International Journal of Recent Trends in Engineering Vol. 1, No. 1, May 2009,

Raga Mining of Indian Music by Extracting Arohana-Avarohana Pattern

Surendra Shetty[1] and K.K. Achary [2]
[1] Department of Computer Applications, NMAMIT, Nitte, Karnataka, India
Email: hsshetty4u@yahoomail.com
[2] Department of Statistics, Mangalore University, Mangalagangothri, Mangalore, Karnataka, India
Email: kka@mangaloreuniversity.ac.in

Abstract—**Data mining is cataloging through data to identify patterns and establish relationships. The relationship among patterns provides information. It can be converted into knowledge about historical patterns and future trends. Raga mining is a part of data mining where audio data is considered in specific. Raga is an attractive combination of notes, engaging tonal context. What distinguish each raga are the characteristic phrases (of notes), sequence of notes, and the treatment given to each note in terms of its timing, rendition, prominence or ornamentation. Generating script for a musical composition and identifying raga in Indian Classical music is a tedious and time consuming task. Here, we depict a system which takes an audio file as an input and converts it into sequence of notes, identifies the raga by extracting its Arohana-avarohana pattern. The system is analyzed for 50 different rags and tested for 20 different ragas with 3-5 songs of each raga totally 90 songs. The system performs with a result of 95%.**

Index Terms—**ANN, Arohana, Avarohana, Indian Classical Music, Pitch, Swara, Raga**

I. Introduction

Ragas are sometimes defined as melody types. The raga system is a method of organizing tunes based on certain natural principles. Tunes in the same raga use the same (nominal) swaras in various combinations and with practice, the listener can pick up the resemblance. Indian classical music is defined by two basic elements - it must follow a Raga (classical mode), and a specific rhythm, the Taal. In any Indian classical composition, the music is based on a drone, i.e. a continual pitch that sounds throughout the concert, which is a tonic [5]. Each raga has a swaroopam (a musical form or image) that is defined by the swaras used, the gamakas given to these swaras, the progression in which the swaras occur etc. This classification is termed as the raga lakshanam. Raga lakshanam usually contains the arohanam, avarohanam, details of raga chaya swaras (the swaras which are chiefly responsible for the characteristic melody of the raga), gamakas, characteristic swara phrases and general usage notes. It is intended more for the performer than for the listener. Arohanam is the sequence of swaras used in a raga in the ascending passages i.e. as the pitch goes up. Avarohana is the sequence of swaras to be used in descent. The arohana and avarohana (or the scale) of a raga provide only emaciated outline upon which the rest of the raga is formed.

Raga Identification is a process of listening to a piece of music, synthesizing it into sequence of notes and analyzing the sequence of notes for identifying the raga it follows. Note Transcription is the first step in raga Identification.

II. Related works

A. Note Transcription

Note Transcription of music is defined to be the act of listening to a piece of music and of writing down the musical notation for the sounds that constitute the piece. In other words, this means converting an acoustic signal into a symbolic representation, which comprises musical events and their parameters [7]. Many attempts are made towards the Note Transcription of Music [7],[12],[2]. Kalpuri proposed a system for the automatic transcription of Western music [2]. Here, Signal processing techniques are introduced that solve different facets of the overall predicament. Main emphasis is laid on finding the multiple pitches of concurrent musical sounds. A.Krishnaswamy [7], has described a method on how Pitch Tracking is useful for Note Transcription of South Indian Classical Music. He presented the results of applying pitch trackers to samples of South Indian classical (Carnatic) music.

He investigated the various musical notes used and their intonation and tried different pitch tracking methods and observed their performance in Carnatic music analysis. Rajeswari Sridhar and T.V. Geetha, have proposed a system for Swara Identification for South Indian Classical Music [12]. It deals with the identification of the swaras in a given Carnatic song. The frequencies associated with the segments are identified and the exact tagging of swara is performed to find the 7-swara combinations in the given music signal.

