

# Raga Classification for Carnatic Music

S.M. Suma and Shashidhar G. Koolagudi

**Abstract** In this work, an effort has been made to identify *raga* of given piece of Carnatic music. In the proposed method, direct *raga* classification without the use of note sequence has been performed using pitch as the primary feature. The primitive features that are extracted from the probability density function (pdf) of the pitch contour are used for classification. A feature vector of 36 dimension is obtained by extracting some parameters from the pdf. Since non-sequential features are extracted from the signal, artificial neural network (ANN) is used as a classifier. The database used for validating the system consists of 162 songs from 12 *ragas*. The average classification accuracy is found to be 89.5 %.

**Keywords** Carnatic music · *Raga* · Machine learning · Artificial neural network · Pitch

## 1 Introduction

Computational Musicology is an emerging field that includes the study of techniques for analyzing various kinds of music. It helps to study science behind music and develops a scientific framework for this art. The analysis of music helps to understand the culture and the heritage of the society the music is evolved in. Most of the studies in this area are dedicated to Western Music whereas very fewer studies have tried to explore Indian Classical Music (ICM) [1, 2]. 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. ICM is characterized by two basic elements—it should follow a specific note sequence known as *raga* and a specific

---

S.M. Suma (✉) · S.G. Koolagudi  
National Institute of Technology Karnataka, Surathkal, Karnataka, India  
e-mail: suma.cs13f02@nitk.edu.in

S.G. Koolagudi  
e-mail: koolagudi@nitk.edu.in

rhythm known as *Taal* [3]. *Raga* is considered to be the backbone of ICM. In a broader sense, *raga* can be understood as the framework that tells all the protocols that should be followed while rendering the piece of music. Since *raga* is the crucial element in the melodic framework of ICM, automatic *raga* identification is one of the important steps in Computational Musicology as far as ICM is considered. Some of the applications of automatic *raga* identification are: development of music recommendation systems, automatic note transcription, music indexing, on-line teaching and learning of music and so on. It may be considered as the first logical step in the process of creating computational methods for ICM.

Some of the characteristics of *raga* are *swara* (notes), *arohana-avarohana* pattern and characteristic phrases. Technically, a note can be defined as an identifiable fundamental frequency component (pitch) of a singer with a beginning and an ending time, of a duration determined by these [2]. The ratio of the fundamental frequencies of two notes is referred to as an interval [2]. The notes are categorised as *Shadja* (*Sa*), *Rishaba* (*Ri*), *Gandhara* (*Ga*), *Madhyama* (*Ma*), *Panchama* (*Pa*), *Dhaivatha* (*Da*) and *Nishadha* (*Ni*). These seven notes are further classified into different types based on the interval system (scale) used. There are three kinds of scales that are generally used in Carnatic and Hindustani music theory: a 12-note scale, a 16-note scale and the scale which contains 22 microtones. Each *raga* is defined using unique sequence of notes that is used in rendering the music clip. Few *ragas* share same set of notes, but they differ in the melody because of different *arohana-avarohana* pattern and characteristic phrases. *Arohana-avarohana* pattern refers to ascending and descending progression of notes. This tells about the transition between the notes that can occur in a given *raga*. Characteristic phrase is defined as a condensed version of the characteristic arrangement of notes, peculiar to each *raga*, which when repeated in a recital enables a listener to identify the *raga* being played [3]. Other than these important features, *raga* is also characterised by the relative position, strength and duration of notes and also *gamakas*. *Gamakas* are the beautification elements that tell about the dynamics of the pitch contour.

