

A Survey of Prospects and Problems in Hindustani Classical Raga Identification Using Machine Learning Techniques

Sreeparna Banerjee

Abstract In this paper we present a survey of current research in Music Information Retrieval in North Indian Classical Music and describe all the characteristics of ragas used for classification. We then describe Bhatkhande's classification scheme and show how it can simplify the classification process of 100 ragas to 10 categories. We also discuss the issues that need to be addressed and the similarities and differences between Hindustani classical music and Western Classical music. Current research efforts on Raga identification are also described.

Keywords Pattern recognition of raga · Music information retrieval · Audio descriptors · Audio signal processing

1 Introduction

Music Information Retrieval (MIR) [1] is a relatively new field in Information Technology that has generated a lot of interest in the last two decades due to its applications in music generation, indexing, processing, recording, reproduction and music oriented services in multimedia. Extensive studies are being conducted in Western classical music [2–5] and notable efforts are also being made in Indian classical music [6–16]. There are two distinct traditions in Indian classical music, namely, Hindustani or North Indian music [6, 7, 9] and Carnatic or South Indian Music [8]. Although they are two distinct genres in Indian music, there are several similarities between them, the most important ones being the concepts of ragas and Talas. The initial work on MIR in Indian classical music has focused on raga identification using soft computing approaches like Support Vector Machine (SVM) [6], swara

Sreeparna Banerjee (✉)

Department of Natural Science and Computer Science and Engineering,
West Bengal University of Technology, Kolkata, India
e-mail: sreeparnab@hotmail.com

histograms [7], data mining algorithms [8] and exponential series analysis [9]. Hindustani Classical Music (HCM) has characteristics described below, some of which resemble Western Classical Music (WCM).

- A raga is essentially a melodic structure in HCM, which conforms to certain rules which dictate its usage of melodic notes (swaras) as well as govern the route of its ascent (arohan) and descent (abarohan) along a musical scale and also possess some preferred patterns or motifs (pakads) during its exposition. These ragas are meant to convey particular emotions. During the presentation of a raga, there are portions of the raga referred to as “gat” that are played using particular rhythms called the Ta ala and in a particular tempo for each of these rhythms.
- Like in Western music, Indian music is also tonal, where one tone, the tonic, provides an anchor around which other tones are organized. However, Western music expresses in tonality through harmony, whereas in Indian music the same is expressed through these musical forms, viz., the ragas. Furthermore, three factors, namely tonal hierarchy, harmonic functions of chords and modulation frequency [2] contribute to Western classical music (WCM). On the other hand, Indian music lacks both modulations and harmonic chord function.
- Tonal hierarchy does exist in Indian music with the melody being organized around the tonic referred to as Sa, which is a solfege of Do, and the fifth note, or sometimes the fourth note being the second most stable note. There is also a predominance of certain notes called the vadi swara followed by the samabadi swara which is generally a perfect fourth or fifth above the vadi swara. There are also anuvadi swaras which exist in the raga as well as vivadi swaras which are foreign to the ragas.
- The ragas must consist of at least five of the twelve notes (swaras) of the musical scale including the tonic and the fourth (Ma corresponding to the solfege Fa in the Western scale) or fifth (Pa corresponding to the solfege So). Indian music uses an additional mechanism called drone [2] which is a continuous sounding of the tonic and usually the fifth scale tone.
- Unlike its Western counterpart, where the musical scores of the composers have to be adhered to, Indian classical music allows ample scope for improvisation. In fact, the beauty and complexity of Indian classical music, be it Hindustani Classical Music (HCM) or Carnatic, is that, shown in Fig. 1, once the rules for raga composition are followed, the performer has complete freedom to compose and render the exposition.

However, this fact has made the job of identifying a raga extremely challenging as no two compositions, even by the same performer, are alike. Since ragas are meant to convey emotions, they have a certain character (devotional, erotic, bold, poignant, etc.), have preferred times for rendition and can also be seasonal.

Fig. 1 Circular block diagram of Thats



2 Characteristics of Hindustani Ragas

2.1 Characteristics

Hindustani or North Indian Classical Ragas can be composed from twelve notes or swards Any Raga must have at least five notes. A Raga with five notes (pentatonic) is called an Orabh raga, with six notes (hexatonic) is called Swarabh raga and all seven notes (septa tonic) is called complete or Sampurna. This characteristic of a Raga is called the jati.

