Hindustani raga representation and identification: A transition probability based approach

Article · January 2011

2 authors:

Arindam (Andy) Bhattacharjee
University of Tuebingen

Narayanan Srinivasan
University of Allahabad

INTERNATIONAL JOURNAL OF MIND, BRAIN & COGNITION
VOL. 2, NO. 1-2, JAN-DEC 2011

Hindustani Raga Representation and Identification: A Transition Probability Based Approach

ARINDAM BHATTACHARJEE
NARAYANAN SRINIVASAN
University of Allahabad, Allahabad, India

ABSTRACT

Most research on perception and identification of complex musical sequences has focused on Western music and very little work has been done with Indian classical music. This paper focuses on an aspect of Indian classical music called the raga *system. The uniqueness of raga system is its wide scope for improvisation, which consequently poses problems for raga identification/classification. Moreover, this also leads to a question that if ragas are always improvised then how are they represented in our mind? In this paper we proposed a cognitively plausible representation – called Transition Probability Matrix (TPM) – for ragas and further evaluated its effectiveness in performing raga identification. We computed ten TPMs (one for each raga considered in this study) where each TPM characterizes the conditional probabilities embedded in* swara *(i.e. musical note) sequences of a raga. We acquired hundred arbitrary sequences from the ten ragas and performed raga classification; classification accuracy was 100%. We also evaluated the TPM representation by extracting relationships between ragas using multiple measures. The excellent classification results and the ability of TPMs to characterize raga relationships show that TPMs constitute an efficient cognitive representation for a raga.*

Keywords: Indian classical music, hindustani music, *raga*, *thaat*, transition probability, music classification, music representation, Markov model.

INTRODUCTION

The ubiquity of music aptly justifies a scientific curiosity: how do we perceive music? (Warren 2008). While a number of studies have investigated this question, most of them are dedicated to the understanding of how we perceive Western classical music; comparatively, fewer studies have investigated how we perceive Indian classical music. In the current paper, we looked into a possible computational mechanism involved in perceiving Indian classical music and propose a cognitively plausible representation for a particular category of Indian classical music. This representation not only captures the statistical as well as hierarchy of information available in the music, but also provides a template to perform efficient music recognition (or classification) within the Indian classical music domain.

Interestingly, Indian classical music do not have written musical scores and all renditions are improvised during performances. This might lead music listeners from other cultures to perceive Indian classical music as elusive and unstructured. Western classical music, in comparison, have written musical score sheets, which could provide researchers with a substrate to create musical stimuli to investigate how we process music. The representation proposed in the current paper might motivate researchers to create stimuli differently, and scientifically study Indian classical music and its cognitive processing.

Indian classical music exists in two different forms – *Hindustani* music and *Carnatic* music, which originated and was primarily practiced in Northern and Southern India respectively. However, both forms of Indian classical music are now practiced and performed beyond any geographical boundaries. In the current study, we focused on Hindustani music to extract and validate a representation that potentially captures the differences between different musical structures and enables us to differentiate between those Hindustani music structures.

The uniqueness of Indian classical music is in its melodic structure, which is based on the *raga* system. In ancient Indian texts, raga has been defined as "*ranjayanti iti raga*" – "THAT which charms is a *rāga*" (see Daniélou 1968: 91) which means raga is a musical rendering that pleases the mind. Jairazbhoy clarifies this vague definition of raga very well by explaining what is not a raga – "A *rāg* does not exist in any precise form in the sense that a symphony can be said to exist in score, but is a complex of latent possibilities. Although this seems to suggest an amorphous quality, each *rāg* is an independent musical entity with an ethos of its own, which becomes manifested through recognizable melodic patterns" (Jairazbhoy 1971: 32). In other words, raga could be considered as a tonal framework for composing different melodies. There are many different ragas practiced in Indian classical music (Reck 2005). The framework of each raga defines the relationship among its musical notes – for example, the hierarchy of note prevalence, or transition between different notes – that establishes typical melodic patterns of a raga and makes it recognizable.

Most of the research pertaining to Hindustani music have been primarily in the field of engineering where the motivation was either to create an automatic raga classification system (Chorid 2004; Chorida & Rae 2007a; Pandey, Mishra & Ipe 2003), or to create a raga note sequence generator (Sahasrabuddhe 1992; 1994; Das & Choudhury 2005). However, very few studies have investigated the structures of raga (Bel 1988; Choudhury & Ray 2003), or have reported any representation of Hindustani music (Chordia & Rae 2007b; Yardi & Chew 2004). Therefore, to enhance our understanding about Hindustani music, we explored the relatedness between ragas and compared those results with the behavioural results reported by Krumhansl and her colleagues (Castellano, Bharucha & Krumhansl 1984) and theoretical results reported by Jairazbhoy (1971).

HINDUSTANI MUSIC

Hindustani music scale is based on the just-intonated scale and the notes of any raga belong to the 12 chromatic tones of this scale (Krishnaswamy 2004). The musical notes (or *swaras*, as they are known in different Indian languages) of the Hindustani music scale are *shadaj*, *rishabh*, *gandhar*, *madhyam*, *pancham*, *dhaivat*, and *nishad*. Collectively these notes are known as *sargam*, equivalent to the Solfège of the Western music system (see Table 1). During singing, only the first syllable is sung, which are pronounced as *Sa*, *Ri*, *Ga*, *Ma*, *Pa*, *Dha*, and *Ni*; these notes are indicated in this study by the symbols *S*, *R*, *G*, *m*, *P*, *D*, and *N,* respectively.

Table 1. *Swaras and their corresponding notes in Western music*

Name	Pronunciation	Symbol	Solfège	Western
Sharaj	*Sa*	*S*	do	C
Komal Rishabh	*Re*	*r*	re(flat)	Db
Shuddha Rishabh	*Re*	*R*	re	D
Komal Gandhar	*Ga*	*g*	mi(flat)	Eb
Shuddha Gandhar	*Ga*	*G*	mi	E
Shuddha Madhyam	*Ma*	*m*	fa	F
Tivra Madhyam	*Ma*	*M*	fa(sharp)	F#
Pancham	*Pa*	*P*	sol	G
Komal Dhaivat	*Dha*	*d*	la(flat)	Ab
Shuddha Dhaivat	*Dha*	*D*	la	A
Komal Nishad	*Ni*	*n*	si(flat)	Bb
Shuddha Nishad	*Ni*	*N*	si	B

Similar to Western classical music, Hindustani music is tonal in nature and follows a hierarchy of notes (Krumhansl 1990). *S* or *Sa* is the tonic note. The scale contains only seven natural notes; however, each member may have two different forms except *S* (the tonic) and *P* (the perfect fifth), which exist in only one form. The natural notes are known as *shuddha swaras*, flat notes are the *komal swaras* and the sharp notes are called *tivra swaras*. The different forms of *R*, *G*, *D*, and *N* are *komal R*, *shuddha R*, *komal G*, *shuddha G*, *komal D*, *shuddha D* and *komal N*, *shuddha N,* respectively. *Shuddha M* and *tivra M* are the two variations of *M*. Thus, the total number of notes or intervals are twelve. A scale with only *shuddha* (or natural) notes is equivalent to the major scale of Western classical music, i.e. the distance between the notes are tone, tone, semitone, tone, tone, tone, semitone, which will complete an octave.

Every raga, in Hindustani music, has a particular ascending pattern of notes called *ārohan*, which is a progression towards the tonic of the next octave and a corresponding descending pattern of notes called *āvarohan*. The *arohan* and *avarohan* refers "to the most characteristic ascending and descending lines of a *rāg*" (see Jairazbhoy 1971: 38). In other words, the *arohan* and *avarohan* sequences of a raga indicate the notes that should be used in the raga during ascending and descending movements. Singing or playing the *arohan* and *avarohan* provides lot of information that might be useful to perform raga identification. Therefore, to validate our representation we generated *arohan* and *avarohan* sequences (i.e. the raga definition sequences) and compared those to the sequences reported in raga literature (Srivastav 1991).