Unlike western music where the frequency of notes is fixed, ICM gives liberty of choosing the frequency of base note to the artists while performing. Western classical music uses equal tempered notes whereas just-intoned notes are used in ICM. In equal tempered scale all notes are equally spaced i.e. all notes have the same frequency ratio with respect to its previous note. But in just-intonation scale the notes are not equally spaced. In ICM, no two performances of the same *raga* by the same singer would be same. ICM is relatively unexplored because of its complicated grammars, extensive use of *gamaka* and the liberty given to singer for improvising while performing. All these complexities in analysing ICM, left it as relatively less explored computational musicology. In ICM, the sequential note information plays an important role in construction of melody unlike chords and rhythm which play an important role in western music. Hence, many of the methods proposed in the past [3–5] have used Hidden Markov Model (HMM) for identification of *ragas*.

In this work, an effort has been made to identify *raga* without the knowledge of the scale using low-level features. Pitch and its derivatives are used as features for

identification of the *raga*. Various kinds of musical clips from both *Janaka* (*ragas* with all 7 notes) and *Janya ragas* (*ragas* having 5 or 6 notes) are used for the study.

The paper is organised as follows. A review of related work is given in Sect. 2. Section 3 explains the proposed method. Experimentation and results along with explanation of database are given in Sects. 4 and 5 concludes the work with some further research directions.

## 2 Literature Review

Since *raga* gives melodic framework attained through pitch manipulation, many research contributions use pitch and its derivatives as the major features for the task of *raga* identification. Some of the works [6, 7] have also used timbre and other features along with pitch derivatives. The literature can be broadly classified into two categories. The first, the works that have used note information explicitly for identification of *raga* and the second, the works that have not used note information. We discuss both kinds of work in this section. One of the main challenges in reviewing the works is lack of standard corpora of ICM. The works that have been carried out are evaluated using different databases collected by the researchers. Hence while reviewing the work the databases considered are also discussed.

In the work [3], HMM is used for the identification of Hindustani *ragas*. The method proposed in this work uses note sequence as a feature. Two heuristics namely Hillpeak heuristics and Note duration heuristics are proposed for note transcription. Many micro-tonal variations present in the ICM makes the note transcription a challenging task even for a monophonic piece of music. The heuristics proposed in this method tries to overcome these variations. Hillpeak heuristics considers the changes in the sign of the slope of the pitch contour for identification of presence of a note. Note duration heuristics assumes that a note sustains for the period of at least 25 ms and estimates the note for every 25 ms. Another main contribution of this work is characteristic phrase analysis. A n-gram model is proposed for identification of repetitive sequence of notes and based on this *raga* is identified. The overall *raga* identification accuracy is claimed to be 77 %. The main drawback of the work is limited data set and accuracy of note transcription. Only two *ragas* are used for validating the system. The database has many limitations such as the musical notes in *arohana-avarohana* sequence are sung externally, the piece of music should be in G-sharp scale. The accuracy of note transcription is not appreciably high even after lot of constraints on the dataset. Similar work has been carried out by Arindam et al. [4]. The note transcription of the given musical piece is manually performed and this sequence is used for evaluating HMM. It is observed that HMM is 100 % accurate in identifying *ragas* if correct note sequence is given as input. However, it is difficult to achieve high accuracy in ICM transcription because of the micro-tonal variations and improvisations. Rajeswari et al. [8] have used *tala* information to identify the number of notes in the given musical piece, further this is used for identification of *ragas*.

Surendra et al. [9] have used similar method of identifying the *raga* in Carnatic music using the note sequence information. Since the data set is restricted to single scale identifying the notes is simpler task. The features extracted are the notes present, number of notes present and type of *arohana-avarohana* sequence. *Arohana-avarohana* patterns are classified as linear or nonlinear based on the pair of notes present in the sequence. They have used ANN as a classifier. The system is tested using 20 *ragas* with 3–5 songs in each *raga*. The paper has claimed 95 % of accuracy in *raga* identification. Ranjani et al. [10] have proposed a method for identification of tonic and other notes which leads to identification of *raga* from the Carnatic music. The property of the tonic note with pitch of highest mean and least variation shown in pitch histogram, is used for identification of tonic pitch value. This method estimates the probability density function (pdf) from the pitch contour. Using Semi-Continuous Gaussian Mixture Model (SC-GMM), tonic frequency and *raga* of the musical piece are identified. The data set used for this experiment consisted of 5 *Sampurna ragas* which is the major limitation. This paper has claimed 86.3 % of accuracy in *raga* identification. However, the experiments conducted were limited only to *arohana-avarohana* sequences. The main drawback of the method that use note information explicitly is that any error in identification of notes, contributes to the error in identification of *raga* as well.