B. Raga Identification

Pandey [5] developed a system to automatically identify ragas Yaman and Bhupali using a Markov model. A success rate of 77% was reported on thirty-one samples in a two-target test, although the methodology was not well documented. An additional stage that searched for specific pitch sequences improved performance to 87%. In an exploratory step, Chordia [11] classified one hundred thirty segments of sixty seconds each, from thirteen raags. The feature vector was the Harmonic pitch class profile (HPCP) for each segment. Perfect results were obtained using a K-NN classifier with 60/40% train/test split. This was further developed in [8] where PCDs and PCDDs were used as features with more sophisticated learning algorithms. In a 17 target experiment with 142 segments, classification accuracy of 94% was attained using 10-fold cross-validation. However, the significance of the results in both cases was limited by the size of the database. A system is developed to identify raags based on pitch-class distributions (PCDs) and pitch-class dyad distributions (PCDDs) calculated directly from the audio signal [9]. Here, a large, assorted database consisting of 20 hours of recorded performances in 31 different raags by 19 different performers was assembled to train and test the system. Classification was performed using support vector machines, maximum a posteriori (MAP) rule using a multivariate likelihood model (MVN), and Random Forests. When classification was done on 60s segments, a maximum classification accuracy of 99.0% was attained in a cross-validation experiment. In a more difficult unseen generalization experiment, accuracy was 75%.

III. TECHNICAL WORK PREPARATION

A. Problem Statement

The objective of the paper is to develop a system which automatically mines the raga of an Indian Classical Music. The fig 3.1 shows the configuration of the proposed system. As a first step Note transcription is applied on a given audio file to generate the sequence of notes used to play the song. In the next step, the features related to Arohana – Avarohana are extracted. These features are given to ANN for training and testing the system.

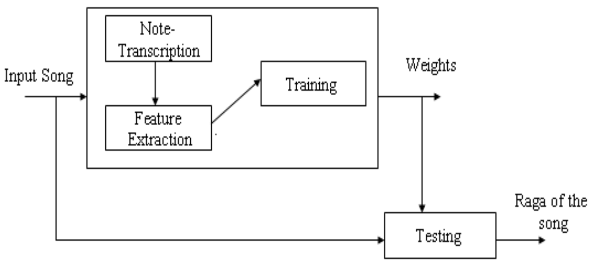

Figure 3.1 System Structure

B. Note Transcription

In order to identify the raga of the songs, the sequence of notes used to play the song must be known. The process of identifying the sequence of notes i.e the swara-script is called as Note Transcription. For Indian classical music, Note Transcription process itself is a very challenging task. Gaurav Pandey used two heuristics, based on the pitch of the audio sample for Note transcription, The Hill Peak Heuristic and Note Duration Heuristic, [5]. Sound onset detection and musical meter estimation was proposed by Kalpuri [2] for Note transcription in Indian Classical Music. Here we illustrate a method for this process and discuss the incurred problems.

Music is represented as signal, so Note Transcription involves signal Processing techniques. Music can also be defined as a melody composed by combination of swaras, where swara is built of set of frequencies – Fundamental frequency and harmonic frequencies. The human perception of these set of frequency is called as Pitch.

The first step in Note Transcription is identifying all the different frequency segments in the song called as 'Frequency Extraction'. Fundamental frequency of each segment is calculated using the Autocorrelation method [12] with a frame-size of 50ms. For simplicity, the type of the audio file taken as input is monophonic song with a single sound-source. System can be enhanced further for recorded Polyphonic music, by applying multi-pitch detection techniques described in [5]. After listing all the frequency segments in the audio file, each such segment has to be converted into its corresponding swara. Fig 3.2 shows a graph where list of fundamental frequencies of a song are plotted. The duration of the song is 34sec.

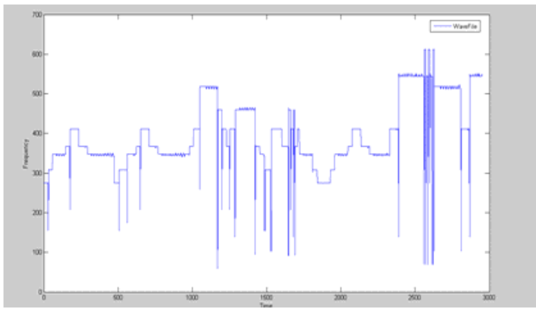

Fig 3.2 Fundamental Frequency List graph for a Monophonic song of 34sec

In Indian Classical Music, the relationship between swara and the frequency used to play it is not fixed. It depends on the fundamental frequency of the note called as 'Sa', which is called as 'Shruti' or 'Scale' of the song. Fundamental frequencies of all swaras are related to Shruti with a defined ratio as shown below.

