ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

# Musical Source Clustering and Raga Identification

Krishnapriya S, Anju.L,

Master of Engineering in Applied Electronics, Sri Venkateswara College of Engineering, Sriperumbudur, Tamil Nadu, India

Assistant Professor, Sri Venkateswara College of Engineering, Sriperumbudur, Tamil Nadu, India

**ABSTRACT**: The role of raga in Indian classical music is never ending. Raga is a collection of different unique notes (swara) with some special properties .These notes helps us to identify the particular raga as well as the musical instrument played in a given polyphonic audio .In the proposed paper, raga identification is done for Carnatic/South Indian Classical music. Here feature extraction is done using short and midterm feature extraction function. K-Nearest-Neighbor (kNN) algorithm is used to classify the features as it does the estimation of number of clusters in a given dataset. Finally the raga is labelled using Support Vector Machine (SVM), a supervised learning algorithm. The main motive behind raga identification is that it can be used as a good basis for music information retrieval of any Carnatic music.

**KEYWORDS**: Indian Classical Music, kNN Algorithm, Support Vector Machine

## I.INTRODUCTION

Computational Musicology is an emerging field that includes the study of techniques for analyzing various kinds of music.It helps to study the science behind music and develops a scientific framework for this art.The analysis of music helps to understand the culture and heritage of the society the music is evolved in.Indian Classical Music(ICM) is broadly classified into Hindustani and Carnatic music.Though these two music types have same base and similar framework,they differ in many factors.

Raga is considered to be the backbone of Indian Classical Music.It is a collection of different unique notes that are having some special properties (e.g. arohana , avarohana , Pakad , Taal,etc).Raga identification has been an effortful task for a long time.The difficulty mainly relies on the following aspects. First, the waveform or given polyphonic audio signal is too complex for computer to process directly.Second, the extracted features has to be classified and finally it has to be labelled with the particular raga identified. In the proposed paper the raga features are extracted using short and mid-term feature extraction function. K-Nearest-Neighbour (kNN) algorithm is used for classifying the audio features. kNN is the simplest algorithm which can be used for binary and multi-level classes. The algorithm used for classifying and labelling is Support Vector Machine (SVM), a supervised learning algorithm .Development of music recommendation systems, automatic note transcription, music indexing on-line teaching ,learning of music e.t.c are some of the applications of automatic raga identification

This paper is organized as follows: Section II deals with feature extraction using short term and mid-term feature extraction procedure. Section III refers about feature classification. Section IV tells about support vector machine and its implementation methods. Section V introduces experiment and evaluates the performance of the proposed system. Section VI concludes the scope and future work.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

**International Journal of Innovative Research in Computer and Communication Engineering**

*(An ISO 3297: 2007 Certified Organization)*

Website: [www.ijircce.com](http://www.ijircce.com)

Vol. 5, Special Issue 3, April 2017

## II.FEATURE EXTRACTION

In order to extract the features, mainly two feature extraction methods are used-Short term and Mid-term feature extraction.

*A.Short-term feature extraction*

The given input is divided in to possibly overlapping short term frames, and the features are extracted from each frame. This type of processing generates a sequence, F, of feature vectors per audio signal. The feature vector dimensionality depends on the nature of the adopted features. Single dimensional and multidimensional features are used for sophisticated audio analysis. Here 23 audio features are extracted for audio analysis by breaking the audio input into short-term windows and computing 23 audio features per window.

*B. Mid-term feature extraction*

In mid-term feature extraction, processing of feature sequence is done on a mid-term basis. According to this method of processing, the input signal is first divided into mid-term segments (windows) and then, for each segment the short-term processing stage is carried out. At a next step, the feature sequence, F, which has been extracted from a mid-term segment, is used for computing feature statistics.

Example, the average value of the zero-crossing rate. In the end, each mid-term segment is represented by a set of statistics which correspond to the respective short-term feature sequences. During mid-term processing, here assume that the mid-term segments exhibit homogeneous behavior with respect to audio type and it therefore makes sense to proceed with the extraction of statistics on a segment basis. In practice, the duration of mid-term windows typically lies in the range 1–10 s, depending on the application domain.