Another important feature of a raga is its anga, which depicts whether the first half of the saptak (or octave) is more important in its rendition, in which case its is Poorvanga, or the second half, in which it is Uttaranga. Poorvanga ragas are generally played or sung during the latter part of the day whereas Uttarang ragas are generally sung in the first half of the day. Nyas swara is a stay note which has a prominent role in the ending of the raga.

Most of the melodies are composed in three octaves, namely, the mandra saptak (lower octave), madhya saptak (middle or most prominent octave) and tar saptak (higher octave). Ragas are built on scales consisting of seven tones. These scales are referred to as thats. The correspondence of the seven notes with the scale notes and their Western solfege syllables of the No major diatonic scale are given in Table 1. Keeping the tonic (Sa) and dominant (Pa) fixed, there other five notes, the remaining five that notes may be altered in one direction only. Re, Ga, Ha and Ni may be lowered by half a step while Ma may be raised by half a step. This gives rise to 32 possible seven note scales [2]. Since these pitch restrictions are absent in Carnatic music there are 72 possible scales known as Melakartha ragas.

The typical rendition of a raga consists of an alap which is free of metrics which invokes the mood of the raga followed by a gat based on a metric pattern. The gat can have different tempos which can conform to different tempos. An additional feature

Table 1 Scale notes/Swaras

Note	Indian note	Ratio to Sa (C)
C (Tonic)	Sa	1
D (Super tonic)	Re (shudh)	9/8
E (Mediant)	Ga (shudh)	5/4
F (Sub dominant)	Ma (shudh)	4/3
G (Dominant)	Pa	3/2
A (Submediant)	Dha (shudh)	5/3
B (Sub tonic/Leading note)	Ni (shudh)	15/8

Table 2 Scale notes/Swaras

That	Note	Vadi	Sambadi
Bilawal	C, D, E, F, G, A, B	A (Ni)	E (Ga)
Khamaj	C, D, E, F, G, A, Bb	E (Ga)	B (Ni)
Kafi	C, D, EbF, G, A, Bb	G (Pa)	C (Sa)
Asavari	CDEb FGAbBb	Ab (Dha)	Eb (Ga)
Bhairavi	CDb Eb F GAb Bb	Ab (Dha)	Db (Re)
Bhairav	CD bEF # GAb B	Ab (Dha)	Db (Re)
Todi	CDb Eb F# GABb	A (Dha)	Eb (Ga)
Marwa	CDb EF # GAbB	Db (Re)	Ab (Dha)
Purvi	CDb EF # GAB	G (Pa)	Db
Kalyan	CDEF # GAB	E (Ga)	B (Ni)

of Hindustani Ragas, is the presence of Shrutis. Subtle nuances of a Raga can be expressed through minor variations (fixed in number for each of the notes) of a note (swara). These variations along with the twelve main notes give rise to 22 shrutis.

2.2 *Bhatkhande's Classification*

The most notable and most popular attempt at classification of ragas has been made by V.N. Bhatkhande. He had made the classification based on the swaras used by the ragas and called them thats. He claimed that the 32 thats mentioned above can be re binned into ten thats which have been named after the popular raga belonging to the that. The ten thats which is described in Fig. 1, their tones and their vadi and samavadi are listed in Table 2. According to Bhatkhande, no two versions of the same note (e.g. B and Bb) can occur during aroha or abaroha. Here b denotes komal (flat) and # denotes tibra (sharp).

3 Problems to Be Addressed in HCM

Some of the concerns that need to be addressed in order to develop suitable raga identification techniques for Music Information Retrieval (MIR) are mentioned in this section.

In addition to the seven notes given in Table 1, Hindustani Classical music ragas use ten micro tones in addition to five semitones corresponding to altered notes of Re, Ga, Ma, Dha and Ni giving rise to a total of 22 Shrutis.

The use of a drone instrument like the tanpura which provides the tonic or absolute frequency along with the fifth (Pa) or fourth (Ma) which provides a continuous pitch reference is an added consideration in HCM.

The presence of melodic motifs called bandish [17] in vocal music and asthayi in instrumental music are provide key phrases used for raga identification. HCM is homo-phonetic in nature. Matching of timbre [18] of the vocal or instrumental performance with the accompanying instruments is an important aspect of music identification.