The works that do not consider note information, mainly concentrate on variations in pitch histogram for *raga* classification. The pitch histograms give better results for stable regions of smaller input clips compared to longer duration inputs. In [11], spectrally derived tone profiles are used as features. A tone profile is a histogram of note values weighted by duration [11]. The tone profile is obtained by taking the DFT of the segment and summing up the energies of the bins surrounding the note centers. It is further normalised using tonic frequency. The music clips were classified using k-NN algorithm. This paper has claimed 100 % of accuracy in Hindustani *raga* identification. However this method cannot be generalised because the data set consisted of very few *ragas* and the music clips were rendered by single artist. Parag et al. in [12] have used Pitch Class Distribution (PCD) and Pitch Class Dyad Distribution (PCDD) as features. PCD is obtained directly from the pitch histograms, that is normalised using the input tonic value and folded into one octave. PCDD is a note bi-gram which is called dyads in musical terms. Dyads are calculated from rough estimation of notes using pitch onsets. This paper has claimed 97 % *raga* recognition accuracy. The system is tested using the data set consisting of instrumental music clips played by a single artist in 17 *ragas*. In the work of Koduri et al. [13], the predominant pitch from vocal region of polyphonic music is used for identification of *ragas* from Carnatic music. The *raga* classification is performed using pitch histogram weighted by duration and frequency of occurrence as the features. The *raga* classification is performed using k-NN classifier. The experiment is performed on the data set consisting of 176 tunes from 10 *ragas* consisting of both vocal and instrumental music. This method is claimed to give 75.6 % accuracy. Pranay et al. [6] have used the concept of *vadi swara* (the dominant note for the *raga*) for identification of *raga* in Hindustani music. From the studies, it is shown that Hindustani note intervals are

nearer to equal tempered than just intonation intervals [14]. In [6], ICM notes are associated with Western note intervals for extracting features. This method uses chromagram features along with MFCCs and timbre features for identification of *raga*. A chromagram is a visual representation of energies in the 12 semitones (or chromas). The most repeating frequency component or semitone is assumed as the *vadi*. The classification is performed using HMM and GMM. 97 % accuracy is claimed using this method. Even though the results are promising, it cannot be generalised since the system is evaluated using only 4 *ragas*. The *vadi swara* for different *ragas* cannot be unique when a larger data set with more than 20 *ragas* is considered.

In this work, an effort has been made to overcome some of the limitations that exist in the past methods. The proposed method considers a comprehensive data set consisting of variety of music clips from different kinds of *ragas*. We try to eliminate the explicit note identification and scale dependency (tonic identification).

### 3 Proposed Method

In the proposed method pitch and its derivatives are used as features for identification of *raga* since melody is the result of variations in the pitch. From [7], it is evident that the pitch histogram for different *ragas* differ. Hence in this study we use some of the parameters obtained from the pdf of pitch contour for identification of *raga*. The flow diagram of the proposed system is shown in Fig. 1. In training phase, the extracted features and corresponding *raga* labels are used to construct a ANN model for classification. In testing phase, the feature vectors are fed to trained ANN model for obtaining the *raga* labels. The accuracy of the system is obtained by comparing them against the actual *raga* labels.

