

Study of suitable Audio Feature Extract and Classification Methods to be used for Indian Classical Music's Singer Identification

Viraj Jamle¹, Sourabh Deshmukh²

¹Pune University, G.H.R.C.E.M, Wagholi, Pune, Maharashtra, India

²Professor, Pune University, G.H.R.C.E.M, Wagholi, Pune, Maharashtra, India

Abstract: *Singer identification is one of the most important applications of Music Information Retrieval (MIR). The process starts with identifying first, the audio descriptors then using these feature vectors, as input, for further classification/ identification of singers using typical classifiers, such as Gaussian Mixture Model (GMM) or Hidden Markov Model (HMM). In this paper, we propose, a Hybrid method of selecting correct audio descriptors for the identification of singer of Indian Classical Music. First, only strong (primary) audio descriptors are released on the system in forward pass and the classification impact is recorded. Then only selecting the top few audio descriptors, having largest impact on the singer identification process, are selected and rest are eliminated in the backward pass. Then selecting and releasing all the less significant audio descriptors from the groups, that had maximum impact on singer identification process, increases the success of correctly identifying the singer. The method reduces substantially the large number of audio descriptors to few, important audio descriptors. The selected audio descriptors are then fed as input to further classifiers.*

Keywords: Indian Classical Music, MPEG 7 standards, Music Information Retrieval (MIR), Audio Descriptors, Timbre, MFCC.

1. Introduction

Singer identification is one of the most important applications of Music Information Retrieval (MIR). The process starts with identifying first, the audio descriptors then using these feature vectors, as input, for further classification/ identification of singers using typical classifiers, such as Gaussian Mixture Model(GMM) or Hidden Markov Model(HMM).

The process becomes chaotic if all audio descriptors are used for finding the feature vector; instead if, the audio descriptors are selected with respect to the kind of input data then the process becomes comparatively simple. In this paper, we propose, a Hybrid method of selecting correct audio descriptors for the identification of singer of North Indian Classical Music.

First, only strong (primary) audio descriptors are released on the system in forward pass and the classification impact is recorded. Then only selecting the top few audio descriptors, having largest impact on the singer identification process, are selected and rest are eliminated in the backward pass. Then selecting and releasing all the less significant audio descriptors from the groups, that had maximum impact on singer identification process, increases the success of correctly identifying the singer.

The method reduces substantially the large number of audio descriptors to few, important audio descriptors. The selected audio descriptors are then fed as input to further classifiers.

2. Existing System

Audio feature extraction addresses the analysis and extraction of meaningful information from audio signals, in order to obtain a compact and expressive description that is machine-process able [3].

It is observed that the audio features or audio descriptors that are extracted, derived and used for a particular task and domain are often tried for other applications and other domains. We restrict here in this paper to, the classification of almost all important audio descriptors for singer identification domain only, especially for north Indian classical vocal.

Audio descriptors are special attributes or characteristic features of the audio segment under consideration. There is a wide variety of these descriptors and identifying them is an essentially first step towards analysis of the audio sample.

3. System Architecture

Solving Approach

We are trying to design the singer identification based on Timbre and audio descriptor. The singer identification process uses three modules viz. an Input module, a Query Module and a Retrieval Module [17]. The algorithm of singer identification uses audio data filtering and standardization as first step in the Input module. By making use of MIR toolbox functions the audio features are extracted in the feature extraction process.

This paper elaborates various audio feature extraction methods applied till now to the best of our knowledge and the classifiers used with the analysis of their input environment, the constraints on the system and the results been generated in controlled result space. Here a comprehensive analysis of various audio features, methods of feature extraction and classification techniques with results is done. Here we have treated human voice, a kind of musical instrument, so that all audio related feature extractors, especially timbre, can be compared.

Proposed System

The singer identification models work in three modules viz. input module (feature extraction module), query module (training and testing) and a classification module (singer identification)[12]. The input audio files have various attributes such as file type (.wav, mp3), sampling rate (44.1k, 11.0 kHz), audio type (mono, stereo) bit rate etc. Some standard feature extraction methods such as Linear Predictive coding (LPC), Mel Frequency Cestrum Coefficient (MFCC), Wavelet Transform (WT), Fourier Transform (FT) etc. are applied. This gives various features of the audio sample. These are audio descriptors. These descriptors are either scalar (one dimensional) or vector (multi-dimensional) quantities.

In order to identify the audio descriptor(s) responsible for detecting the gender of a singer following algorithm is applied. The algorithm divides the process into two parts.

Part A: identification of the singer and calculating the accuracy of complete singer identification process and Part B: backward detection and analysis of audio descriptors responsible for correct singer identification. We use two datasets DB1 and DB2, each containing studio, noise free recordings of popular North Indian classical singers.

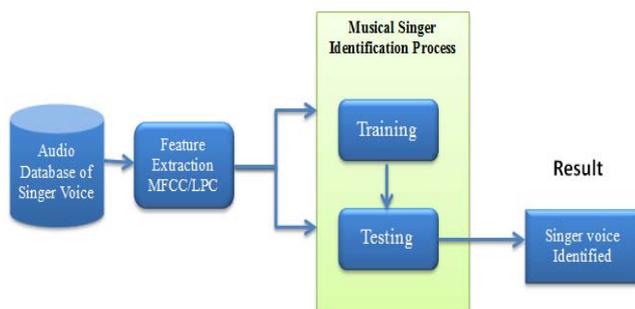


Figure 1: Proposed System block diagram of Musical Singer identification

Above figure shows five elements in design phase of Singer Identification Process.