The number of notes used in the *arohan* or ascending scale-pattern may not be equal to the number of notes used in the *avarohan* or descending scale-pattern. Each raga has a minimum of five notes and a maximum of seven notes (Reck 2005). Some notes are more frequently used than others during the rendition of any particular raga. The most frequently used note or *swara* is known as the *vādi swara* – the sonant note, and the *swara* that complements the *vadi swara* is known as the *samvādi swara* – the consonant note, which is often perfect fifth or fourth away from the *vadi swara*. The characteristic phrases of a raga are known as *pakad*, according to Hindustani music terminology, which establishes the ragas' identity and mood.

While, there are many ragas reported in raga literature, the most popular ones count to some 150 to 200 (Srivastav 1991). Hence, there was a need to classify them based on some organizing principle. According to Vishnu Narayan Bhatkhande (1860-1936) – one of the most influential musicologists in the field of North Indian classical music in the twentieth century – each one of the several traditional ragas is based on, or is a variation of, ten basic *thaats* (Srivastav 1991). *Thaats* – similar to "modes" in Western music – are musical scales or frameworks for classifying ragas based on the notes used in the raga. It is a system in which ten complete scale of seven notes each, in ascending order, are formulated to categorize maximum number of ragas within it. The ten *thaats* are *Bilawal*, *Kalyan*, *Khamaj*, *Bhairav*, *Poorvi*, *Marwa*, *Kafi*, *Asavari*, *Bhairavi* and *Todi*, which are shown in the first column of Table 2. Except one *thaat*, nine out of ten *thaats* mentioned above share their name with nine prominent ragas in Hindustani classical music. *Kalyan thaat* is the only exception; *Yaman raga* and *Kalyan thaat* corresponds to each other. These ragas are known as *ashray ragas*. It is important to note here that the structure of a *thaat* is for categorization purpose only and does not convey the meaning or characteristics of the ragas. Other ragas that have been categorized under one of these *thaats* are known as non-*ashray ragas*. The names of the *ashray ragas* are shown in the second column of Table 2 and a name of one non-*ashray raga* for each corresponding *thaat* are given in the third column. The scale and interval sizes of each *thaat* are also given in Table 2.

Table 2. *First column in this table shows the 10* thaats *considered in Hindustani music. Second column shows the list ragas considered in this study, which are also* ashray ragas *and third column shows the list of non-*ashray ragas. *The notes for each* thaat *are also shown and the numbers below the notes are the interval sizes in reference to the tonic note*

Name of *thaat*	Name of *ashray raga*	Non-*ashray raga*	Scale or *swaras*, and interval size relative to the tonic						
Asavari	*Asavari*	*Jaunpuri*	*S*	*R*	*g*	*m*	*P*	*D*	*n*
			Tonic	2	3	5	7	9	11
Bhairav	*Bhairav*	*Ahir-Bhairav*	*S*	*r*	*G*	*m*	*P*	*d*	*N*
			Tonic	1	4	5	7	8	11
Bhairavi	*Bhairavi*	*Malkauns*	*S*	*r*	*g*	*m*	*P*	*d*	*n*
			Tonic	1	3	5	7	8	10
Bilawal	*Bilawal*	*Shankara*	*S*	*R*	*G*	*m*	*P*	*D*	*N*
			Tonic	2	4	5	7	9	11
Kafi	*Kafi*	*Bageshri*	*S*	*R*	*g*	*M*	*P*	*D*	*N*
			Tonic	2	3	6	7	9	11
Khamaj	*Khamaj*	*desh*	*S*	*R*	*G*	*m*	*P*	*D*	*n*
			Tonic	2	4	5	7	9	10
Marwa	*Marwa*	*Bibhas*	*S*	*r*	*G*	*M*	*P*	*D*	*N*
			Tonic	1	4	6	7	9	11
Poorvi	*Poorvi*	*Puriya dhanashri*	*S*	*r*	*G*	*M*	*P*	*d*	*N*
			Tonic	1	4	6	7	8	11
Todi	*Todi*	*Multani*	*S*	*r*	*g*	*M*	*P*	*d*	*n*
			Tonic	1	3	6	7	8	10
Kalyan	*Yaman*	*Kedar*	*S*	*R*	*G*	*M*	*P*	*D*	*N*
			Tonic	2	4	6	7	9	11

Among the various characteristics of Hindustani music, which have been discussed above, improvisation during performance is a key element. Performance of a particular raga always shares the same basic scale and number of *swaras* or musical notes according to the defining patterns of that raga. However, there are lots of variations of the same raga that are performed at different occasions even by the same musician. These variations can be attributed to the mood of the performer, the style of the musician, creative capacity or number of years of musical experience. Interestingly, in spite of such differences, we are still able to identify the raga. Conversely, in another instance, two ragas may have identical notes and yet be very different in its form and melodic progressions; for example, *Shree* and *Puriya Dhanashri*, both have identical notes yet they are distinctly different in their melodic characteristics. Evidently, a simple knowledge about the notes used in a raga is not sufficient to identify the raga, and hence extracting a raga representational structure as well as evaluating its usefulness in raga classification could be particularly challenging and interesting.

While the ability to improvise during a performance allows musicians to express their creativity, this makes it almost impossible to transcribe every melody of each raga in Hindustani music. Therefore, we took a probabilistic approach to the problem. Due to the dynamic nature of Hindustani music and the relationships between *swaras* in a *swara* sequence, we used a Markov model to capture the statistical regularities between *swaras* in ragas. Markov models have been widely used in speech (Gold & Morgan 2004; Rabiner 1989) and music (Meyer 1957) – both Indian classical music (Hariram, Sheila, Krishna, Varadarajan & Venketeswaran 2001; Sinith & Rajeev 2007) and Western music (Chai & Vercoe 2001; Shao, Xu, & Kankanhalli 2004). We created a representation based on the sequences of notes in each raga for 10 different ragas listed in Table 2 (column 2). We then validated our representation by creating definition sequences (i.e. the *arohan-avarohan* sequences) of those ragas and by comparing the generated sequences to unique identifying sequences published in raga literature (Srivastav 1991).

Apart from validating our representation using definition sequences, we also performed classification of different snippets of music from those ragas. The purpose of performing raga classification were two folded: a) it provided an alternative approach to validate the efficacy of our representation, and b) it allowed us to investigate whether conditional probabilities incorporated within the sequence of notes of a raga are sufficient to attain the highest raga classification accuracy. We finally investigated the relationship between all the ragas considered in this study by estimating the Euclidean distance between ragas and inter-raga correlations.

METHODS

Ideally, to create the most precise representation of a raga all available recordings of that raga could be used; however, due to practical limitations we created our representations based on two musical pieces for each raga of approximately one hour in duration. We used two performances of each raga to add variation to our representation so that the representations are not merely a reflection of one performer's depiction of each raga. Moreover, this is our first attempt at creating a quantitative representation of ragas and acquiring raga performances from two different performers each seemed to be a reasonable beginning.

An important component in this study was transforming the music into distinct units i.e. *swaras* (or musical notes), which was necessary for creating the representation as well as validating the representation. We (the first author is trained in Hindustani music) manually transcribed the musical recordings to *swara* sequences by repeatedly listening to the raga performances. While manual transcription of the recordings might appear rudimentary and unimpressive, we reiterate that the objective of this study was not to create an automatic raga classification system. Our aim in this study was to create a cognitively plausible representation of the ragas. We realized that automatic *swara* identification (or musical pitch tracking in other musical forms) is a difficult task and has been a separate topic of scientific investigation (Cheveingne & Kawahara 2002; Hariram et al. 2001; Krishnaswamy 2003; Rossignol, Desain, Honing 2001; Sengupta, Dey, Nag & Datta 2002; Sridhar & Geetha 2006). Also, because our representation is computed from only two performances, errors in *swara* sequences could possibly weaken the reliability of the representations of the considered ragas. We recognize, however, that an automatic *swara* identification algorithm with perfect identification capability will make the transcription process easier, which in turn could be used to transcribe large corpora of recordings for each raga and to generate representation of those ragas.