The Ratios are as below

Sa---1	Ma2---729/512
Re1---256/243	pa ---3/2
Re2--- 9/8	dha1---128/81
Ga1--- 32/27	Dha2---27/16
Ga2--- 81/64	ni1----16/9
Ma1---- 4/3	Ni2----243/128

Shruti is highly variable in Indian Classical Music (It can be different for different singer). So, in this process one of the important tasks is to find out which fundamental frequency corresponds to the frequency of swara 'Sa'. In our system the scale of the song is fixed to 240Hz.

Once the Scale is known, other frequency segments are mapped according to the defined ratio into the 36 swaras of Mandra, Madhya and Taara saptaka.

C. Feature Extraction

A raga is defined by: The choice of notes (from the 12 notes), Ascending and descending sequences *(arohana &*

avarohana), The nature of inflexion on different notes *(gamaka/meend),* Characteristic phrases (groups of notes) *(Swara sanchara/chalan/pakad).* The type of the features extracted here are, the choice of notes and the Arohana-Avarohana pattern. These features help to identify the raga. A *raga* is constructed of 5 to 7 consistent *svaras* (melodic steps). Each *raga* has an ascending (*arohana*) and descending (*avarohana*) form; these forms may be different i.e. the *arohana* and *avarohana* may contain different svaras; *Arohana* is the ascending sequence of notes which the raga follows. Any ascending sequence in improvised portions of the raga follows the pattern defined in *Arohana* strictly. Similarly, *Avarohana* is the corresponding descending sequence. Table 1 shows the Arohana/Avarohana and Swara combination of some of the ragas

In our system the features which are extracted from the input files are, Swara Combination, Numbers of swaras used in the raga, Vakra pairs in Arohana & in Avarohana.

TABLE 1
AROHANA-AVAROAHANA PATTERN IN RAGAS

RAGA	SWARA combination	AROHANA/ AVAROHANA
1 Todi	S r1 g1 m1 p d1 n1	S r1 g1 m1 p d1 n1 s'/ s' n1 d1 p m1 g1 r1 s
2 Dhanyasi	s r1 g1 m1 p d1 n1	S g1 m1 p n1 s' /s' n1 d1 p m1 g1 r1 s
3 Varali	s r1 g1 m2 p d1 n2	s r1 g1 m2 p d1 n2 s'/s' n2 d1 p m2 g1 r1 s
4 Mayamalavagaula	s r1 g2 m1 p d1 n2	s r1 g2 m1 p d1 n2 s'/s' n2 d1 p m1 g2 r1 s
5 Saveri	s r1 g2 m1 p d1 n2	s r1 m1 p d1 s'/s' n2 d1 p m1 g2 r1 s
6 chakravak	s r1 g2 m1 p d2 n1	s r1 g2 m1 p d2 n1 s'/s' n1 d2 p m1 g2 r1 s
7 Gaula	s r1 g2 m1 p n2	S r1 m1 p n2 s'/s' n2 p m1 g2 r1 s
8 Kamavardhini	s r1 g2 m2 p d1 n2	s r1 g2 m2 p d1 n2 s'/s' n2 d1 p m2 g2 r1 s
9. Saurashtra	s r1 g2 m1 p d2 n1 n2	s r1 g2 m1 p d2 n2 s'/s' n1 d2 n1 d2 p m1 g2 r1 s
10. Abheri	s r2 g1 m1 p d2 n1	s g1 m1 p n1 s'/s' n2 d2 p m1 g1 r2 s
11. Anandabhairavi	s r2 g1 m1 p d2 n1	S g1 r2 g1 m1 p d2 p s'/s' n1 d2 p m1 g1 r2 s
12. Bhairavi	s r2 g1 m1 p d1 d2 n1	s r2 g1 m1 p d2 n1/s' n1 d1 p m1 g1 r2 s

Swara combination is nothing but basic set of notes(swara) which is used to compose a musical script. Swara combinations are represented in bits. The binary sequence is converted into decimal value

For Example, Mohana Raga ,The swara combination is : s r2 g2 p d2. So,

S	R1	R2	G1	G2	M1	M2	P	D1	D2	N1	N2
1	0	1	0	1	0	0	1	0	1	0	0

But, as 's' is present in all the ragas we can ignore the value and the binary sequence will be, 01010010100 – 660. Another feature is also one of the characteristics of raga. It tells about the number of distinguished swaras used in the raga.