Generally, any singer identification system works in following way. There are usually three modules viz an Input module, a Query Module and a Retrieval Module. In input module, from audio files, various audio features of the singer are extracted and stored in a feature database. When the query module come up with a query of whether the singer data in hand is known or known, the retrieval module gets a query along with feature set of singer to be identified. The

Retrieval Module then uses some similarity comparison method or a set of such modules and then gives a feedback, whether there is a match available in the feature database or not. As shown in Fig 2.

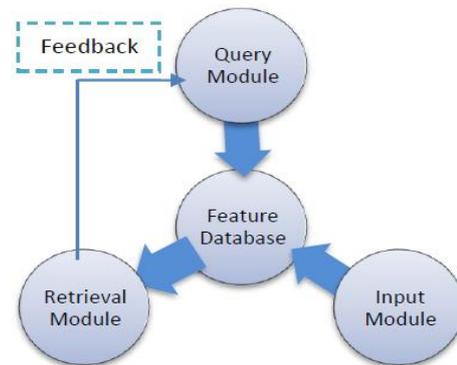


Figure 2: General Structure of Singer Identification Process

Input: I is set of voice contents (eg. Singer 1 voice, Singer 2 Voice)

Output: O is output of dataset
 $O = \{y: y \in \text{correct voice of singer}\}$

Process:

A is set of Singer Voice i.e. Audio DB
 $A = \{a: a \in \text{Input taken from Dataset}\}$
 F is set of Feature Extracts (eg. MFCC, LPCC, Timber)
 $F = \{e: e \in \text{feature extract selection}\}$
 C is set of Classifier (eg. KMEANS, 1NN, 2NN, 3NN, M-SVM)
 $C = \{c: c \in \text{classifier selection}\}$

Proposed System:

$I = \{A, F, C\} \in O$
 $F \cap C \{ \{MFCC, KMEANS\}, \{MFCC, 1NN\}, \{MFCC, 2NN\}, \{LPCC, KMEANS\}, \{LPCC, 1NN\}, \{LPCC, 2NN\}, \{Timber, KMEANS\}, \{Timber, 1NN\}, \{Timber, 2NN\} \}$

To derive LFPC the signal frames are passed through a bank of 12 band-pass filters spaced logarithmically, and the coefficient for each band is obtained by computing the power of the band divided over the band bandwidth and expressed in decibels [New et al., 2004], as follows,

$$LFPC_t(m) = 10 \log_{10} \left| \frac{S_t(m)}{N_m} \right|, \quad S_t(m) = \sum_{k=f_{m-1}}^{f_m} X_t(k)^2, \quad m = 1, 2, \dots, 12$$

4. Advantages & Disadvantages

4.1 Advantage

- 1) It provides the highest accuracy by multiple combinations of Classifiers and feature extracts.
- 2) It provides correct and efficient classifier is used along with it.

4.2 Disadvantage

- 1) To design this combination required lots of efforts for complex combinations.
- 2) Noisy data may badly impact on results obtained.

"International Workshop on Artificial Intelligence and Music"(IJCAI-07), Hyderabad, India, 2007

5. Conclusion

The audio descriptor selection is a crucial process in identifying a singer in given database of singers. As much as, to identify correctly a classical singer from a given database, the quality of input is important, it is also important that correct set of audio descriptors should be used. To classify these audio descriptors special attention has to be given to the kind of input given and the kind of final classifier is used.

In case of identifying a singer from North Indian classical music, the traditional audio descriptors selection method does not consider the aspects of the specialties of the stylizations in this type of music. With the help of this hybrid selection method the descriptor size drastically reduces. In this a straight approach of selecting strong descriptors at the first place, releasing them and then finding the most important descriptor classes, then backward eliminating the less important and again firing all the less relevant audio descriptors from the important groups back leads to a successful singer identification in North Indian Classical Music.

6. Acknowledgment

We would like to thank all the authors of different research papers referred during writing this paper. It was very knowledge gaining and helpful for the further research to be done in future.

References

- [1] A Hybrid Selection Method of Audio Descriptors for Singer Identification in North Indian Classical Music. Saurabh Deshmukh¹, Dr. S.G. Bhirud² Head IT Dept, GHRCEM, Pune, India¹, Professor, Computer Engineering Department, VJTI, Mumbai, India²
- [2] "North Indian classical Music's singer identification by Timbre recognition using MIR toolbox"
- [3] B. Whitman G. Flake and S Lawrence, "Artist detection in Music with Minnow- Match," in proceedings of the 2001 IEEE workshop on Neural Network for Signal Processing, Falmouth ,MA, 2001, pp. 559-568.
- [4] "Analysis and application of suitable audio feature extraction and classification method to be used for North Indian Classical Music's singer identification problem"
- [5] Kumar Rakesh, Subhangi Dutta, and Kumara Shama, "GENDER RECOGNITION USING SPEECH PROCESSING TECHNIQUES IN LABVIEW," International Journal of Advances in Engineering & Technology, vol. Vol. 1 , no. Issue 2 , pp. pp.51-63, May 2011.
- [6] A. Bapat, V. Rao and P. Rao, "Melodic contour Extraction from indian Classical Vocal Music," in