Additionally, because we were performing raga classification to validate the representation of the ragas, it was important to have the test sequences without any erroneous notes as well, which might occur during automatic *swara* identification. Moreover, we were also interested in estimating raga classification accuracy by using only the conditional probability information of *swara* transitions available in the *swara* sequences. Previous studies have performed automatic raga classification using probabilistic approach; however, none of the studies were able to attain very high classification accuracy (Chordia & Rae 2007a; Pandey, Mishra & Ipe 2003). Therefore, it is unclear whether excellent classification accuracy was unattainable due to poor *swara* identification, or whether probabilistic information alone is insufficient to attain perfect raga classification accuracy. By performing manual *swara* transcription, we eliminated the influence of problems in *swara* identification that could potentially affect the performance of raga classification results. Hence, based on the motivation and objectives of this study, manual transcription of *swaras* presumably was the best method for transforming the music into *swara* sequences.

Computation of the transition probability matrix

We represented each raga in the form of a matrix that quantified the transitions between the *swaras*. Therefore, we called the representations as Transition Probability Matrix (TPM). After meticulously transcribing the performances of two musicians' rendition of any particular raga we identified musical phrases where the musicians used *swaras* other than the middle octave. We then shifted up (or down) those phrases from the lower (or higher) octave to the middle octave phrases. Although, by performing this operation we lost information about the octave where the musicians preferred to sing or play certain combination of *swaras*, we were able to capture the information about transitions from one *swara* to another *swara* which was our primary interest irrespective of the octave information. Moreover, by using performances from only two musicians' rendition of a raga encompassing all the octaves, our matrix representation would have been very sparse and the transition information between the *swaras* would have been less informative.

To create the matrix, we considered all 12 *swaras* of Hindustani music scale even though rarely any raga uses all 12 *swaras*. This, however, would capture two kinds of information: a) an immediate glance at the matrix will reveal which *swaras* are used in any particular raga, and b) because there are ragas where note transition between *swaras* are unidirectional (i.e. the sequence of musical notes are used only during *arohan* or only during *avarohan*), the matrix will represent the complete raga and not any particular direction of movement.

After we individually created a corpus of *swaras* transcribed from each raga, we determined how often any one particular *swara* was followed by another *swara* by counting the exact number of transitions from one *swara* to the other *swaras*. We repeated this procedure for all 12 *swaras* such that we calculated the frequency of transitions from every *swara* to each of the 12 *swaras*, which included transition to the same *swara* too. We then divided that number by the sum of all the transitions originating from any one particular *swara*. This depicted the probability of appearance of a *swara* conditioned on the transition-originating *swara*.

In the current study, we used a first order Markov model for raga classification as illustrated in Figure 1 (Gold & Morgan 2004). While using higher order Markov model might capture higher order statistical information about the raga, it would be computationally expensive to generate a representation. We wanted to investigate a simple Markov model and evaluate its feasibility before we explore more complicated statistical models. Because we decided to use 12 *swaras*, we obtained a two dimensional 12-rows-by-12-columns matrix where (i,j)th element represents the probability of transition from i^{th} *swara* to j^{th} *swara*. Mathematically, if we were to denote a *swara* at time t as $\omega(t)$, then a particular sequence of arbitrary length T could be denoted by ω = $\{\omega(1),\omega(2),\ldots,\omega(T)\}$. For example, a typical phrase of raga *Yaman* might be depicted as $\omega = \{N,R,G,R,N,R\}$, where $\omega(1)$ and $\omega(5)$ = N, and $\omega(2)$, $\omega(4)$ and $\omega(6)$ = R. The system can revisit a state at different steps, and not every state need to be visited. We also recognized that the transition probabilities may not be symmetric i.e. $a_{ij} \neq a_{ji}$ and a particular state may be visited in succession i.e. $a_{ii} \neq 0$ is a valid movement. Thus we created a TPM for each raga listed in Table 2 (column 2).

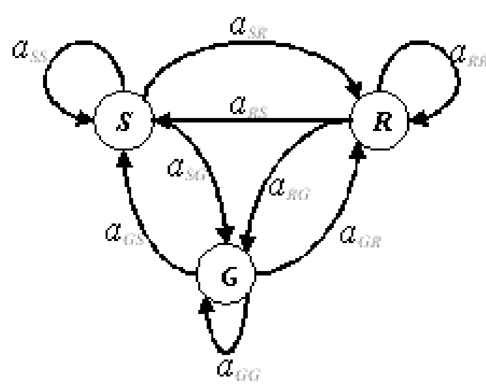

Figure 1. *Basic Markov model showing all possible transition between three example swaras - S, R, & G. In any raga the values of all these transitions will never be equal, which has been characterized in TPMs*

Validation of the TPM representation using arohan-avarohan *sequences*
Having computed the TPM for the ragas, we then validated whether the TPMs efficiently represent the ragas. Guided by the probability scores between transitions from one *swara* to other *swaras*, we generated the *arohan* and *avarohan* sequences (i.e. the definition ascending-descending *swara* sequence) of each raga considered in this study. To create the *arohan* sequence starting from *S*, which is the tonic note, we chose the *swara* that had the highest probability of

transitioning from *S* and we continued until we were to choose the *S* of the next octave; thus, completing a stepwise progression through a complete octave starting and ending at the tonic note *S*. The procedure to create the *avarohan* sequence was similar to the *arohan* sequence generation except the direction of progression was from the higher octave *S* to the lower octave *S*. Mathematically, the generation procedure to create the arohan-avarohan sequences can be described by transition probabilities $P(\omega_j(t{+}1) | \omega_i(t)) = a_{ij}$ – the time-independent probability of having swara ω_j at step t+1 given that the *swara* at time t was ω_i. We generated arohan-avarohan sequence for all the ragas considered in this study.

During the process of generation of arohan-avarohan sequences from the TPMs of each raga, if a conflict occurred between the highest transition probability value and the direction of movement along the octave, we considered the *swara* corresponding to the highest transition value in the direction of movement, i.e. towards the higher octave while extracting the *arohan* or ascending sequence and towards the lower octave while extracting the avarohan or descending sequence. Since *arohan-avarohan* sequences could be used to identify ragas, we generated those sequences based on TPMs computed for each raga and compared the generated definition sequence with the corresponding definition sequence for that particular raga that were already available to us through text and literature on Hindustani classical music (Srivastav 1991).

Validation of the TPM representation using raga classification

To test the efficiency of TPMs as a raga representation system, we performed raga classification using swara sequences manually transcribed from audio clips of 5 seconds in length from one to many renditions of each raga. By listing ten clips per raga, we obtained a total of hundred *swara*-sequences. We then assigned a score to each test sequence. The score was calculated by adding the logarithm of the transition probability values for the corresponding transition of *swaras* in the given sequence using the TPM of each raga, where TPM of each raga worked as a look-up table. Because we created TPMs for 10 ragas in this study, each test sequence received 10 different scores representing a score for each raga. The raga that yielded the highest score for each sequence was reported as the source raga of that particular test sequence. Finally, the classification accuracy was computed for all the sequences in the testing set.

Relationship between ragas

To understand the raga relationship and to evaluate the efficiency of TPM as raga representation, we calculated the Euclidean distance among the ragas using the TPMs. Based on those Euclidean distances as a metric, we then mapped the relationships between the ragas to a two-dimensional space representation using classical multidimensional scaling method. In addition to distance measure, the relationship between the ragas was also studied by analyzing the usage of *swaras* in a raga. We counted the total number of occurrences of each *swara* during the rendition and calculated the percentage of its appearance corresponding to each raga. Using the percentages of occurrence of *swaras* for all ragas as a measure for similarity, we then calculated the correlation coefficients between ragas. We reasoned that more similar ragas would show greater usage of same notes thus creating higher frequency for that note and eventually higher correlation coefficients between ragas. This could provide very important information about the tonal hierarchies of ragas as shown by Castellano, Bharucha & Krumhansl (1984).

All computations related to the creation of TPMs, i.e. validation of TPMs as a representation for Hindustani ragas, and investigation of relationships between ragas was performed using MATLAB.