4 Comparison Between HCM and WCM

In both musical genres, the tonic plays a crucial role [19]. However, WCM is based on harmony, in which a group of notes called chords are played simultaneously [20]. HCM is based on melodic structures called ragas. These ragas can be rendered in several sub-genres like Dhrupad, Khayal, Thumri, etc., The micro tonal scale form based on a natural harmonic series reflects the mood of the raga.

The sub-genres of WCM include symphony orchestra and chamber music. Symphony is a musical composition with several movements, usually four, with each movement having a different tempo. Symphonies are usually orchestras with multiple instruments and sometimes, voices are included. Chamber music is composed of a small group of musical instruments.

HCM belongs to the chamber music category with one or two performers along with an accompanying background drone from the tanpura to provide fundamental frequency (pitch) reference and also an instrument to provide the tala or rhythm, which can be played with different layas (or tempo).

5 Music Information Retrieval(MIR)

As ragas form the backbone of HCM, MIR consists essentially of their classification, performed by first extracting low level features or audio descriptors [21] and subsequent identification using high level features [22], a process called audio data mining and use of metadata [23]. A typical low level search process could use LPI combined with a high level process like LSI [24].

After signal acquisition and signal thumb nailing to extract the most repetitive portion of the musical piece, which could provide a representative summary of the raga, feature extraction [25] is carried out. This feature extraction consists of low level features and high level features referred to as music content and music context, respectively [26].

The low level features [25] or music content are features that can be extracted directly from the audio signal can be classified as follows:

- 1 Time domain features like mean energy, zero crossings, number of silent frames.
- 2 Frequency domain based features like spectral centroid, roll-off, spectral flux, etc.,
- 3 Coefficient based features like MFCC (Mel frequency cepstral coefficients) LPC (Linear Predictive Coding), etc.,
- 4 Time-frequency domain features like pitch. This is the most widely used feature.

Loudness and timbre are low level features. Also sometimes referred to as the color or quality of sound. Timbre is related to three main properties of music signal. These are spectral envelop of shape and time variation of spectrum and time evolution of energy [27]. Audio descriptors are also a special set of low level features. High level features or musical context cannot be obtained directly from the audio signal but can be derived from it. These features include genre, melody, mood, rhythm and timbre in context of the melody and artist identification.

Some of the subfields of MIR [26] are feature extraction, similarity search, classification and applications like audio fingerprinting, which is a compact signature for content based audio retrieval. Feature extraction involves processes like timbre description, music transcription and melody extraction. Query based approaches fall under the similarity search category. Classification embraces emotion and mood recognition, genre classification, instrument classification, composer and performer identification among other techniques. Content based query and retrieval is another example of applications.

6 Related Work

The related work can be classified as work related to raga identification and works addressing concerns of Sect. 3.

6.1 Raga Identification

Since MIR in Indian Music is a relatively new field there is a small amount of literature available. One of the initial attempts was by Sahasrabuddhe and Upadhye [10] who used finite state automata to model a raga based on its swara constituent pattern. Pandey et al. [11] used the Hidden Markov Model (HMM) on swara sequences,

treating them as word sequences based on swara alphabets in their Tansen raga recognition system. A Gaussian Mixture Model based on HMM using three features, namely, Chroma, MFCC (Mel Frequency Cepstral Coefficients), was used by Dighe et al. [12]. By combining these three features, they obtained a 62 dimensional feature vector. Pitch-class profile distribution with K-NN classifier and K-L divergence was used in [13]. Some of the research focused on a particular characteristic of a raga like [14] who compared the arohan and avarohan, [15] who studied vadi and samavadi and [16] who studied pakar.

6.2 Works Addressing Issues of Section 3

Attempts made to address the problems discussed in Sect. 3 are described below.

The harmonium is a keyboard system like the piano, but, unlike the piano where there are pauses between musical notes, the notes from the harmonium linger, thus producing continuous swaras. In [28], a novel method for raga recognition by utilizing the continuous sound producing property of the harmonium has been presented. At first, audio signals from the harmonium were segmented into separate frames for note detection purpose. Because the swaras are played continuously, proper onset and end-point detection is very critical for identification and this was done using two approaches, namely spectral flux determination and fundamental frequency estimation in the time-frequency domain analysis of each frame. Raga identification was treated as a template matching problem using a database of ragas to match the query audio signals with the prototypes in the database.