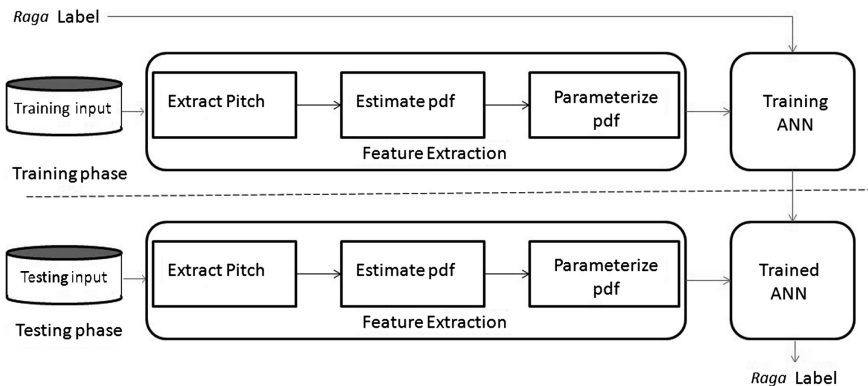
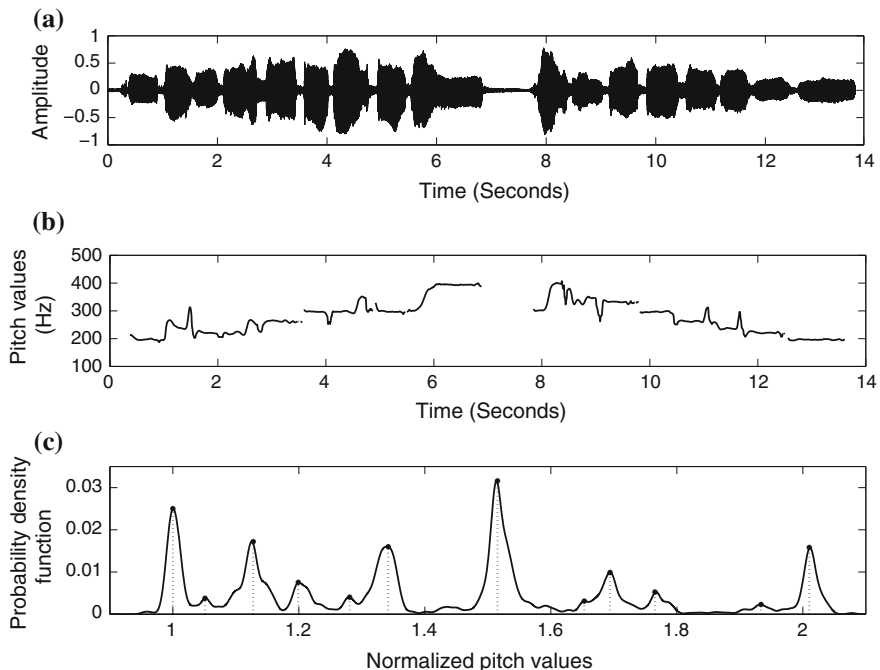


Fig. 1 Flow diagram of the proposed system

### 3.1 Feature Extraction

The pitch contour is extracted from the input signal (piece of music) using auto-correlation method [15]. The average pitch values are obtained for the frame of every 50 ms with 50 % overlapping to avoid transition cliques. The pitch contour of *arohana-avarohana* sequence is shown in Fig. 2b. Different pitch values corresponding to different notes are rendered in the piece of music. A sharp rise and fall in the contour is observed when the transition from a note to another takes place. From the pitch values, pdf is obtained using kernel density estimation method (Gaussian kernel is assumed). Bandwidth range is taken into account from minimum and maximum pitch values. The pdf obtained is shown in Fig. 2c. One can observe that the prominent peaks correspond to the notes present in the *raga*. Even though Carnatic music is defined using 16-note interval system, only 12 distinct frequency components are present [14]. Each *raga* is defined using 5, 6 or 7 note combination. Hence from the pdf curve, parameters of the 12 prominent peaks are considered for characterizing the *raga*. From the pdf, frequency of 12 peaks are obtained and their mean (height of the peaks in pdf) and variances are calculated. Variance of the peak is calculated by considering 20 frequency values on both side of the peak frequency. Since the database consists of clips rendered by different