RESULTS

Computation of the transition probability matrix

In this study we computed TPMs for ten ragas listed in Table 2 (column 2). As an example, the TPM for *Bilawal* raga is shown in Table 3. In this 12-by-12 matrix, the rows represent the origin note and the column represents the transitioned note. The numerical values in this table represent the probability of transition from a particular *swara* to another *swara* or the same *swara* in raga Bilawal. According to the TPM of this raga, for example, if a given *swara* is *S* then the probability of transition to *S* again is 0.07. Similarly, the probability of transition from *S* to *R* is 0.316, from *S* to *G* is 0.158, from *S* to *m* is 0, *S* to *P* is 0.053, *S* to *D* is 0.123 and *S* to *N* is 0.281. The reason for higher transition probability from *S* to *N* compared to that between *S* to *D* is because during *avarohan* the most used *swara* is *N*. Following the rules of *Bilawal* raga, the transition probability values from and to *r*, *g*, *d* and *n* are zero, which denotes that these *swaras* are not used in *Bilawal* raga. However, it should be noted that though *m* is used in the raga, the transition probability for *S* to *m* is zero, which means the likelihood of transition from *S* to *m* is extremely small. For some other raga, the probability of transition from *S* to *m* might be very high, which may be the characteristic of that raga, for example raga *Malkauns*.

Table 3. *Transition Probability Matrix* – Raga Bilawal. *The first row and column shows the 12 chromatic notes. The combination of row and column number will provide transition probability value when the transition occurs from the note in the row to the note in the column*

	S	*r*	*R*	*g*	*G*	*m*	*M*	*P*	*d*	*D*	*n*	*N*
S	0.070	0	0.316	0	0.158	0	0	0.053	0	0.123	0	0.281
r	0	0	0	0	0	0	0	0	0	0	0	0
R	0.484	0	0	0	0.463	0.021	0	0.021	0	0.011	0	0
g	0	0	0	0	0	0	0	0	0	0	0	0
G	0	0	0.272	0	0.008	0.344	0	0.376	0	0	0	0
m	0	0	0.402	0	0.578	0	0	0.020	0	0	0	0
M	0	0	0	0	0	0	0	0	0	0	0	0
P	0.010	0	0	0	0.010	0.495	0	0.067	0	0.410	0	0.010
d	0	0	0	0	0	0	0	0	0	0	0	0
D	0.010	0	0	0	0.010	0.052	0	0.448	0	0	0	0.479
n	0	0	0	0	0	0	0	0	0	0	0	0
N	0.343	0	0.015	0	0	0	0	0.015	0	0.612	0	0.015

Validation of the TPM representation using arohan-avarohan *sequences*

As described in the methods section, we created *arohan-avarohan* sequences for all ten ragas considered in this study. For example, here we discuss the generation of the *arohan-avarohan* sequences in Bilawal raga. It can be seen in Table 3, that the maximum probability of transition from *S* to any other *swara* is 0.316, which corresponds to *R*, so *S* is deemed to be followed by *R*. After *R* the *swara* that should follow is *S* according to the highest transition value of 0.484. Since in ascending sequence or *arohan*, the direction of movement is higher up the octave, the next highest transition value from *R* has to be considered, which is 0.463. This represents the transition to *G* i.e. the next *swara* that comes after *R* is *G*. The *swara* that should come after *G* could be either *m* (0.344) or it could be *P* (0.376), both of which are allowed for *Bilawal* raga (Srivastav, 1991). Similarly, we created the *avarohan* sequence for *Bilawal* raga. Interestingly, the TPM shows that there is a high probability of transition from *m* to *G* (in the *avarohan*) that is characteristic of *Bilawal* and the TPM reliably captures this characteristic. Therefore, the *arohan* and *avarohan* sequences determined using the transition probability values from *Bilawal* TPM are *S R G m P D N S* and *S N D P m G R S* respectively. This agrees well with the definition sequences for *Bilawal* as defined in Hindustani music literature (Srivastav 1991). *Arohan-avarohan* sequences generated using TPM for all ten ragas matched completely to those reported in Hindustani music texts (Srivastav 1991). Successful matching of the computed *arohan-avarohan* sequences with the defined *arohan* and *avarohan* sequences from TPMs demonstrates that TPMs can be a valid representation and could be useful for further identification or other applications based on ragas.

Validation of the TPM representation using raga classification

To test the validity of TPMs as a representation for Hindustani raga music we obtained ten *swara*-sequences from each raga (see Table 2 for list of ragas in column 2). Following the procedure of calculating *swara*-sequence score, as described in the methods section, we determined the likelihood that any particular *swara*-sequence belongs to one of those ten ragas. The TPM that yielded the highest score was considered as the parent raga that could generate a sequence similar to the test *swara*-sequence. Hence, based on the computed score, all 100 *swara*-sequences were classified into 10 ragas. Since we acquired the *swara*-sequences from different renditions of the ragas considered in this study, we compared whether the TPM score based raga identification matched the actual raga name. Interestingly, all *swara*-sequences were correctly classified into their parent ragas yielding 100% classification accuracy. The perfect classification accuracy obtained with arbitrary *swara*-sequences show that TPMs could be successfully used for raga identification.

Relationship between ragas

To explore the relationship between the ragas, we calculated the Euclidean distances between the ragas (as described in the methods section). The resulting distances between the ragas computed using TPMs are shown in Table 4. In the 10-by-10 matrix (because we considered 10 ragas in this study), the values in the (i,j)th position provides the Euclidean distance between ith raga and jth raga. As the distance measure, which were computed based on the 144 values (12-by-12 matrix of *swaras*) in a TPM, were obtained in the a multidimensional space, it posed a difficulty in interpretation; so, we chose to project it to a 2-dimensional pictorial representation of the raga distances using multidimensional scaling shown in Figure 2 where each point in the 2-dimensional map represents a raga.

Table 4. *Euclidean distance look-up table for ragas computed between the TPMs of each raga*

	Asavari	*Bilawal*	*Bhairav*	*Bhairavi*	*Kafi*	*Khamaj*	*Marwa*	*Poorvi*	*Todi*	*Yaman*
Asavari	0	5.401	4.877	3.918	2.779	5.998	6.139	6.641	6.343	6.372
Bilawal	5.401	0	3.538	5.290	3.812	3.052	4.037	4.553	5.213	2.776
Bhairav	4.877	3.538	0	2.981	4.607	4.333	3.678	2.506	3.298	5.053
Bhairavi	3.918	5.290	2.981	0	3.624	6.534	4.797	4.832	3.811	5.960
Kafi	2.779	3.812	4.607	3.624	0	3.415	5.009	6.262	5.578	4.823
Khamaj	5.998	3.052	4.333	6.534	3.415	0	6.404	6.217	6.445	5.131
Marwa	6.139	4.037	3.678	4.797	5.009	6.404	0	3.179	4.461	3.056
Poorvi	6.641	4.553	2.506	4.832	6.262	6.217	3.179	0	3.226	4.133
Todi	6.343	5.213	3.298	3.811	5.578	6.445	4.461	5.000	0	4.354
Yaman	6.372	2.776	5.053	5.960	4.823	5.131	3.056	4.133	4.354	0

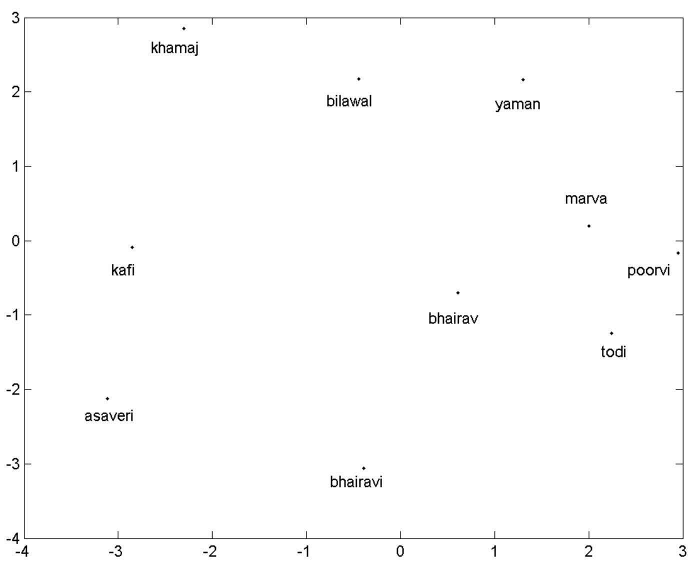

Figure 2. *Raga representation in two dimensions*

As another measure of relatedness between the ragas, we calculated the percentage of occurrences of every *swara* for each raga that were considered in this study (see Table 5), which allowed us to create *swara* -frequency profiles for each raga. Figure 3 shows the *swara*-frequency profile for *Bilawal* and *Khamaj* ragas, where the x-axis shows all the 12 *swaras* for the Hindustani music scale and the y-axis shows the percentage values of *swara* occurrence frequency. Next we computed the correlation coefficients between the ragas. These inter-raga correlations produced a matrix of values that indicates the degree of relatedness between each pair of ragas, which are shown in Table 6. High correlation in suggests similar usage of *swaras* and hence strong relatedness between ragas. For example, the correlation value for *Bilawal* and *Khamaj* is 0.865, which is quite strong and it is evident in the *swara* profile for *Bilawal* and Khamaj shown in Figure 3.

