Notice of Violation of IEEE Publication Principles

"A Tool for Identifying Ragas Using MIDI (Musical Instrument Devices) for CMIR (Classical Music Information Retrieval)"
by R. Sudha,
in the Proceedings of the 8th IEEE International Conference on Dependable, Autonomic and Secure Computing, December 2009, pp. 837-842

After careful and considered review of the content and authorship of this paper by a duly constituted expert committee, this paper has been found to be in violation of IEEE's Publication Principles.

This paper contains portions of text from the paper(s) cited below. A credit notice is used, but due to the absence of quotation marks or offset text, copied material is not clearly referenced or specifically identified.

"Raga Mining of Indian Music by Extracting Arohana-Avarohana Pattern"
by Surendra Shetty and K.K. Achary
International Journal of Recent Trends in Engineering, Vol 1, No 1, May 2009, pp. 362-366

A Tool for Identifying Ragas using MIDI (Musical Instrument Devices) for CMIR (classical Music Information Retrieval)

R. Sudha
Assistant Professor, Department of Computer Science and Engineering
B.S.Abdur Rahman University, Tamil Nadu, Chennai-48

ABSTRACT

The main motive behind raga identification is that it can be used as good basis for music information retrieval of carnatic music songs or film songs based on carnatic music. The input polyphonic music signal is analyzed and made to pass through a signal separation algorithm to separate the instrument and the vocal signal. The frequency components of the signal are then determined and we map these frequency in to swara sequence ant thereby determine the raga of the particular song.30-40 a sample of Base Ragas from 72 is being identified which will help more for music learners and musicians. Still now only vocal type of identification is being carried out and we provide the identification with string type Instruments of 1 or 2 Instruments mixture (violin/sitar/both).

Keywords: Polyphonic, Arohana, Avarohana

1. INTRODUCTION

Ragas are the central structure of Indian Classical Music, each consisting of a unique set of complex melodic gestures. The gestures are sequence of notes that are often inflected with various micro-pitch alterations and articulated with an expressive sense of timing. Music Information retrieval is a part of Multimedia Information Retrieval. A lot of work has been done in the other components of multimedia like text, video and the one that is yet to be fully developed is audio. Audio processing involves processing speech and music. Raga Identification is a process of listening to a piece of music, synthesizing it in to sequence of notes and analyzing the sequence of notes for identifying the raga it follows. Here we identify patterns and establish relationships. A large, diverse database consisting of 30 hours recorded performances in 30 different ragas by 22 different performers was assembled to train and test the system. Classification was performed using support vector machines, also classification was done on 60 segments, and in unseen generalization accuracy was 75%.Where as our experiment in discriminating ragas with musical differences subtle. 20 different ragas each with 3-5 songs of each totally 90 songs are tested. The system performs with a result of 88%.

Improvements in audio compression together with increases in hard disk capacity and network bandwidth have made possible the creation of large personal music collections. The automatic analysis of music stored in audio format. Musical genres are categorical labels created and used by humans in order to structure the vast universe of music. This experiment covers manual annotation, automatic methods and usage based methods such as collaborative filtering.

2. CHARACTERISTIC OF INDIAN CLASSICAL MUSIC (ICM)

Indian music is broadly classified in to South Indian Carnatic Music and North Indian Hindustani Music. Both are rich in their own style and Carnatic Music is much more complex in the way the notes are arranged and rendered. Each raga has a swaroopam, the gamakas given to these swaras is termed as the raga laksanam which contains the arohanam(ascending passages i.e. pitch goes up), avarohanam(descent i.e pitch goes down).

Indian Music is based on Raga and Talam equivalent to melody and rhythm in western music. Raga is more complex than melody and scale in western music. Raga invokes the emotion of song. A raga is classified in to melakarta raga (parent) and

978-0-7695-3929-4/09 $26.00 © 2009 IEEE
DOI 10.1109/DASC.2009.155

janya raga (child).There are seven basic notes in carnatic music:

s	Sa	Sadjamam
r	Ri	Rishabam
g	Ga	Gandharam
m	Ma	Madhyamam
p	Pa	Panchamam
d	Da	Daivatam
n	Ni	Nishadham

Once the seventh note or the higher Sa is reached, the notes begin to descend in frequency from Sa to Ni to Da and so on by the same interval. These seven notes or swaras are not specific to Carnatic Music but are also common to Hindustani, Western, and other systems of music. In Carnatic music and Hindustani music, we call the seven swaras as Sa, Ri, Ga, Ma, Pa, Da, and Ni and in Western Music, the same seven swaras or notes are called doh, ray, me, fa, soh,lah,te,respectively.