Example 1: Mohana Raga, The swara combination is: s r2 g2 p d2. Numbers of swaras are: 05

Example 2: Todi, The swara combination is, S r1 g1 m1 p d1 n1, Number of swaras are:07

The Arohana and Avarohana of the raga can be either linear or non-linear. In linear, the Arohana and Avarohana pattern is same as the swara combination in ascending and descending sequence respectively. In case of non-linear the Arohana-Avarohana pattern may have some of the pairs which are not present in ascending/descending sequence of swara-combination. Such pairs in non-linear arohana-avarohana are called as vakra. For example,

1. Chakravak raga (linear)
   Swara combination sequence , s r1 g2 m1 p d2 n1
   Linear arohana: s r1 g2 m1 p d2 n1 s'
   Linear Avarohana: s' n1 d2 p m1 g2 r1 s
2. Dhanyasi raga (Non-linear)
   Swara Combination sequence, s r1 g1 m1 p d1 n1
   Aroahana: S g1 m1 p n1 s'
   Avaraohana : s' n1 d1 p m1 g1 r1 s
   Vakra pairs are, (s g1) and (p n1)

Here we find out the set of 2 vakra pairs in arohana and 2 vakra pairs in avarohana. Each possible pair is assigned a value according to which the features are generated. The feature value will be zero if raga has a linear arohana/avarohana. Totally 6 features are extracted from the generated sequence of notes.

D. Training and Testing

A neural network is constructed by highly interconnected processing units (nodes or neurons) which perform simple mathematical operations. Neural networks are characterized by their topologies, weight vectors and activation function which are used in the hidden layers and output layer. The topology refers to the number of hidden layers and connection between nodes in the hidden layers. The activation functions that can be used are sigmoid, hyperbolic tangent and sine. A very good account of neural networks can be found in [13]. The network models can be static or dynamic. Static networks include single layer perceptrons and multilayer perceptrons. A perceptron or adaptive linear element (ADALINE) [3] refers to a computing unit. This forms the basic building block for neural networks.

The input to a perceptron is the summation of input pattern vectors by weight vectors. In Figure 3.3, the basic function of a single layer perceptron is shown.

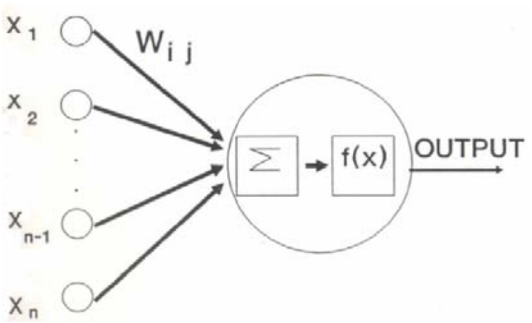

Fig 3.1 Operation of a Neuron

In Figure 3.4, a multilayer perceptron is shown schematically. Information flows in a feed-forward manner from input layer to the output layer through hidden layers. The number of nodes in the input layer and output layer is fixed. It depends upon the number of input variables and the number of output variables in a pattern. In this work, there are six input variables and one output variable. The number of nodes in a hidden layer and the number of hidden layers are variable. Depending upon the type of application, the network parameters such as the number of nodes in the hidden layers and the number of hidden layers are found by trial and error method [6].