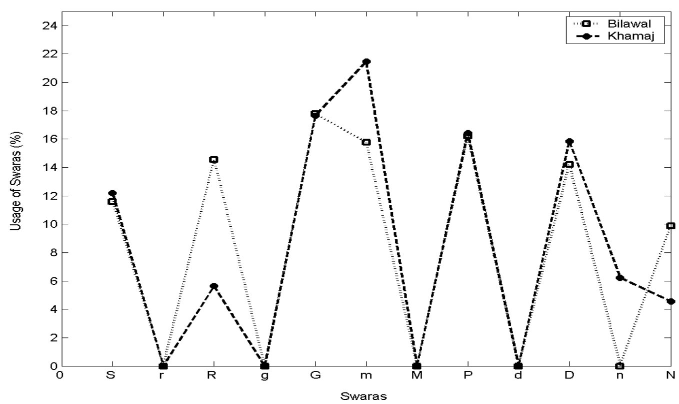

Figure 3. *Relation between* Khamaj *and* Bilawal

Table 5. *Percentage of* swara *occurrence of* swaras *or notes in every raga presented in this study*

	S	*R*	*R*	*g*	*G*	*m*	*M*	*P*	*d*	*D*	*n*	*N*
Asavari	13.50	0	14.88	7.21	0	16.41	0	26.23	17.02	0	4.75	0
Bilawal	11.59	0	14.53	0	17.77	15.77	0	16.23	0	14.22	0	9.89
Bhairav	14.84	16.41	0	0	14.65	19.53	0	14.45	13.09	0	0	7.03
Bhairavi	18.64	11.56	0	15.87	0	15.72	0	16.18	12.17	0	9.86	0
Kafi	12.25	0	16.62	15.07	0	15.63	0	19.86	0	11.41	9.16	0
Khamaj	12.18	0	5.63	0	17.66	21.46	0	16.44	0	15.83	6.24	4.57
Marwa	6.06	23.94	0	0	17.89	0	18.45	0	0	19.58	0	14.08
Poorvi	10.53	16.37	0	0	22.81	4.53	15.35	11.99	10.09	0	0	8.33
Todi	15.14	20.95	0	22.43	0	0	14.59	2.84	14.46	0	0	9.60
Yaman	12.13	0	20.26	0	14.84	0	12.98	11.27	0	12.41	0	16.12

Table 6. *Correlation coefficient chart showing the raga relatedness computed based on percentages of note occurrence in each raga*

	Asavari	*Bilawal*	*Bhairav*	*Bhairavi*	*Kafi*	*Khamaj*	*Marwa*	*Poorvi*	*Todi*	*Yaman*
Asavari	1	0.296	0.371	0.614	0.657	0.289	-0.789	-0.170	-0.168	-0.095
Bilawal	0.296	1	0.319	-0.165	0.429	0.865	-0.034	0.066	-0.697	0.594
Bhairav	0.371	0.319	1	0.499	-0.087	0.427	0.013	0.574	0.043	-0.240
Bhairavi	0.614	-0.165	0.499	1	0.440	0.087	-0.544	-0.090	0.359	-0.650
Kafi	0.657	0.429	-0.087	0.440	1	0.451	-0.647	-0.600	-0.333	0.026
Khamaj	0.289	0.865	0.427	0.087	0.451	1	-0.082	0.045	-0.720	0.215
Marwa	-0.789	-0.034	0.013	-0.544	-0.647	-0.082	1	0.535	0.176	0.264
Poorvi	-0.173	0.066	0.574	-0.093	-0.604	0.045	0.535	1	0.207	0.157
Todi	-0.168	-0.697	0.043	0.359	-0.333	-0.720	0.176	0.207	1	-0.365
Yaman	-0.095	0.594	-0.240	-0.650	0.026	0.215	0.264	0.157	-0.365	1

DISCUSSION

In the current study, we created matrix representations for ragas by quantifying the transition probabilities of *swara* (or note) sequences. We tested the validity of the representations by comparing TPM based *arohan-avarohan* sequences and the *arohan-avarohan* sequences informed by the Hindustani music literature (Srivastav 1991). As an additional validation step, we performed raga classification and we achieved excellent results. We also explored the relationship between the ragas and compared our results to that of previously published studies. Interestingly in the current study, the relationship between the ragas derived from the TPMs independently matched the observations reported in a theoretical study (Jairazbhoy 1971) and behavioural study (Castellano, Bharucha & Krumhansl 1984). Thus, the overall results suggest that the probability based representation holds lot of information about Hindustani ragas and that TPMs could be used as a good template to perform raga classification and *swara*-sequence generation.

Raga representation

Representation of ragas can be investigated at two different levels – a) a structural representation that might capture different aspects of Hindustani music, and b) a cognitive representation that might model how we perceive Hindustani music. To the best of our knowledge there has been only one study that truly investigated Hindustani music to create structural representation of ragas. Yardi & Chew (2004) constructed representations of ten (*ashray*) ragas using “Harmonic Network” – a concept used in Western tonal music to study pitch relations (Cohn 1998). The harmonic network is a network of twelve notes repeated several times, where each note is connected to six other notes. This network shows the harmonic relationship between three-note chords (also called triads) connected in a triangular fashion. Yardi & Chew (2004) depicted a raga on to the harmonic network by connecting the musical notes that are used in the raga, which allowed them to obtain a geometric shape for each raga. Based on those shapes the authors predicted the emotion that the corresponding raga should evoke, which they compared to the raga-emotions reported in the Hindustani music literature. While this model proposed by Yardi & Chew (2004) predicted the emotion of the ragas, it did not provide any information about the tonal hierarchy, or the co-occurrence statistics of the notes – a feature that could be useful in performing raga identification. Interestingly, although TPM representation does not create any geometrical shapes, the 12-by-12 matrix structurally provides quick information about the *swaras* used in a particular raga similar to the representations created by Yardi & Chew (2004). We, however, acknowledge that we have not tested whether the TPMs capture the emotions of the raga; therefore, commenting on use of TPMs in raga-emotion prediction is beyond the scope of this study.

In the current study, our primary objective was to create a cognitive representation of Hindustani music perception. To the best of our knowledge this is the first study to create an explicit cognitive representation of Hindustani raga music. We derived our motivation from techniques in automatic speech processing (Lu & Hankinson 2000; Gold & Morgan 2004) and music cognition studies performed by Krumhansl and her colleagues (Krumhansl 1990; Krumhansl Toivanene, Eerola, Toiviainen, Jarvinen & Louhivuori 2000). When there are finite elements, for example, the number of phonemes in a language, statistical regularities can be captured by quantifying the transition between the elements. This also justifies the use of Markov chain model in creating the raga representation.

Interestingly in the field of music, Krumhansl et al. (2000) observed that two-note transitions could capture musical information in yoiks – a style of music performed by Sami people who are indigenous to Scandinavian Peninsula – and people from Sami culture were particularly sensitive to the experimental manipulation of statistical regularities in the yoik music stimuli. These studies suggested that, since Hindustani ragas use finite number of notes and there are statistical regularities in music, transition probabilities between the notes would be able to capture this information and support the plausibility of a cognitive representation for Hindustani ragas. Additionally, research has shown that from a very early age humans become sensitive to the statistical structure of their acoustic environment. For example, research by Saffran and her colleagues showed that 8-month old infants are sensitive to conditional probabilities embedded in syllable sequences (Saffran, Aslin & Newport 1996) and tone sequences (Saffran, Johnson, Aslin & Newport 1999), which eventually leads to the development of sensitivity to distribution of pitch or tonality in music (Kessler, Hansen & Shepard 1984; Oram & Cuddy 1995; Krumhansl et al. 2000). Therefore, the presumption of a cognitive representation for Hindustani ragas that are based on transition probabilities seems to be highly plausibly.