The next process takes as input short-term feature sequences that have been generated and returns a vector that contains the resulting feature statistics. If, for example, 23 feature sequences have been computed on a short-term basis and two mid-term statistics are drawn per feature (e.g. the mean value and the standard deviation of the feature), then, the output of the mid-term function is a 46-dimensional vector. The structure of this vector is the following: elements 1 and 24correspond to the mean and standard deviation of the first short-term feature sequence, elements 2 and 25 correspond to the mean and standard deviation of the second audio short-term sequence, and so on.

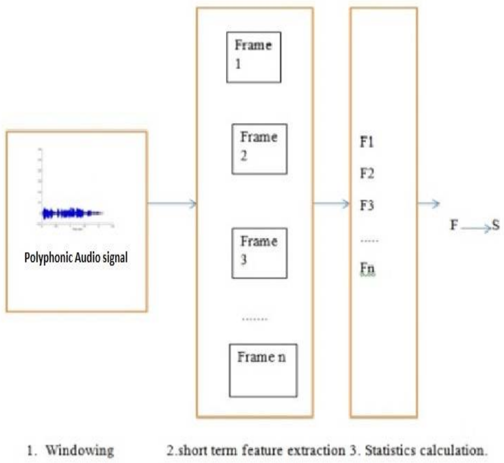


Fig 1 Mid-term feature extraction

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

## III. FEATURE CLASSIFICATION

In the proposed paper, for feature classification kNN classifier is used.

k-Nearest-Neighbor Classifier

One of the simplest classifier in machine learning and Data mining is kNN. Despite of its simplicity kNN is well suited for both binary and multi-class problems.The highlight feature of KNN is that it does not require a training stage in the strict sense. When the test pattern (unknown feature vector), x , is given, we first detect k-nearest neighbors in the training set and count how many of those belong to each class. In the end, the feature vector is assigned to the class which has accumulated the highest number of neighbors.

## IV.SUPPORT VECTOR MACHINE (SVM)

Support vector Machine is a supervised learning algorithm. SVM is used mainly in Machine learning and Data mining for classification and regression analysis. This algorithm is based on the key observation that, for the simple case of linearly separable classes, the optimal decision hyperplane (line on the two dimensional space) is the one that maximizes the 'margin' among the training data of the two classes. To this end, supporting hyperplanes are used, parallel to the decision hyperplane, in order to define the margin that minimizes the classification error. This approach is extended to cover the non-separable case by permitting misclassified training data.

*Main parameters of SVM are*

Type of kernel : According to the SVM methodology , a kernel function is used in order to map the feature vectors to the 'kernel space'. Typical kernals are linear, polynomial (of order 3 or higher), and radial basis function (RBF) kernel, etc.

Kernel properties: depending on the selected kernel, certain kernel parameters need to be set for the training and evaluation of the SVM classifier

Constraint parameter, C. It is related to the cost function of the SVM training procedure. As the value of C increases, the cost of misclassified samples also increases. Care must be taken so that overfitting must be avoided.

The support vector machine function do the following steps while executing

➢ Loads the dataset of feature vectors

➢ Randomly splits the dataset into two equally sized subsets, one for testing and one for training .Validation techniques are then employed to achieve better reliable performance measurements, the random sub-sampling step has to be repeated several times.

➢ At last computes the classification accuracy, i.e. the fraction of the correctly classified samples.