The melakarta ragam will have all the seven swaras and janya ragam will have at least 5 of seven swaras. There are essentially 175 talams which forms the integral form of song for identification. The most stressed note is called the vadi and the second most stressed, traditionally a fifth or fourth away is called the samvadi. Indian Classical music (ICM uses approximately one hundred ragas, in which 50 are common. The micro tonal conforms to one of the twelve chromatic pitches of a standard just intoned scale.

3. EXISTING WORK

Krumhansl and Shephard, Castellano et al. have shown the stable pitch distributions gave rise to mental schemas that structure expectations and facilitate the processing of musical informations. David Huron has show that emotional adjectives used to describe a tone are highly correlated with that tone's frequency in a relevant corpus of music.

Anssi Kalpuri proposed a system for the automatic transcription of western music. Main emphasis is laid on find in it is multiple pitches of concurrent musical sounds. A.Krishnaswamy has described a method on how pitch tracking is useful for note Transcription of south Indian Classical Music

Parag Chordia classified one hundred thirty segments of sixty seconds each from thirteen ragas and stated about the Automatic raag classification of pitch tracked performances using pitch-class and pitch-class dyad distribution.Gaurav Pandey developed a system to automatically identify Yaman and Bhupali using a markov model for Identifying Hindustani music. A success rate of 77% was reported on thiry one sample in a two target test.

H.V.Sahasrabuddhe et al in their work said ragas have been modeled as finite automata which were constructed using information codified in standard texts on classical music. Rajeshwari Sridhar and T.V.Geetha done in swara identification and singer identification of carnatic music.

4. RELATED WORK

4.1 NOTE TRANSCRIPTION

Note transcription of music is listening to a piece of music and of writing down the musical notation for the sounds that constitute the piece or else the process of identifying the sequence of notes i.e. the swara script. Also converting an acoustic signal in to a symbolic representation and comprising of musical events and their parameters. Trying with various musical notes used and their intonation and should perform pitch tracking method to observe their performance in carnatic music is done.

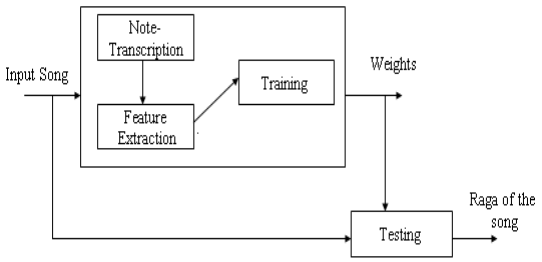

Figure 1: System Structure

Identification of the swaras in a given carnatic song is done with frequencies associated with the segments identified and the exact tagging of

swara is performed to find the 7- swara combinations in the given music signal. To implement this we have followed kalpuri sound onset and musical meter estimation in Indian Classical Music.

Music is represented as signal (polyphonic song), so note transcription involves signal processing techniques. Music is combination of swaras which inturn is set of frequencies (fundamental and harmonic frequencies). To human perception this frequency is called pitch. And we use multi pitch detection technique. Identifying all the different frequency segments in the song is called "frequency extraction". Fundamental frequency of each segment is calculated using the Auto correlation method with a frame-size of 50ms.

The relationship between swara and the frequency is not fixed it depends on the fundamental frequency of the note called as 'Sa' which is called 'Shruthi' or 'scale'.Shruthi is highly variable element. The scale of song is fixed to 24º Hz. Thus the frequency is mapped according to the defined ration in to the 36 swaras of Mandra, Madhya and Taara Saptaka.

4.2 FREQUENCY MAPPING AND RAGA DETERMINATION

Here we identify the individual swaras available and use the swaras and compare it with a database of Raga to identify the raga of the given song. The raga database is a table consisting of the name of the raga, the arohanam, avarohanam of raga in the form of swara components. We use a simple string matching to compare the identified swara pattern with the raga database and determint the raga of the input carnatic song.

4.3 RAGA IDENTIFICATION / ANNOTATION

In order to facilitate the creation of pitch profiles relevant to the particular tuning of the performances, each raga sample was labeled with the frequency value for the tonic. This was done manually, by tuning an oscillator and noting the value in HZ. Specific pitch sequence is improved by 87%.

In a 17 target experiment with 142 segments, classification accuracy of 94% was attained using 10- cross-validation due to limited by the size of database. Here a large assorted database consisting of 20 hours of recorded performances in 30 different ragas by 21 different performers was assembled to train and test the system. When a classification is done for 60 segments accuracy was 99% with cross validation whereas with unseen generalization accuracy was 75%.