In Hindustani music research, the idea of statistical regularities has been demonstrated previously, albeit in an implicit manner by Sahasrabuddhe (1992, 1994) and more recently by Das & Choudhury (2005). All three studies used finite state machine to create swara sequences for particular ragas. The states or nodes for both models were the notes of the raga which the authors considered in their raga. Authors of both studies were successful in demonstrating that raga information were well captured by the statistical models they used, which once again supports the use of statistical regularities as a substrate to create a cognitive representation for Hindustani ragas in the current study.

Validation of the TPM representation

We tested the validity of the TPM representation using two different approaches – a) *swara* sequence synthesis approach, which could be compared to the approach reported by the raga synthesis studies (Sahasrabuddhe 1992, 1994; Das & Choudhury 2005), and b) raga classification approach. The excellent results of both approaches strongly support TPMs as a valid representation for Hindustani ragas. While the results obtain using both validation techniques were very encouraging, we were particularly interested in the results observed in the raga classification approach; the reasons are – a) we believe that raga classification approach is relatively more robust technique for the validation process and we acquired perfect classification results, which provides reassurance of the objective of our study, and b) although, statistical information of tonality in ragas was used earlier in automatic raga recognition studies (Chordia & Rae 2007a; Pandey, Mishra & Ipe

2003), the true effect of the statistical regularities in raga classification was not clear; as the aim of those studies of were to perform automatic *swara* recognition followed by raga classification.

In the current study we eliminated the step of automatic *swara* recognition, which provided us with the opportunity to utilize only the statistical regularities embedded in *swara* sequences in different ragas. The perfect classification accuracy achieved using this procedure suggests that – a) statistical regularities hold sufficient information necessary to perform raga classification, and b) the results obtained by the above mentioned studies might be "contaminated" by the accuracy of automatic swara recognition. Since the studies by Pandey, Mishra & Ipe (2003) and Chordia & Rae (2007a) used statistical information to perform raga classification (or recognition) we here further discuss these studies in light of the results obtain in the current study.

Pandey, Mishra & Ipe (2003) chose two ragas to test their model – a) *Yaman kalyan*, and b) *Bhoopali*, both belonging to the same *thaat* (see section: Hindustani music for definition). The *swaras* (or musical notes) used in *Yaman* are: *S R G M P D N* (*M* is sharp compared to *m* in *Bilawal*, see Table 2); in Western nomenclature if C is considered as the tonic note then by replacing *S* with C, the notes used in *Yaman* will be: C D E F# G A B. The *swaras* used in *Bhoopali* are: *S R G P D*; according to Western nomenclature considering C as the tonic note and replacing *S* with C, the notes of *Bhoopali* will be: C D E G A. The authors performed automatic note transcription and then used hidden Markov model for identifying whether a given sequence belongs to *Yaman* or *Bhoopali*. They tested all together 31 samples of *Yaman* (15 samples) and *Bhoopali* (16 samples) and acquired 77% average classification accuracy; relatively higher classification of *Yaman* (80% accuracy) samples was observed compared to *Bhoopali* samples (75% accuracy). Since the test sequences were not published by the authors, we do not know the reason for slightly higher classification of *Yaman* samples; however, it must be noted that *Bhoopali* does not contain two notes *N* and *M*. While it is unclear why Pandey et al. (2003) only chose two ragas, they however, chose ragas from the same thaat such that both ragas share a subset of the *swaras* of *Kalyan thaat.* They reasoned that having ragas belonging to the same note class or *thaat* would make the task of the classifier more difficult. Although these two ragas belong to the same set of notes, as mentioned above, the subset of notes are different; mere presence of the notes *N* and *M* should make it easily identifiable that the *swara* sequence belongs to Yaman raga instead of *Bhoopali*. Of course, it would be very interesting to see the classification accuracy of *swara* sequences that did not contain the notes *N* and *M*. In the current study, our technique did not have any information about which notes were present in the test sequences; rather it solely relied on transition probability scores to determine whether a sequence belonged to *Yaman* or *Bhoopali*.

Chordia & Rae (2007a) chose an impressive set of 31 ragas. The authors performed automatic *swara* identification; however, they did not report the accuracy of this procedure. The authors then used neural networks techniques to train their network and further perform raga classification. They used two different approaches for raga classification – a) pitch-class distributions, which is similar to pitch histogram or *swara* usage profiles for every raga, and b) pitch-class dyad distribution, which is similar to distribution of note transitions. The authors trained their network with a huge corpus of *swara* sequences and then performed the validation process using the approaches mentioned above. Additionally, during testing the authors used sequences that were part of the neural network training set, and sequences that were unfamiliar to their network specialized to perform raga classification. When the authors used sequences that were part of the training set, they obtained 78% classification accuracy using pitch frequency information and obtained 97.1% classification accuracy using note-pair distribution information. However, when the authors used sequences that were unfamiliar to their classifying network, the classification accuracy dropped to 75.2% for pitch frequency approach and to 57.1% for note-pair distribution information approach. By using cross-validation procedure, i.e. testing a sequence with both approaches, they attained 99% classification accuracy for sequences used in the training set and 73.7% classification accuracy for sequences that were unfamiliar to the raga-classifying network. Interestingly, a previous work by the lead author of this study acquired perfect classification accuracy by using pitch frequency information (Chordia 2004); however, only 13 ragas were used in that study.

Among the known studies that performed raga recognition in Hindustani classical music, there is a particularly interesting study using signal processing techniques (Datta, Sengupta, Dey & Nag 2001). The authors used only four ragas representing four different thaats; however, they used impressively large corpus of *swara* sequences. Datta et al. (2001) reported the best average raga recognition accuracy of 52.5%, which is slightly lower than the accuracy rate reported by Chordia & Rae (2007a) for the unfamiliar sequences. This suggests that statistical information imbibed in *swara* sequences contain lot of information and might be a good substrate to perform raga classification.

Relationship between ragas

While the TPM representation in the current study resulted in excellent classification accuracy, the studies by Datta et al. (2001) and Chordia & Rae (2007a) observed several misclassifications. Such misclassifications could be a result of similarity between ragas and it would be important to capture the similarity between ragas. Therefore, using the TPM representations for all ten ragas, we computed relatedness between those ragas. We calculated Euclidean distances between the ragas and projected in a 2-dimensional representation, shown in Figure 2 where each point denotes a raga. Interestingly, the 2-dimensional representation of raga space computed in this study is very similar to that reported by Bel (1998) and the *thaat* space reported by Chordia & Rae (2007b). To create the representational map, the authors in both studies considered a Boolean representation for the ragas where they scored 1 or 0 for notes that are present or absent, respectively, in the raga. Chordia & Rae (2007b) reportedly used Self-Organizing Maps (a neural network technique) to estimate the Euclidean distances between the ragas. Bel (1998) used classical multidimensional technique (similar to the current study) to calculate the Euclidean distances between the ragas.

Although Bel (1998) classified 30 ragas in an attempt to study the deviations in ragas, comparison between the current study (Figure 2) and the study by Bel (1998, Figure 6) shows similar trends. For example, the distance between *Marwa* and *Poorvi* is smaller than the distance between *Marwa* and *Yaman*; *Marwa*, *Todi*, *Bhairavi, Jaunpuri* and *Darbari Kanada* (both ragas belong to *Asavari thaat*) are depicted in the periphery of the representation similar to *Marwa*, *Todi*, *Bhairavi*, and *Asavari* as shown in Figure 2 in the current study. While we used TPM representations to calculate the Euclidean distances, Bel (1998) used a Boolean representation, and that might be the reason for some variations observed between the two graphical representations, i.e. Figure 2 of the current study and Figure 6 from Bel (1998).

