machine learning because it is computationally efficient and robust inhigh-dimension .Support vector machine is trained with dataset of musical instrument. During the training of SVM feature extractor converts each input value to feature set and these feature sets capture basic information about each input. And these sets are used for classify the feature unit.

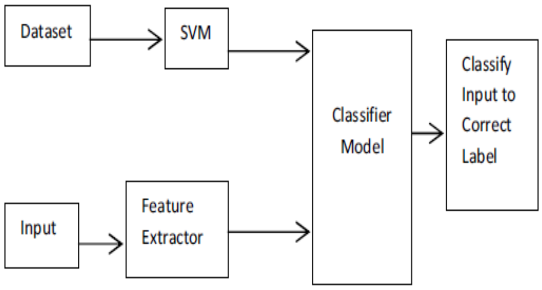

Figure. 1 Framework of SVM Classifier

B. MLP

In this system MLP is used for classification of musical instruments. It consists of two input neurons in first layer and hidden layer consist of six neurons with one output neuron. MLP uses batched back proposition algorithm. Activation functions for MLP. MLP uses supervised training on multiple layers of interconnected perceptrons. MLPs contain at least one layer of hidden neurons.

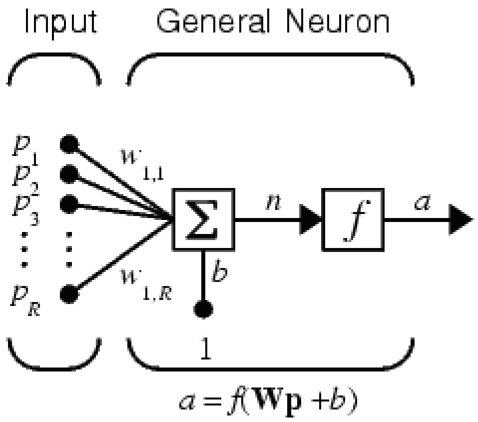

Figure 2: Multi Layer Perceptron Network

C. Feature Selection and xtraction

As commented before, each specific pair of instruments has a featureassociated with it. Therefore, the features to be extracted depend directly on which instruments are being considered. Four instruments - alto saxophone (S), violin (V), piano (P) and acoustic guitar (G) - were chosen to validate the strategy. Hence, there are six possible pairs of instruments. The instruments in the pairs SP, SG, VP and VG have considerably different characteristics (temporal waveform, spectral content, etc), while the instruments in the pair SV have some similar characteristics and the instruments in the pair PG are closely related, as discussed in [13]. In this way, the technique can be tested under different levels of difficulty.

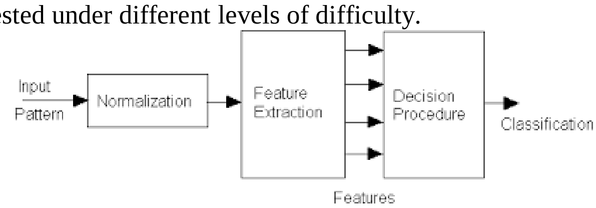

Figure3:Process of Feature Extraction

D. MFCC

MFCC are the most useful coefficients which are used for speech recognition because of their ability to represent speech amplitude spectrum in a compact form. Figure 2 shows the process of creating MFCC features [7,]. Speech signal is divided into frames by applying a hamming windowing function at fixed intervals. Cepstral feature vectors are generated using each frame.