The homo-phonic aspect of HCM and timbre, whereas single melody line is accompanied by instruments like tanpura or harmonium was explored in [29]. The authors separated the singer from the accompanying instrument using a hybrid selection algorithm applied to six audio descriptors, namely, attack time, attack slope, zero crossing rate, roll-off, brightness, roughness and MFCC with 20 coefficients from the MIR toolbox of MATLAB. The k-means and a statistical classifier k nearest neighbors are used to identify and differentiate between human voice and instrument, although they have similar timbres. Since tonic is important in HCM, its identification should be the first step in automated raga identification, and this has been attempted by [29]. They used a four step process for music extraction. In the first step spectral peaks were isolated from the audio signal. Then the salience function determination was performed using the sum of weighted energies extracted at integer multiples (harmonics) of that frequency, followed by tonic candidate generation obtained from the pitch histogram of the entire audio signal and finally selecting the tonic as the highest peak. Classification was performed using the C4.5 decision tree of the Weka software.

Tuning of HCM was studied by [30] Stable fundamental frequency collection was done followed by construction of interval histogram. Their objective was to detect the twelve notes of the octave as well as further subdivisions of the octave to look for the presence of micro tones. Reference [24] extracts formants from audio files using

LPC for detection and then term document matrix which is decomposed using Single Vector Decomposition for Latent Semantic Indexing (LSI). The detected formant (query) is matched with the document.

Vocal expressions like glides (meend) and vibrato (andolan) can play an important role in raga identification. An example is the distinction between ragas Bhupali and Shudh Kalyan. An algorithm has been described by [27] to perform this operation. Firstly, the pitch curve was estimated and singing voice frames were identified. A pitch envelope was used to develop a canonical representation and, finally, templates were used for identification of expressions for creating a transcription of an audio signal.

7 Prospects

In summary it might be remarked that, although MIR in Hindustani Classical Music is in a nascent state, specially as compared with its Western counterpart, there has been a lot of progress in the past decade and, so, it is hoped that the future will witness remarkable progress in Raga identification, which is the mainstay of MIR in HCM.

References

1. Markus Schedule, Emilia Gmez, Julin Urbano, Music Information Retrieval: Recent Developments and Applications, Foundations and Trends R in Information Retrieval, Vol. 8, No. 2–3, pp. 127261, (2014).
2. Mary A. Castellano, J. J. Bharucha, Carol L. Krumhansl, Tonal Hierarchies in the Music of North India, *Journal of Experimental Psychology: General*, 113 (3), pp. 394–412 (1984).
3. G. Tzanetakis, G. Essl, P. Cook, Automatic Music Genre Classification of Audio Signals, Proceed. 2nd International Symposium of Music Information Retrieval, Bloomington, Indiana, USA, pp. 205–210, (2001).
4. Felix Weninger, Noam Amir, Ofer Amir, Irit Ronen, Florian Eyben, Bjorn Schuller, Robust feature extraction for automatic recognition of vibrato singing in recorded polyphonic music, Proceed. 37th International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, pp. 85–88, (2012).
5. A. Ghias, J. Logan, D. Chamberlin and B.C. Smith, “Query by Humming- Musical Information Retrieval in an audio database” Proc. ACM Multimedia, pp 231–236 (1995).
6. Vijay Kumar, Harit Pandya, C.V. Jawahar, Identifying Ragas in Indian Music, Proceed. 22nd International Conference on Pattern Recognition (ICPR), pp. 767–772 (2014)
7. Pranay Dighe, Harish Karnick, Bhiksha Raj, Swara histogram based structural analysis and identification of Indian classical ragas, 14th. Conference of the International Society for Music Information Retrieval, Curitiba, Brazil, pp. 35–40, (2013).
8. K Oriya, R. Geetha Ramani, Shomona Gracia Jacob, Data Mining Techniques for Automatic recognition of Carnatic Raga Swaram notes, *International Journal of Computer Applications*, 52(10) pp. 4–10, (2012).