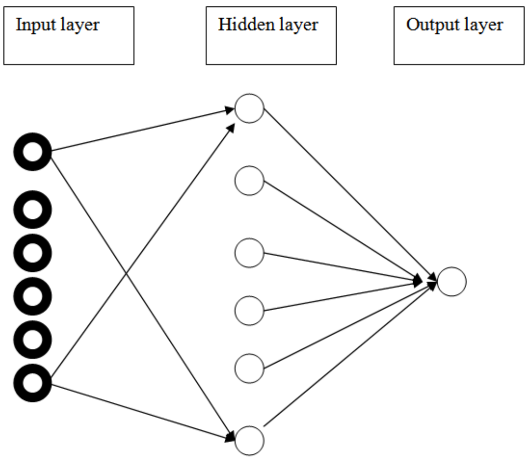

Fig 3.4 Multilayer Perceptron

In most of the applications one hidden layer is sufficient. The activation function which is used to train the ANN is the sigmoid function and it is given by:

$$f(x) = \frac{1}{1+\exp(-x)} \qquad (1)$$

where, f (x) is a non - linear differentiable function,

$$x = \sum_{i=1}^{N_n} W_{ij}(p)\, x_i^n(p) + \theta(p),$$

If testing is done for other ragas which are not trained, it gives random output.

where,

N_n is the total number of nodes in the n^{th} layer, $W_{ij,}$ is the weight vector connecting i^{th} neuron of a layer with the j^{th} neuron in the next layer, θ is the threshold applied to the nodes in the hidden layers and output layer and p is the pattern number. In the first hidden layer, x_i is treated as an input pattern vector and for the successive layers; x_i is the output of the i^{th} neuron of the proceeding layer. The output x_i of a neuron in the hidden layers and in the output layer is calculated by:

$$X_i^{n+1}(p) = \frac{1}{1+\exp(-x+\theta(p))} \qquad (2)$$

For each pattern, error E (p) in the output layers is calculated by:

$$E(p) = \frac{1}{2}\sum_{i=1}^{N_M}(d_i(p) - x_i^M(p))^2 \qquad (3)$$

where,

M is the total number of layer which includes the input layer and the output layer, N_M is the number of nodes in the output layer. $d_i(p)$ is the desired output of a pattern and $X_i^M(p)$ is the calculated output of the network for the same pattern at the output layer.

The total error E for all patterns is calculated by:

$$E = \sum_{p=1}^{L} E(p) \qquad (4)$$

where, L is the total number of patterns.

IV. Experimental studies

In our system, we extract 6 features as described above and these features are associated with variables x1, x2, x3, x4, x5 and x6. Here x1 indicates the number of distinguished swaras, x2 indicates the swara combination. Each possible pair in the swara script is assigned a value for unique identification of the pair. The arohana vakra pairs values are stored in x3, x4 and the avarohana vakra pairs values are stored x5 and x6. We give these values x1-x6 as input to the training system. We took 4 songs of each raga and the features were extracted. Then the features of two/three songs are selected randomly and given as input to the training system. Totally songs of 20 ragas are considered. The system generates the weights according to the inputs given. Then it is tested by extracting features of one/two songs of each raga, which is given as input along with weights generated in the training part. The system gives the output indicating the identification of each raga. Unique labeling is done for each raga. The system identifies the set of rags for which it is trained.

V. Results and discussions

Note Transcription is done for monophonic song (audio signal), as the frequency can be easily determined. Note transcription process generates the sequence of swara used to play song. It gives perfect result with a known Scale of the song. The Feature Extraction is done for the songs belonging to 50 different ragas. The Table 2 shows the results of Feature extraction for the songs belonging to the 20 different ragas. The table shows the output for one song belonging to a raga. Any other song will give the same results as every song follows the Arohana/Avarohana pattern of the raga. If the Arohana/Avarohana is linear then the output will have x3, x4, x5, and x6 as zero, which means there is no vakra-pair. If arohana and avarohana is non-linear, then the output will have values for x3-x6. For most of the raga, we get this value zero, as it will be linear. For the set of ragas which has same swara combinations the vakra-pair features will differ. The feature extraction is done for 90 songs of 50 ragas. Out of which, the features of 60 songs are used to train the network. The features of remaining 30 songs are used for testing purpose. It performs with a result of 95%. Matlab tool is used for the implementation.