**Fig. 2** Feature extraction for *raga Ananda Bhairavi*. **a** A music piece rendered in *raga Ananda Bhairavi*. **b** Pitch contour of the input music clip. **c** Peak selection from the pdf of pitch contour

singers in different scale, there is a need to normalise the frequencies of peaks. The frequency that lies in between 100 and 200 Hz having the highest mean is used as a normalising factor. For some clips played by instruments, all the frequency components lie above 200 Hz. In such cases, the frequency having highest mean is used as a normalising factor. All 12 prominent frequencies are divided by the normalising frequency. The peaks and mean values extracted are illustrated in Fig. 2c.

The feature vector obtained using the above procedure is of 36 dimensions (12 normalised frequency values, 12 mean values of the peak frequency and 12 variances). Features 1–12 are the normalised prominent pitch values, features 13–24 are the height of the corresponding peaks and the features 25–36 are the variance of the corresponding peaks. This is illustrated in the Fig. 2c.

### 3.2 Raga Identification Model

ANN consists of number of simple processing elements, called neurons, that are interconnected to each other [16]. A typical multilayer network, consists of a number of layers namely input layer, output layer. One single neuron makes the simple operation of a weighted sum of the incoming signals and a bias term (or threshold), fed through an activation function resulting in the output value of the neuron [16].

In the proposed system, ANN with feed-forward backdrop learning algorithm is used as a classifier (shown in Fig. 3a). It consists of 1 hidden layer. The input layer consists of 36 neurons which is equals to the size of feature vector. The number of neurons in the hidden layer is taken as 40. This number is chosen after experimenting with various number of neurons ranging from 20 to 50. The accuracy of the system is maximum when the number is 40. The number of neurons in the

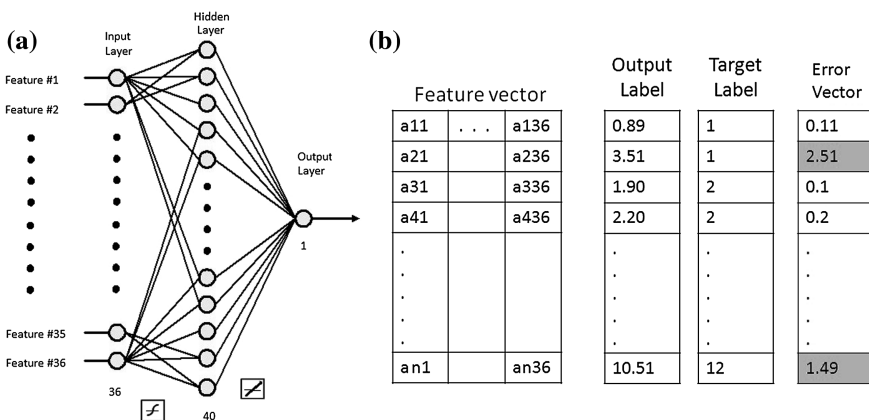


Fig. 3 Raga classification model. a Architecture of ANN. b Error vector calculation and validation

output layer is fixed as 1. Activation function for hidden and output layers is chosen as tan-sigmoid and pure-line respectively. For each *raga* different value is expected as an output. For example, if 2 *ragas* are considered for classification, if input is *raga* 1, then 1 is expected as outcome and if input is *raga* 2, then 2 is expected as output value. ANN is trained accordingly with different labels for different *ragas*. During testing, the *raga* label obtained from the system (output label) is validated using the actual *raga* labels (target labels refer Fig. 3b). The difference between system output and actual output is calculated resulting in the error vector. If the error value is strictly less than 0.5, then the input is considered to be correctly classified. This process is illustrated in Fig. 3b. The gray marked values in error vector indicates mis-classification.