9. Soubhik Chakraborty, Saurabh Sarkar, Swarima Tewari and Mita Pal, An interesting application of simple exponential smoothing in music analysis, *International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI)*, 2(4), pp. 37–44, (2013).
10. H. Sahasrabudde and R Upadhye, “On the computational model of raag music of India”, *Workshop on AI and Music, European Conference on AI (1992)*
11. G. Pandey, C. Mishra and P. Ipe, “Tansen: A system for automatic raga identification”, *Proc. Indian International Conference on Artificial Intelligence*, pp. 1350–1363 (2003).
12. Pranay Dighe, Parul Agrawal, Harish Karnick, Siddhartha Thota and Bhiksha Raj, Scale independent raga identification using chroma gram patterns and swara based features, *IEEE International Conference on Multimedia and Expo Workshops*, 2013.
13. Koduri Gopala-Krishna, Sankalp Gulati and Preeti Rao, A Survey of Raaga Recognition Techniques and Improvements to the State-of-the-Art, *Sound and Music Computing*, 2011.; OP. Chordia and A. Rae, Raag recognition using pitch-class and pitch-class dyad distributions, *International Society for Music Information Retrieval*, 2007.
14. S. Shetty and K. Achary, Raaga Mining of Indian Music by Extracting Arohana-Avorahan patterns, *International Journal of Recent trends in Engineering*, 2009.
15. S. Shetty and K. Achary, Raga Identification of Carnatic Music for Music Information Retrieval, *International Journal of Recent trends in Engineering*, 2009.
16. Vignesh Iswar, Shrey Dutta, Ashwin Bellur, Hema A. Murthy, Motif Spotting in an Alpana in Carnatic Music, *International Society for Music Information Retrieval*, 2013.
17. Joe Cheri Ross*, Vinutha T. P. and Preeti Rao, Detecting melodic motifs from audio for Hindustani classical music, *13th International Society for Music Information retrieval conference (ISMIR 2012)* 2012.
18. Saurabh H. Deshmukh and S. G. Bhirud, North Indian Classical Musics Singer Identification by Timbre Recognition using MIR Toolbox, *Intl. J. Computer Applications*, 91(4) 1–4 (2014).
19. Justin Salamon, Sankalp Gulati and Xavier Serra, A multipitch approach to tonic identification in indian classical music, *Proceed. International Society for Music Information Retrieval Conference 2012*.
20. Parul Agarwal, Harish Karnick and Bhiksha Raj, A comparative study of indian and western music forms, *Proceed. ISMIR 2013*.
21. Saurabh H. Deshmukh, Divya P. Bajaj, S.G. Bhirud, Audio descriptive analysis of singer and musical instrument identification in north indian classical music, *International Journal of Research in Engineering and Technology*, 4(6), 505–508, June 2015.
22. Surendra Shetty, K K Achary and Sarika Hegde, Feature Extraction Computation and Automatic Raga Identification for Carnatic Ragas, *International Journal of Information Processing*, 7(2), 41–51, 2013
23. Preeti Rao, Audio metadata extraction: the case for Hindustani classical music, *Proceed. of SPCOM 2012, IISc Bangalore* (2012)
24. P Kirthika, Rajan Chattamvelli, Frequency based audio feature extraction for raga based musical information retrieval Using LPC And LSI, *Journal of Theoretical and Applied Information Technology*, 69(3) 2014.
25. Balaji Thoshkahna, Algorithms for Music Information Retrieval, *Master of Science Thesis, IISc Bangalore*, April 2006.
26. Markus Schedl, Emilia Gomez and Julian Urbano, Foundations and trends in Information Retrieval, 8(2–3), 127–161 (2014)
27. Sai Sumanth, Miriyala Kalika, Bali Ranjita, Bhagwan Monojit Choudhury, Automatically identifying vocal expressions for music transcription, *Proceed. ISMIR 2013*.
28. Rajshri Pendekar, S. P. Mahajan, Rasika Mujumdar, Pranjali Ganoo, Harmonium Raga Recognition, *International Journal of Machine Learning and Computing*, Vol. 3, No. 4, August 2013.
29. Justin Salamon, Sankalp Gulati and Xavier Serra, A multipitch approach to tonic identification in indian classical music, *Proceed. ISMIR Conference 2012*.
30. Joan Serra, Gopala K. Koduri, Marius Miron and Xavier Serra, Assessing the tuning of sung indian classical music, *Proceed. ISMIR 2011*