4.4 PITCH DETECTION (ONSET/OFFSET SEGMENTATION)

Pitch detection was done using Harmonic Product Spectrum (HPS) algorithm. Each segment was divided in to 40ms frames using Gaussian Window. The threshold value of the harmonic is chosen so that on repeated iterations the signal gets separated. With first iteration the signal is separated in to voice and music. Next iteration extract output of first iteration and the signal is segmented to identify the frequency components present in the signal for swara identification yielding to raga identification.

The pitch detection distribution were calculated simply taking histograms of the pitch tracks. The bins corresponded to each note of five octave of chromatic scale centered about the tonic for the segment. The ratios of the just intoned scale and the tonic frequency were used to calculate the center of each bin, and the edges were determined as the log mean. The five octave were then folded in to one and the values normalized to create a pitch class distribution. This nulled any significance of octave errors in the HPS algorithm.

The segmentation algorithm determines the onset and offset followed by two level segmentation processes. There is a special song which is normally sung in a combination of Ragam where, every 4 or 5 lines will be sung in one Ragam the next 4 or 5 lines in another Ragam and so on. Such song is called as Ragamalika means composed of many ragas. We are trying to identify songs that are based on Ragam and Talam.

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 n2 o 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 g2	S r1 g2 m2 p d1 n2 s'/s'n2 p m2 g2 r1 s
9.Saurashtra	S r1g2 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 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

Table 1: Vakra Pairs

Any song is an integral multiple of a pre-specified Talam which indicates the order to identify the swaras of the song to find out the beginning of musical note or sound in which the amplitude rises from zero to an initial peak. Onset operates in time domain, frequency domain, phase domain or complex domain and includes changes like:

1. Increase in spectral energy
2. Changes in spectral energy distribution (spectral flux) or phrase)
3. Changes in detected pitch
4. Spectral patterns recognizable by machine learning techniques such as neural networks.

We take copy of the signal and convert it in to frequency domain and observe the change in spectral energy.

4.5 RHYTHMIC /PITCH CONTENT FEATURES

In each segment there is a possibility for the presence of more than one frequency component. After identifying frequency component corresponding to a swara in each segment, we determine the one with the highest energy as the one that corresponds to the swara. We now have to map this frequency component to the corresponding swara after identifying the fundamntal frequency of the signal. Since a singer can use a frequency in the range of 240 to 400 MHZ as the fundamental frequency:

Swara	Ratio	Swara	Ratio
S	1	M2	27/20
R1	32/31	P	3/2
R2	16/15	D1	128/81
R3	10/9	D2	8/5
G1	32/27	D3	5/3
G2	6/5	N1	16/9
G3	5/4	N2	9/5
M1	4/3	N3	15/8

Table 2: Swaras and their ratios with 'S'

The nature of inflexion on different notes (gamaka meend), characteristic phrases (swara sanchara chalan/pakad), the choice of notes and the Arohana- Avarohana pattern are noted here. A raga is constructed of 5 to 7 consistent swaras (melodic steps). In our system the features which are extracted from the input files are, swara combination, numbers of swaras used in raga, vakra pairs in Arohana and Avarohana. Swara combinations are represented in bits. This is then converted in to decimal value.

For example Mohana raga the swara sombination is : s r2 g2 p d2, so number of swara is : 05 and the equivalent bits are :

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 'S' is present in all the ragas so we can ignore the value and the binary sequence will be (01010010100) – 660 which says the number of distinguished swaras used in the raga. Similarly for Todi ragam the swara combination is: s r1 g1 m1 p d1 n1 and the number of swara are: 07

The Arohana and Avarohana can be linear or non-linear. In linear the pattern will be same as the swara combination and in non-linear some of the pairs which are not present in ascending/descending sequence of swara combination will be there. Such

pairs in non-linear are called vakra and some vakra pairs are listed below. The feature will be zero if the raga has linear arohana/avarohana.

5. EXPERIMENTAL STUDIES / RESULTS

Here we extract 6 features and named those as x1,x2,x3,x4,x5,x6 respectively. Here x1 indicates the number of distinguished swaras, x2 indicates the swara combination. The arohana vakra pairs values are assigned to x3,x4 and avarohana vakra pairs to x5,x6. Each possible pair in the swara script is assigned a value for unique identification of the pair. Unique labeling is done for each raga.