## 4 Result and Analysis

### 4.1 Database

Monophonic and polyphonic clips from 12 *ragas* listed in the Table 1 are considered to form a database. *Ragas* are chosen in such a way that database includes all notes that are present in Carnatic Music. Particularly, *Kalyani* and *Shankarabharanam* both *ragas* are included because *Kalyani* when perceived in a different tonic becomes *Shankarabharanam*. There is high chance of mutual mis-classification between these two *ragas* if the identification is highly tonic dependent. The proposed method is tonic independent and hence the classification of *Kalyani* and *Shankarabharanam* is crucial aspect of this work. In the same way, we also include the *ragas* which share the same set of notes such as *Ananda Bhairavi* and *Reethigowlai* (These two *ragas* share same set of notes), as there is high level of mis-classification between them if set of notes is used for classification. The

**Table 1** Database: *Ragas* used and number of clips

S. No.	<i>Raga</i> name	Number of clips
1	Abhogi	16
2	Harikhambojhi	11
3	Kalyani	13
4	Malayamarutha	10
5	Mayamalavagowla	14
6	Hindolam	17
7	Mohanam	12
8	Hamsadwani	12
9	Nattai	13
10	Shankarabharanam	13
11	Ananda Bhairavi	16
12	Reethigowlai	15



polyphonic music clips consist of vocals accompanied by Tambura or Tabla. The database is made rich in variety by including *arohana-avarohana*, *krutis*, *alapanas* from different *ragas* rendered by instruments and different singers including both male and female. We have also included both *Janaka* and *Janya ragas* in our database. *Janaka ragas (Sampurna ragas)* contain all seven notes whereas *Janya ragas* contain fewer than seven notes.

## 4.2 Performance Evaluation

The experiment is performed using different number of neurons in the hidden layer. Considering 40 number of neurons in the hidden layer yielded the best results. The system is validated using three experiments. Initially the *raga identification* system is developed with 131 music clips from the first 10 ragas of the Table 1, using 70 % of them for training and 30 % for testing. The average *raga* recognition accuracy is found to be 91.5 %. The accuracy of recognition for each *raga* in this experiment is given in Table 2. In the second experiment, the system is tested to verify if the features are able to distinguish between *ragas* that are composed using the same set of notes (*Ananda Bhairavi* and *Reethigowlai*). The system is validated using only these two *ragas*. Using 21 clips for training and 10 clips for testing the accuracy of classification in this experiment is found to be 80 %. This can be justified using the normalisation performed using the pitch value having highest mean and the other parameters such as mean and variances. This either corresponds to tonic or highly repeating note which is called *Jeeva Swara* in Carnatic music. *Jeeva Swara* is different for different *ragas* when they are composed of same notes. If normalisation

**Table 2** Accuracy of the *raga* classification

S.No.	Raga name	Classification accuracy (in percentage)	
		Training and testing with first ten <i>ragas</i>	Training and testing with twelve <i>ragas</i>
1	Abhogi	100.0	100.00
2	Harikhambojhi	100.0	100.00
3	Kalyani	80.00	92.30
4	Malayamarutha	100.00	90.00
5	Mayamalavagowla	80.00	85.71
6	Hindolam	100.00	100.00
7	Mohanam	100.0	91.67
8	Hamsadwani	100.0	91.67
9	Nattai	80.00	84.61
10	Shankarabharanam	80.00	84.61
11	Ananda Bhairavi	NA	80.00
12	Reethigowlai	NA	80.00

factor is tonic, then the means and variances of the peaks differ because the importance given for various notes is different in different *ragas*. Hence classification of *raga* becomes efficient using the proposed features.

In the third experiment, the *raga* recognition system is validated using 5-fold validation method. Complete database consisting of 162 clips from 12 *ragas* is divided into 5 sets, each set consisted of songs from all *ragas*. From these sets, 4 are used for training the network and the remaining one is used for testing. The experiment is repeated for 5 times so that the system is validated against each set. The average result of 5 experiment for each *raga* is listed in Table 2. The average classification accuracy is 89.51 %.