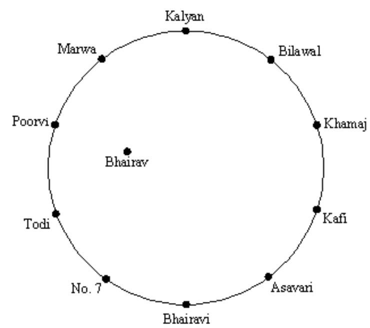

Figure 4. *The ten ragas placed on the circle of thaats based on Castellano, Bharucha & Krumhansl (1984)*

It is interesting to note that the *thaat* space representation created by Choridia & Rae (2007b) and the theoretical map proposed by Jairazbhoy (1971) called the circle of *thaats* are very similar. Additionally, it is very encouraging to observe that the raga-space representation in the current study is similar to the map proposed by Jairazbhoy (1971) except that the raga-space representation is not circular in Figure 2. In the theoretical map proposed by Jairazbhoy (1971), all the ragas (except *Bhairav*) lie in the perimeter of a circle and are equally spaced (see Figure 4). We observed similar trends in Figure 2 where the lines connecting near by ragas is not a circle but *Bhairav* is within the perimeter of the polygon. We believe that the asymmetrical relationships between one raga and another raga, may suggest a more plausible raga-space cognitive map displaying relationships between ragas.

While the above-mentioned studies were computational (Bel 1998; Chordia & Rae 2007b) and theoretical (Jairazbhoy 1971), we also observed similarities between the results obtained in the current study to that obtained in a behavioural study (Castellano, Bharucha & Krumhansl 1984). This similarity in results is very encouraging because the aim of the study by Castellano, Bharucha & Krumhansl (1984) was to probe the tonal hierarchy of Hindustani ragas in Indian listeners and in Western listeners. Statistical regularities give rise to tonal hierarchies, which the authors tested using a procedure called "probe tone" method. In this procedure, a *swara* sequence of certain number of notes are played to the participant and then the participant is asked to judge how well the last *swara* fit with the sequences of *swaras*.

Probing each of the 12 notes following different sequences, Castellano, Bharucha & Krumhansl (1984) created a *swara*-fitting profile. Interestingly, the authors noticed that *swaras* that occurred more often in a raga were also rated as highly fitting "last note" in the probe tone procedure.

We compared *swara*-frequency profile computed in the current study to the rating profile reported by Castellano, Bharucha & Krumhansl (1984) and found similarities between them. Based on rating profile, Castellano, Bharucha & Krumhansl (1984) also created raga-space using multi-dimensional scaling. Once again, the results obtained by Castellano, Bharucha & Krumhansl (1984) depicting in the raga-space showed similar trends as shown in Figure 3 in the current study. For example, based on probe-tone rating profile Castellano, Bharucha & Krumhansl (1984) observed smaller raga distance between *Marwa* and *Poorvi* compared to *Marwa* and *Yaman*, a trend obtained in current study as well (see Figure 2). We noticed another similarity between the results reported by Castellano, Bharucha & Krumhansl (1984) and the results obtained in the current study. Castellano, Bharucha & Krumhansl (1984) using the rating profiles obtained from their participants calculated inter-raga correlations and observed that Yaman and Bilawal are more correlated (coefficient = 0.775) compared to the correlation between Yaman and Bhairavi (coefficient = -0.081). In the current study, using swara-frequency profiles, we calculated correlations between each raga to all the other ragas (see Table 6). We also observed that the correlation between *Yaman* and *Bilawal* (coefficient = 0.594) is stronger than the correlation between *Yaman* and *Bhairavi* (coefficient = -0.65). Therefore, the similarity between the results obtain in the current study and that reported in Castellano, Bharucha & Krumhansl (1984) strongly support the plausibility of TPMs as a plausible cognitive representation of Hindustani ragas.

The relationship between the ragas from examined through various aspects makes the current study very interesting. In one hand, the distance measures between the ragas, calculated using the TPM of each raga, provide information about the layout of ragas in a two dimension space. It also helps to graphically compare the 2-dimensional representation to the theoretical representation proposed by Jairazbhoy (1971). On the other hand, the relationship based on swara profile helps to compare the results of this study to that of human performance data reported by Castellano, Bharucha & Krumhansl (1984). Moreover, it is interesting to notice that, while the correlation matrix (Table 6) is computed using the percentage of *swara* usage and does not contain the transition information between the *swaras*, the distance representation (Figure 2) is computed using the TPMs and both are highly similar. Higher the positive correlation coefficients between two ragas in Table 6 closer are those ragas in the 2-dimensional representation in Figure 2. Negative correlations predict that the ragas are in the opposite half of the 2-dimensional representation. The relationship between the current study, behavioural observation of Castellano, Bharucha & Krumhansl (1984) and theoretical representation of ragas of Jairazbhoy (1971) is highly encouraging to consider TPMs as a simple yet efficient approach for representing ragas.

Implications, limitations, and future scope

While the primary objective of the current study was to create an effective representation for Hindustani raga music, we also demonstrated that TPM representation could be used as a template for performing raga classification. The technique for creating TPM representations developed in this study could be used to create TPMs for many different ragas and their relatedness could be explored. Additionally, researchers interested in testing behavioural implications of transition probabilities can use similar techniques to generate TPM representations as well as experimental stimuli for conducting scientific investigations.

Recently we conducted a behavioural study to investigate whether TPM based representation scheme is dependent of enculturation; for example, are listeners sensitive to transition probabilities for a raga and does this sensitivity depend on the musical culture of the listener (Bhattacharjee, Srinivasan, & Trainor, in revision). It has only been possibly to conduct such an experiment due to the presence of TPM representations. In that study, Indian and Western listeners were presented with *swara* sequences that either conformed or violated the TPM representations of *Yaman* raga and only Indian listeners were able to identify the sequences that violated the transition probabilities between the *swaras*. The results of that study strongly suggest that listeners internalize the statistical relations and probably use that information to classify sequences of *swaras* into different ragas.

A limitation of TPM representation of a raga is that although TPMs effectively captures the melodic progression of ragas, it fails to capture any temporal aspect associated with melodic progression, for example, the duration of a note sung or played, or tempo of a given piece of music. Additionally, the current TPM represents only single-step transition Markov model. If the order of the steps are increased the efficiency of TPMs in performing raga classification might be excellent for whole set of Hindustani ragas.

While we have proposed and evaluated a computationally simple cognitive representation of Hindustani ragas based on Markov models, another interesting approach would be to explore the rules of Hindustani ragas from the linguistic point of view, i.e. the grammar in Hindustani ragas. Linguists have been interested in the fact that infants learn grammatical rules of their first language without specific instruction. In music too the rules of music are implicitly learned by the listeners without any explicit instructions (Bhattacharjee, Srinivasan & Trainor, in revision). Such an approach has been investigated for Carnatic ragas (Vijayakrishnan 2007) and it would be of interest to explore the parallels in Hindustani music system. However, given the scope of the current paper we limited ourselves to tackle only one question at time, i.e. creating a simple cognitively plausible representation for Hindustani ragas.

At present, TPMs were computed for ten ragas that belong to ten thaats. For future work, this may be extended to more ragas per *thaat* that will enable us to classify all the ragas in Hindustani Music. That classification might be better than the *thaat* classification, which is often criticized as a poor form of raga classification (Choudhury & Ray 2003). The TPM based representation approach can also be extended to the ragas in Carnatic music system and may enable us to quantitatively study the similarities and differences between the two music systems.

CONCLUSIONS

In conclusion, the current study successfully demonstrated the effectiveness of TPMs as a representation for ragas. We explored the relationship between ragas using Euclidean distance measure and inter-raga correlational measure. Successful classification results indicate the effectiveness of statistical information present in sequences of *swaras* for raga identification. Our aim was not to develop an automated system to perform raga recognition; we were rather interested in investigating the structures underlying ragas for understanding raga perception and identification. The reasonable closeness of behavioural data with probe tone ratings and percentage of *swaras* in ragas also point to the usefulness of statistical relationships in music perception. Recently Janata and colleagues (Janata, Birk, Van Horn, Leman, Tillman & Bharucha 2002) were able to identify cortical correlates of tonal structure underlying Western music. Perhaps in future, researchers might be able to identify neural substrates that correlated with a multidimensional cognitive map of ragas. Meanwhile, we derive a hint from the results of the current study that there is a possibility of a multidimensional cognitive map of musical entities based on statistical relationships in our mind; perceptual evidence based on further behavioural experiments for TPMs and other musical characterizations are necessary for finding the cognitive basis of raga classification by humans.