Raga No		Features					
		X1	X2	X3	X4	X5	X6
1	Abhogi	5	836	75	130	124	166
2	Anandbhairavi	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	Hansadhwam	5	657	60	103	96	151
7	Kalyani	7	1209	0	0	0	0
8	Kamavardhini	7	854	130	0	0	0
9	Kambhoji	7	725	0	0	0	0
10	Kanada	7	721	44	0	0	0
11	Kannada	7	726	103	75	0	0
12	Kedar	6	534	0	0	96	0
13	Khamaj	7	726	0	0	0	0
14	Madhyamavathi	7		320	102	68	138
15	Mayamalava	5	1241	60	0	0	0
16	Mohana	7	660	81	130	96	166
17	Saranga	7	693	0	0	0	0
18	Shankara	7	725	0	0	0	0
19	Todi	7	1370	117	0	0	0
20	Dhanyasi	7	1370		130	0	0

Table 3: Feature Extraction Result

Note transcription is done for polyphonic song (audio signal). The feature extraction is done for 90 songs of 50 ragas out of which the feature of 70 songs are used to train the network. The remaining 20 songs are used fo testing purpose which result in 95% accuracy. Matlab tool is used for the implementation.

6. RESULTS AND EVALUATION

A neural network is constructed by highly interconnected processing units 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, hyperbbolic tangent and sine. A very good account of neural networks can be found. The network models can be static or dynamic. Static netwokrs include single layer perceptron or adaptive linear element perceptrons. A perceptron or adaptive linear element refers to a computing unit. This forms the basic building block for neural networks.

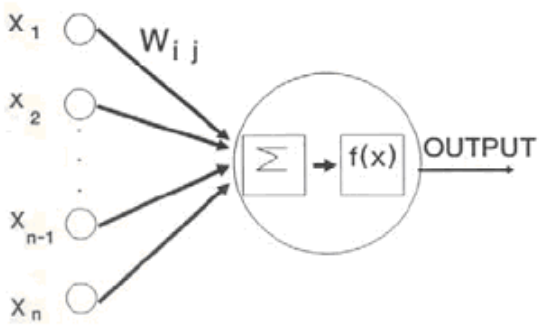

Figure 2 : Operation of a Neuron

The input signal is sampled at 44.1 KHz for five Melakarta Ragam for evaluating this algorithm. Here the protruding lines indicate the points of segmentation.

7. CONCLUSION AND FUTURE WOEK

Raga in an Indian Music is a very complex structure. To play song we analyze the sequence of notes for Raga Identification. The system works well for Polyphonic songs too. Thus eliminating qualms found in prior comparison. As we deal with only the rhythmic instruments and yet superimpose frequency components in the spectrum.

This work is based on low level signal features for Raga Identification. The best accuracy is 95% which is better than Gaussian Mixture Model (GMM) with Hidden Markov Model.

Allowing the system to recognize stylistic distinctions for which a human require extensive immersion in the genre to learn. As maintained larger database it's possible to make comparisons between instrument type and performance style. Thus analyzing melodies with gliding tones was developed.

8. REFERENCE

[1] Parag Chordia and Alex Rae, RAAG Recognition using pitch-class and pitch-class dyad distribution, Austrian Computer Society (OCG), 2007.

[2] Surendra Shetty and K.K.Achary, Raga Mining of Indian Music by Extracting Arohana-Avarohana Pattern,Acamey Publisher, 2009.

[3] Rajeswari Sridhar and T.V.Geetha. Raga Identifcation of Carnatic Music for Music Information Retrieval,2009.

[4] Tao Li and George Tzanetakis. Factors in Automatic Musical Genre Classification of Audio Signals, 2003.

[5] Emilia Gomez and Perfecto Herra. Comparative Analysis of Music Recordings from Western and Non-Western traditions by Automatic Tonal Feature Extraction, 2008.

[6] Hidden Markov Model based Recognition of Musical Pattern in South Indian Classical Music,2005.

[7] David Huron, Sweet Anticipation: Music and the Psychology of Expectation. MIT press 2006

[8] Carol Krumhansl, cognitive Foundations of Musical Pitch. Oxford University Press, 1990.

[9] Anssi Kalpur: Signal Processing Methods for the Automatic Transcription of Music, Thesis for the degree of doctor of technology, USA, 2004

[10] Garuav Pandy, Chaitanya Mishra, and Paul Ipe: Tansen Proceedigns of the 1st Indian international Conference on Artificial Intelligence,2003.

[11] A.Krishnaswamy : Application of Pitch Tracking to south Indian Music, Application of signal processing to Audio and Acoustics,2003, IEEE workshop.