A. Performance measures

Confusion matrix is an important tool for analyzing the performance of binary class methods. This provides the means to group the classification results into a single matrix and helps the designer of the classifier to understand the types of errors that occur during the testing and training stage. The confusion matrix, CM, is a Nc x Nc matrix, whose rows and columns refer to the true (ground truth) and predicted class labels of the dataset, respectively. In other words, each element, CM (i , j), stands for the number of samples of class i that were assigned to class j by the adopted classification method. It follows that the diagonal of the confusion matrix captures the correct classification decisions (i=j).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Special Issue 3, April 2017

It is useful to normalize the confusion matrix so that its elements become probabilities and not simple counts of events. This can be basically achieved in two ways, the first of which is to divide each element of CM by the total number of samples in the dataset, i.e. by the sum of the elements of the confusion matrix is given by the equation

..... (4.1)

The second class-specific performance measure is class precision, Pr(i), which is defined as the fraction of samples that were correctly classified to class i if we take into account the total number of samples that were classified to that class. Class precision is, therefore, a measure of accuracy on a class basis and is defined according to the equation

........ (4.2)

Where the quantity in the denominator, stands for the total number of samples that were classified to class i.

Finally, a widely used performance measure that combines the values of precision and recall is the F1-measure, which is computed as the harmonic mean of the precision and recall values.F1 is given by the equation

.......... (4.3)

*B.Cross-Validation*

Two cross-validation methods are used to check the accuracy.

1. Repeated hold-out validation: to avoid overfitting, the hold-out method partitions the dataset into two non-overlapping subsets: one for training and the other for testing. For example, one third of the samples are used for testing.

2.Leave –one-out validation: The leave-one-out method is actually a variation of the k-fold cross validation approach, where k = M, i.e. the number of folds is equal to the total number of samples available in the set. In other words, each fold consists of a single sample. Therefore, during each iteration, all the samples, apart from one, are used for training the classifier and the remaining sample is used in the testing stage. The leave-one-out method is an exhaustive validation technique that can produce very reliable validation results.

## V.EXPERIMENTAL RESULTS

*A. Dataset*

A collection of carnatic ragas with audio and instrumental clips has been collected from ‘carnaticcorner’ database, an authorised database website.

Raga audio samples are also collected from an open Indian Classical Music database called ‘Kosha”

B. Feature Extraction:

Features extracted from audio files using short-term and mid-term feature extraction methods. After calculating mid-term feature extraction statistical measurement of features also computed and plotted.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

**International Journal of Innovative Research in Computer and Communication Engineering**

*(An ISO 3297: 2007 Certified Organization)*

Website: [www.ijircce.com](http://www.ijircce.com)

**Vol. 5, Special Issue 3, April 2017**

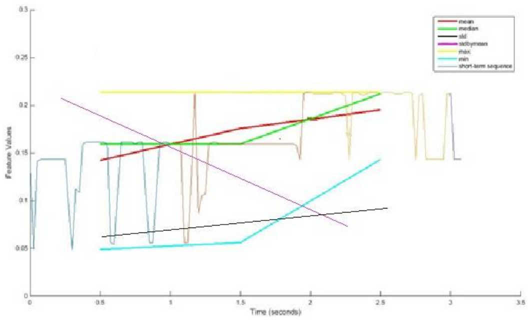

Figure 2 Statistical feature measures of Harikamboji raga.

*C. Audio Classification*

Using the extracted audio features they are classified in to classes. Here for simplicity 2 ragas are chosen and which is shown in table.

| Name | Description | Class | Samples per class |
|---|---|---|---|
| Ragas | Harikamboji Vs Charukeshi | 2 | 120 |

Table 1 : Harikamboji Vs Charukeshi

*D Raga labelling*

Raga labelling is done using support vector machine (SVM) classifier. Support Vector Machine algorithm is applied to the classified features and get the labeled features. The labelling is done with different values of cost function C to understand how the cost function will affect the classification.