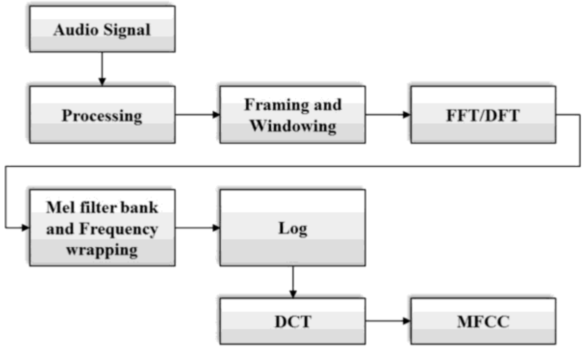

Figure 4: The process of creating MFCC

The final step is to compute the discrete cosine transform (DCT) of the log filter bank energies in MFCC. But only 12 of the 26 DCT coefficients are kept because the higher DCT coefficients represent fast changes in the filter bank energies which reduce the performance of system. So it gives small improvement by dropping them. Thirteen MFCC coefficients (×28 to ×40) are used by our proposed approach which is extracted for each segment from IVS.

E. ANN (Artificial neural network)

Singer Identification (SID) is the process of retrieving identity of the singer in a song through features of voice. The voice recognition system is most prominent technique to identify of singers voice. Every person has unique voice quality. The unique qualities of a singer's voice make it relatively easy to identify a song of particular artist. The identity of a singer can be identified by using Artificial Neural Network (ANN). The singer identification comes under speaker identification or voice biometrics.

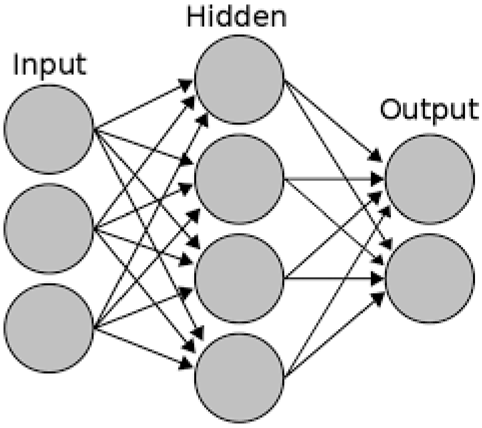

Figure 5: Artificial Neural Network

We used MFCC's and its deltas for feature extraction because of its more robustness to noise than other technique. Singer models are built from speaker features. After thisthe features are ready to feed into artificial neural network. The ANN did training and testing and classifies the input voices in different classes of singers.

IV. Importance of Research

Sound, Music, musical Instrument and Voice are very fascinating areas of research for digital signal processing applications.

1. Musical Instrument Identification [1], Singer Identification [2] [3], Speaker Recognition [4], Music Melody Extraction, Music Database Indexing are some promising areas of applications of Music Information Retrieval (MIR)research.
2. In most popular music, the singing voice section captures one of the most important characteristics, and thus detection of the singing voice has attracted a number of researchers in a music information retrieval (MIR) community for many years. Detection of singing voice or vocal/non-vocal discrimination can be used in many other related applications in MIR such as singer identification, singing voice separation, query-by-humming (QBH), structural analysis of music or lyrics-audio alignment, and so on.

3. Music applications, especially for vocal/non-vocal discrimination because it requires a great deal of human labor to manually annotate vocal/non-vocal boundaries as he/she listens to a number of audio files.

V. Disscussion

In this paper, we focused on various technique which are used in identification of singer voice. Two separate ANNs are develop for language and Speech recognition system and trained using back propagation algorithm and radial basis function network. Here use of MFCCs and delta-MFCCs as acoustic features. The present work is limited to small vocabulary and four languages. In future we will develop system for large vocabulary and more languages. The use of vocal and instrumental features and ANN can provide successful music classification. To provide benefit of Information technology in each and everywhere in India, voice based interface for various computer related task is most suitable. To develop a voice based interface for countries like India is very difficult task as here approximate twenty five languages spoken. The Speech signal conveys many levels of information like what is spoken, language, speaker, gender, sentiments etc.

VI. Conclusion

In this review, the basic techniques used in voice identification of singer are discussed and its recent researches and work is highlighted. Some approaches available for developing an Singer Identification system are clearly explained with its merits and demerits, techniques of feature extraction, model training. The performance of the Singer Identification system based on the adapted feature extraction technique and the speaker identification and recognition approach for the particular individual is compared in this paper. In recent years, the there is huge requirement on singer identification and recognition on huge datasets. In researches it has been seen that ANN approach along with MFCC features is more suitable for these requirements and offers good recognition result. Where ever the above combination will be used to identification will help to generate large powerful systems.