TABLE 2
FEATURE EXTRACTION RESULTS

No	Raga	Features					
		x1	x2	x3	x4	x5	x6
1	abhogi	5	836	75	130	124	166
2	Anandabhairavi	7	854	4	0	0	0
3	Arabhi	7	725	32	130	0	0
4	Bilahari.	7	725	60	130	0	0
5	Chakravak.	7	1238	0	0	0	0
6	Hansadhwani	5	657	60	103	96	151
7	Kalyani	7	693	0	0	0	0
8	Kamavardhini	7	1209	0	0	0	0
9	Kambhoji	7	726	130	0	0	0
10	Kanada	7	854	0	0	0	0
11	Kannada	7	725	44	75	0	0
12	Kedar	6	721	103	0	96	0
13	Khamaj	7	726	0	0	0	0
14	Madhyamavathi	5	594	32	102	68	138
15	Mayamalava	7	1241	0	0	0	0
16	Mohana	5	660	60	130	96	166
17	Saranga	7	693	81	0	0	0
18	Shankara	7	725	0	0	0	0
19	Todi	7	1370	0	0	0	0
20	Dhanyasi	7	1370	117	103	0	0

VI. CONCLUSIONS

Raga in an Indian Music is a very complex structure. The sequence of notes used to play the songs is based on the raga. Here we analyze the sequence of notes for Raga identification. The arohana – avarohana pattern is well defined for each raga so it is very useful feature in identification of the raga. The system works successfully for monophonic song. . This method works well for even polyphonic songs, however with some qualms in frequency extraction. This is due to the fact that rhythmic instruments do not contribute to the rag, but only the beat, and yet superimpose frequency components in the spectrum. The system can be enhanced further by generating a complete swara-script containing notes and rhythm information which will be very useful for musicians. The method of raga identification can be improved by using rhythmic information.

REFERENCES

[1] Alai de Cheveigne, “YIN, a Fundamental frequency estimator for speech and music”, Journal of *Acoustical Society of America*., Vol. 111, No. 4, April 2002

[2] Anssi Klapuri, “Signal Processing Methods for the Automatic Transcription of Music”, *Thesis for the degree of Doctor of Technology, USA*, 2004

[3] Bernard Widrow (1990), 30 Years of adaptive neural networks: Perceptron, madaline and back-propagation, Proc. of the IEEE, 18(9), pp 1415 - 1442.

[4] Fortuna L, Graziani S, LoPresti M and Muscato G (1992), “Improving back-propagation learning using auxiliary neural networks,” Int. J of Control , 55(4), pp 793-807.

[5] Gaurav Pandey, Chaitanya Mishra, and Paul Ipe, “Tansen: A system for automatic raga Identification”, *In Proceedings of the 1st Indian International Conference on Artificial Intelligence, pp 1350–1363, 2003.*

[6] Hirose Y, Yamashita K Y and Hijiya S (1991), Back-propagation algorithm which varies the number of hidden units, *Neural Networks*, 4, pp 61-66.

[7] Krishnaswamy, A, “Application of pitch tracking to South Indian classical music”, *Applications of Signal Processing to Audio and Acoustics, 2003 IEEE Workshop* on 19-22 Oct. 2003.

[8] Parag Chordia, “Automatic raag classification of pitch-tracked performances using pitch-class and pitch-class dyad distributions”, *In Proceedings of International Computer Music Conference, 2006*

[9] Parag Chordia, Rae, A. “Raag recognition using pitch-class and pitch-class dyad distributions”, *In Proceedings of International Conference on Music Information Retrieval, (2007)*

[10] Parag Chordia, "Understanding Emotion in Raag: An Empirical Survey of Listener Responses." In *Proc. of the 2007 International Computer Music Conference* (ICMC)

[11] Parag Chordia, “Automatic rag classification using spectrally derived tone profiles”, *In Proceedings of the International Computer Music Conference, 2004,* volume 129, pp 83-87

[12] Rajeswari Sridhar, and T.V. Geetha, “Swara Indentification for South Indian Classical Music“, *9th International Conference on Information Technology (ICIT'06)*, (2006), pp 143-144

[13] Simon Haykin, *Neural Networks-A Comprehensive foundation,* Prentice Hall; 2 edition

[14] S. K. Padmanabha Acharya, “PanchamaVeda”, Published by S.K. Padmanabha Acharya, Kasargod, India, 2001.