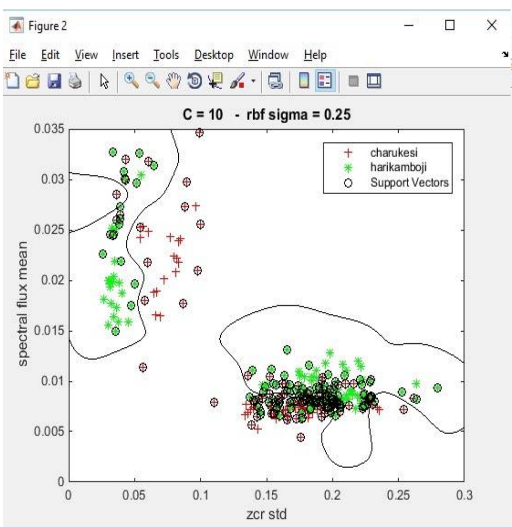


Figure 3 SVM labelling with cost function =10

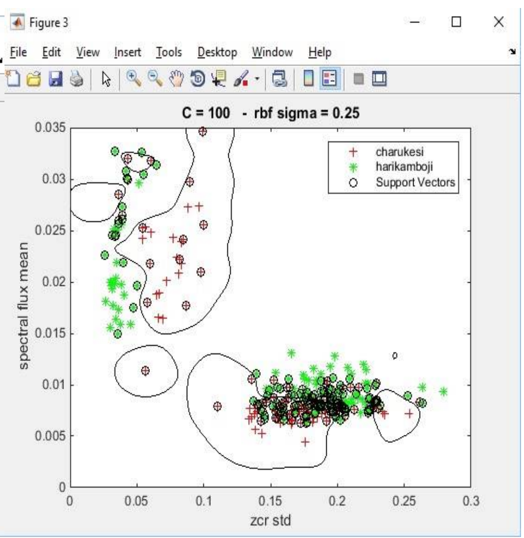


Figure 4. SVM labelling with cost function C=100

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

**International Journal of Innovative Research in Computer and Communication Engineering**

*(An ISO 3297: 2007 Certified Organization)*

*Website: [www.ijircce.com](http://www.ijircce.com)*

**Vol. 5, Special Issue 3, April 2017**

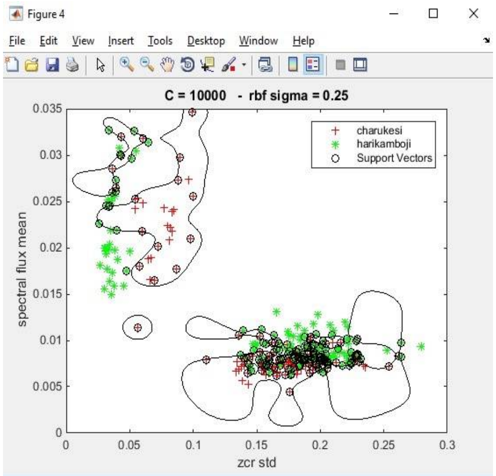


Figure 5 SVM labelling with cost function c=10000

From the plot we can understand that while C is increasing overfitting problem will happen. It can be seen that, for very high C values, the resulting classification scheme manages to classify correctly almost every sample of the training set. However, the classification accuracy on the test set decreases considerably and shows a performance accuracy of 96.15%.

## VI. CONCLUSION AND FUTURE WORK

In the proposed paper,the features are extracted using short-term and mid-term feature extraction methods. Feature classification of raga classification is done using kNN classifier. Raga labelling is done using Support Vector Machine (SVM), a supervised learning algorithm with an accuracy measure of 96.1%.

In future, musical source clustering along with raga identification of a given polyphonic audio will be done simultaneously with the help of Artificial Neural Networks.

## REFERENCES

1.S Suma,Shashidhar G.Koolagudi :Raga Classifiaction for Carnatic Music;Springer-2015.
2.Paul Sgeba,L.L.M;Dept of ECE,Saranathan College of Engineering,Nagar India,Revathy A;Raga Identificaton using clustering algorithm;IEEE-2014.
3.Automatic musical instrument classification using fractional Fourier transform based-MFCC features and counter propagation neural network D.G.Bhalke 1& C.B.Rama Raol, & D.S.Bormanae ;16 April 2015 #Springer Science+Business Media New York 2015.
4. Musical Source Clustering and Identification in Polyphonic Audio ;Vipul Arora and Laxmidhar Behera,Senior Member,IEEE-2014.