REFERENCES

Bel, B. 1988. Raga: Approches conceptuelles et expérimentales. Actes du colloque "Structures Musicales et Assistance Informatique", Marseille: France. Available online: <http://hal.archives-ouvertes.fr/hal-00008280/en/>

Bhattacharjee, A., Srinivasan, N. & Trainor, L.J. (in revision). Sensitivity to sequential structure in ragas: a comparison of Indian and Western listeners.

Castellano, M. A., Bharucha, J. J., & Krumhansl, C. L. 1984. Tonal hierarchies in the music of north India. *Journal of Experimental Psychology,* 113, 394-412.

Chai, W. & Vercoe, B. 2001. Folk music classification using hidden markov models. In proceedings of *International Conference on Artificial Intelligence*, Las Vegas, NV.

Cheveingne, A. D. & Kawahara, H. 2002. YIN, a fundamental frequency estimator for speech and music. *Journal of Acoustical Society of America,* 111, 1917-1930.

Chordia, P. 2004. Automatic rag classification using spectrally derived tone profiles. In proceedings of the *2004 International Computer Music Conference* (ICMC), University of Florida, Miami.

——. & Rae, A. 2007a. Automatic Raag recognition using pitch-class and pitch-class dyad distributions. In proceedings of the *7th International Conference on Music Information Retrieval* (pp. 431-436), Vienna, Austria.

——. & Rae, A. 2007b. Modeling and visualizing tonality in North Indian classical music. In *Neural Information Processing Systems, Music Brain Workshop,* Vancouver, Canada.

Choudhury, M. & Ray, P. R. 2003. Measuring similarities across musical compositions: an approach based on the raga paradigm. In proceedings of *International Workshop in Frontiers of Research in Speech Music* (pp. 25-34), Kanpur, India.

Cohn, R. 1998. Introduction to neo-riemannian theory: A survey and historical perspective. *Journal of Music Theory,* 42, 167-180.

Daniélou, A. 1968. *The rāga-s of Northern Indian Music*. London: Barrie and Rockliff, The Cresset Press.

Das, D. & Choudhury, M. 2005. Finite state models for generation of Hindustani classical music. In proceedings of *Frontiers Research in Speech and Music,* Bhubaneshwar, India.

Datta, A. K., Sengupta, R., Dey, N. & Nag, D. 2001. Studies on identification of raga using short pieces of taan: A signal processing approach. In proceedings of the *6th International Workshop on Recent Trends in Speech, Music, and Allied Signal Processing*. New Delhi.

Gold, B. & Morgan, N. 2004. *Speech and Audio Signal Processing: Processing and Perception of Speech and Music*. Singapore: John Wiley & Sons.

Hariram, S., Sheila, C., Krishna, S., Varadarajan, A. & Venkateswaran, N. 2001. Real time adaptive note and swaras recognition using HMM and neural network. In proceedings of *4th International EUROSIM Congress* (pp. 1-7), Delft, Netherlands.

Jairazbhoy, N. A. 1971. *The rāgs of North Indian music: Their Structure and Evolution*. London: Faber and Faber.

Janata, P., Birk, J. L., Van Horn, J. D., Leman, M., Tillman, B. & Bharucha, J. J. 2002. The cortical topography of tonal structures underlying Western music. *Science,* 298, 2167-2170.

Kessler, E. J., Hansen, C. & Shepard, R. N. 1984. Tonal schemata in the perception of music in Bali and the West. *Music Perception,* 2, 131-165.

Krishnaswamy, A. 2003. Application of pitch tracking to south Indian classical music. In *IEEE International Conference on Acoustics, Speech, and Signal Processing* (pp. 557-560), Hong Kong, China.

——. 2004. Results in music cognition and perception and their application to Indian classical music. In proceedings of *Frontiers Research in Speech and Music* (FRSM). Chidambaram, India.

Krumhansl, C. L. 1990. *Cognitive Foundations of Musical Pitch*. Oxford: Oxford University Press.

——., Toivanen, P., Eerola, T., Toiviainen, P., Jarvinen, T. & Louhivuori, J. 2000. Cross-cultural music cognition: Cognitive methodology applied to North Sami yoiks. *Cognition,* 76, 13-58.

Lu, G. & Hankinson, T. 2000. An investigation of automatic audio classification and segmentation. In proceedings of *International Conference on Signal Processing* (pp. 776-781).

Meyer, L. B. 1957. Meaning in music and information theory. *The Journal of Aesthetics and Art Criticism*, 15/4, 412-424.

Oram, N. & Cuddy, L. L. 1995. Responsiveness of Western adults to pitch-distributional information in melodic sequences. *Psychological Research,* 57, 103-118.

Pandey, G., Mishra, C. & Ipe, P. 2003. TANSEN: A system for automatic raga identification. In proceedings of *1st Indian International Conference of Artificial Intelligence* (pp. 1350-1363). Hyderabad, India.

Rabiner, L. R. 1989. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77/2, 257-285.

Rossignol, S., Desain, P. & Honing, H. 2001. State-of-the-art in fundamental frequency tracking. In proceedings of the *Workshop on Current Research Directions in Computer Music* (pp. 244-254), UPF, Barcelona.

Reck, D. B. 2005. India/South India. In J. T. Tinton & L. Fujie (Eds.), *Worlds of Music: An Introduction to the Music of the World's Peoples* (pp. 196-231). Belmont, CA: Thomson Learning Inc.

Saffran, J. R., Aslin, R. N., & Newport, E. L. 1996. Statistical learning by 8-month-old infants. *Science,* 274, 1926-1928.

——., Johnson, E. K., Aslin, R. N. & Newport, E. L. 1999. Statistical learning of tone sequences by human infants and adults. *Cognition,* 70, 27-52.

Sahasrabuddhe, H. V. 1992. Analysis and synthesis of Hindustani classical music. Available online: <http://www.it.iitb.ac.in/~hvs/paper_1992.html>.

——. 1994. Searching for a common language of ragas. In proceedings of *Indian Music Computation. Can 'Mindware' and software meet?* Available online: <http://www.it.iitb.ac.in/~hvs/paper_1994.html>.

Sengupta, R., Dey, N., Nag, D. & Datta, A. K. 2002. Automatic extraction of swaras and srutis from kheyal renditions. *Journal of Acoustical Society of India,* 30.

Shao, X., Xu, C. & Kankanhalli, M. S. 2004. Unsupervised classification of music genre using hidden Markov model. In proceedings of *IEEE International Conference on Multimedia and Expo*, Taipei.

Sinith, M. S. & Rajeev, K. 2007. Pattern recognition in South Indian classical music using a hybrid of HMM and DTW. In *International Conference of Computational Intelligence and Multimedia Applications* (pp. 339-343).

Sridhar, R. & Geetha, T.V. 2006. Swara identification for south Indian classical music. In proceedings of the *9th International Conference of Information Technology* (ICIT'06).

Srivastav, H. 1991. *Raga Parichay, Part – One*. Allahabad: Sangeet Sadan.

Vijayakrishnan, K. G. 2007. *The Grammar of Carnatic Music (Phonetics and Phonology)*. Berlin: Mouton de Gruyter.

Warren, J. 2008. How does the brain process music? *Clinical Medicine,* 8, 32-36.

Yardi, S. & Chew, E. 2004. Giving ragas the time of day: Linking structure, emotion, and performance time in North Indian classical music using harmonic network. In proceedings of *8th International Conference of Music Perception and Cognition* (pp. 247-288), Evanston, IL.

ARINDAM BHATTACHARJEE
CENTRE OF BEHAVIOURAL AND COGNITIVE SCIENCES,
UNIVERSITY OF ALLAHABAD, ALLAHABAD 211002, INDIA.

&

NARAYANAN SRINIVASAN (Corresponding Author)
CENTRE OF BEHAVIOURAL AND COGNITIVE SCIENCES,
UNIVERSITY OF ALLAHABAD, ALLAHABAD 211002, INDIA.
PH: 91-532-2460738
FAX: 91-532-2460738
E-MAIL: <NSRINI@CBCS.AC.IN>