Refrerences

[1]. Vipul Arora and Laxmidhar Behera "Instrument Identification Using PLCA Over" Department of Electrical Engineering, Indian Institute of Technology, Kanpur 978-1-4799-2361-8/14/$31.00 c 2014 IEEEStretched Manifolds.

[2]. Swati D. Patil1, Tareek M. Pattewar "Musical Instrument Identification Using SVM & MLP withFormalConcept Analysis" Department of Computer Engineering, SES's R. C. Patel Institute of Technology, Shirpur, Maharashtra, India 978-1-4673-7910-6/15/$31.00_c 2015 IEEE.

[3]. Emmanouil Benetos, Margarita Kotti, and Constantine Kotropoulos,"Large Scale Musical Instrument Identification"Department of Informatics.

[4]. TetsuroKitahara,iMasatakaGoto,$ and Hiroshi G. Okunot "Musical instrument identification based on fo-dependent multivariate normal distribution" 0-7803-7663-3/03/$17.00 02003 IEEE.Dept. of Intelligence Science and Technology Graduate School of Informatics, Kyoto University Sakyo-ku, Kyoto 606-8501, Japan.

[5]. M. Abulaish, "Ontology Engineering for Imprecise Knowledge Managemen"t.Saarbrücken, Germany: Lambert Academic, 2008.

[6]. PamomaolJincahitru "Polyphonic instrument identification using independent subspaceanalysis" CCRMA, Stanford University Stanford, CA94305, USA pj97@ccrma.stanford.edu 2004 IEEE International Conference on Multimedia and Expo (ICME).

[7]. Juan Pablo Bello, Laurent Daudet, SamerAbdallah, Chris Duxbury, Mike Davies, and Mark B. Sandler "A Tutorial on Onset Detection in Music Signals" Manuscript received August 6, 2003; revised July 21, 2004. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gerald Schuller.IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 13, NO. 5, SEPTEMBER 2005.

[8]. Jayme Garcia ArnalBarbedoGeorge Tzanetakis "INSTRUMENT DENTIFICATION IN POLYPHONIC MUSIC SIGNALS BASED ON INDIVIDUAL PARTIALS"State University of Campinas DECOM/FEEC Campinas, SP, Brazil978-1-4244-4296-6/10/$25.00 ©2010 IEEE.

[9]. Mari Okamura, Masanori Takehara, Satoshi Tamura, and Satoru Hayamizu "Toward Polyphonic Musical Instrument Identification using Example-based Sparse Representation" Mari Okamura, Masanori Takehara, Satoshi Tamura, and Satoru Hayamizu Department of Information Science, Gifu University, Gifu, Japan.

[10]. Vipul Arora and Laxmidhar Behera "Instrument Identification Using PLCA Over Stretched Manifolds" Department of Electrical Engineering, Indian Institute of Technology, Kanpur 978-1-4799-2361-8/14/$31.00 c 2014 IEEE.

[11]. Corneliu Octavian DUMITRU, Inge GAVAT, "A Comparative Study of Feature Extraction Methods Applied to Continuous Speech Recognition in Romanian Language", 48th International Symposium ELMAR-2006, 07-09 June 2006, Zadar, Croatia.

[12]. DOUGLAS O'SHAUGHNESSY, "Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis", Proceedings of the IEEE, VOL. 91, NO. 9, September 2003, 0018-9219/03$17.00 © 2003 IEEE.

[13]. D. Fragoulis, C. Papaodysseus, M. Exarhos, G. Roussopoulos, T. Panagopoulos, and D. Kamarotos, "Automated classification of piano-guitar notes," IEEE Trans. Audio Speech Lang. Process., vol. 14, pp. 1040–1050, 2006.