## 5 Conclusion and Future Scope

In this work, a system for direct *raga* classification without use of explicit note sequence is proposed. Set of low-level features obtained from the signal are used as features. A comprehensive data set consisting of variety of music clips is used for validating the proposed method. An effort has been made to come up with a *raga* classification system that is independent of tonic note information. From the study, it is evident that pdf and primitive features give good accuracy even for *ragas* composed of same set of notes. The accuracy of the system clearly demonstrates that the features capture more details about *raga* rather than just the set of notes.

As a future work, the system shall be tested with larger database having hetero-phonous music clips (composed of more than one melody such as vocal accompanied by violin) and also film songs that are composed based on *ragas*. The proposed features shall be tested with other classifiers as well. Further, the *raga* information shall be used for developing systems for transcription of notes, music recommendation based on *raga* information and so on.

## References

1. Bello, J.P., Daudet, L., Abdallah, S., Duxbury, C., Davies, M., Sandler, M.B.: A tutorial on onset detection in music signals. *IEEE Trans. Speech Audio Process.* **13**(5), 1035–1047 (2005)
2. Klapuri, A., Davy, M.: *Signal processing methods for music transcription*. Springer, New York Inc., Secaucus NJ USA (2006)
3. Pandey, G., Mishra, C., Ipe, P.: Tansen: A system for automatic raga identification. In: *Proceedings of Indian International Conference on Artificial Intelligence*, pp. 1350–1363 (2003)
4. Bhattacharjee, A., Srinivasan, N.: Hindustani raga representation and identification: a transition probability based approach. *IJMBC* **2**(1–2), 66–91 (2011)
5. Sinith, M., Rajeev, K.: Hidden markov model based recognition of musical pattern in South Indian classical music. In: *IEEE International Conference on Signal and Image Processing* (2006)

6. Dighe, P., Agrawal, P., Karnick, H., Thota, S., Raj, B.: Scale independent raga identification using chromogram patterns and swara based features. In: IEEE International Conference on Multimedia and Expo Workshops (ICMEW), pp. 1–4 (2013)
7. Koduri, G.K., Serra, J., Serra, X.: Characterization of intonation in carnatic music by parametrizing pitch histograms. In: International Society for Music Information Retrieval, pp. 199–204 (2012)
8. Sridhar, R., Geetha, T.: Swara identification for South Indian classical music. In: 9th International Conference on Information Technology (IEEE) (2006)
9. Shetty, S., Achary, K.K.: Raga mining of indian music by extracting arohana-avarohana pattern. *Int. J. Recent Trends Eng.* **1**(1), 362–366 (2009)
10. Ranjani, H.G., Arthi, S., Sreenivas, T.V.: Carnatic music analysis: Shadja, swara identification and raga verification in alapana using stochastic models. In: IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, pp. 29–32 (2011)
11. Chordia, P.: Automatic rag classification using spectrally derived tone profiles. *Int. Comput. Music Conf.* **2004**, 1–4 (2004)
12. Chordia, P., Rae, A.: Raag recognition using pitch-class and pitch-class dyad distributions. In: Proceedings of International Conference on Music Information Retrieval (2007)
13. G.Koduri, Gulati, S., Rao, P.: A survey of raaga recognition techniques and improvements to the state-of-the-art. *SMC* (2011)
14. Krishnaswamy, A.: On the twelve basic intervals in south Indian classical music. In: 115th AES Convention, New York (2003)
15. Rabiner, L.R.: On the use of autocorrelation analysis for pitch detection. *IEEE Trans. Acoust. Speech Signal Process.* **25**(1), 24–33 (1977)
16. Suykens, J.A.K., Vandewalle, J.P.L., Moor, B.L.R.D.: *Artificial Neural Networks for Modelling and Control of Non-Linear Systems*. Springer, US (1996)