Object Oriented Classification and Pattern Recognition of Indian Classical Ragas

Sayanti Chakraborty
Department of CSE
West Bengal University of Technology,
BF-142, Sector 1, Salt Lake City, Kolkata 700064,
West Bengal, India
sayanti.chakraborty21@gmail.com

Debashis De
Department of CSE
West Bengal University of Technology,
BF-142, Sector 1, Salt Lake City, Kolkata 700064,
West Bengal, India
dr.debashis.de@gmail.com

Abstract: **Musical pattern recognition is a challenging area of research as music is a combination of audio and speech; a signal of a particular music cannot be expressed through a linear expression only. For this reason musical pattern identification by mathematical expressions is critical. In this paper a new method is being proposed for cataloguing different melodious audio stream into some specific featured classes based on object oriented modeling. The classes should have some unique features to be specified and characterized; depending upon those properties of a particular piece of music it can be classified into different classes and subclasses. The concept is developed considering the non-trivial categorization problem due to vastness of Indian Classical Music (ICM). It can be expected that the novel concept can reduce the complication to the objective analysis of the classification problem of ICM as well as the world music.**

Keywords: Musical Patteren Recognition; Classification of audio data; Object oriented analysis; Musicology; Indian Classical Music.

I. INTRODUCTION

In world music there is a pressing need to recognize and categorize a musical sequence with the help of pattern recognition and feature extraction. In western music the music is totally cord based and there are some specified positions or calculations to build up a specific cord. In Indian Music it is very doubtful to specify any cord or any scale at a particular position. Scientific investigations in Indian music have fallen far behind those in western music. The present technology can provide ample scope for quantitative as well as qualitative approach to investigate shrutis, swaras, intervals, octaves (saptak), consonance (vaditya), musical quality, rhythm etc. [1-2] The power available with the computing technology and the advancement in the area of artificial intelligence, approximate reasoning and natural language processing may allow us to understand the higher level tasks like composition, improvisation and aesthetics associated with classical music [3-5]. Musical scales in India pose a confusing picture. A Raga is a melodic abstraction around which almost all Indian classical music is prepared. A raga is most easily explained as a collection of melodic gestures and it is a technique for developing them. The gestures are sequences of notes that are often inflected with various micro-pitch alterations and articulated with an expressive sense of timing. Longer phrases are built by joining these melodic atoms together [17].

Object Oriented Programming is a technique where a group of objects can be classified into some different classes on the basis of some unique characteristics. All the objects of the similar characteristic are put in the same class and the objects of different characteristics are classified into other different types of classes depending on their characteristics. All the components that are to be compared with others based on some properties are the Objects and the sections formed depending upon some properties or characteristics are the Classes. Every object should belong to a class. One object cannot have different properties other than its parent class. Depending upon one or more properties it can be put into some different classes. When a class is formed depending on more than one property then it can be sub-grouped into a subclass. This subclass should have the properties of its parent class and it can also have some properties beyond the scope of its higher class. This phenomenon is called the Inheritance.

As the complication of categorization of ICM is being encountered the specifications and features of the musical pieces are taken into consideration based on Indian musical conception. In Indian Music the classes are stated as Thhats and subclasses are different Ragas. All the ornamentations of different Ragas are the objects of the raga class. In Indian Classical Music there are enormous numbers of ragas belonging to different thhats. These Thhats are composed of seven primary notes. Among these Thhats there are clear differentiations of different ragas depending upon the occurrence of notes, consonance of notes and shrutis, and also use of notes. So we have opted the method to compare the concept of different Thhats with the concept of different Classes and Inheritance. The ragas are being treated as subclasses and the Thhat from which it is belonging to is being considered as the parent class and the properties based on which one raga is belonging to particular class is being mentioned as inherited properties.

In Indian Classical Music there are certain properties on the basis of which the Ragas are defined. The Properties are Aroha, Aboroha, Jati, Time of Singing, Pakarh, Register,

978-1-4577-0697-4/12/$26.00 ©2012 IEEE

Position of notes, Badi Notes, Sambadi Notes. Based on these properties we can use the object oriented approach to interpret the Indian Classical Ragas. We have observed that most of the characteristics are dependent on the notes. So, Note Identification is the most important and the fundamental objective. If twelve notes and the twenty two shrutis are recognized then it will be possible to identify a raga with the help of the notes. To identify the notes error-minimization algorithm is used.

II. RELATED WORK

A. K. Datta, R. Sengupta and N. Dey have merged two concepts of theoretical model based on vaditya of Indian Classical Music with the shruti position and the representation of Pitch ratios based on different opinions. In a paper of A K Datta, R Sengupta, N Dey, B M Banerjee and D. Nag have discussed about the pitch differences rather differences between a same scale in different voices that is caused by the shruti position between the notes used in Indian Classical Music [1]. Further in other papers [2-4] the concept of consonance of Swaras are elaborated and established. In these papers the presence of more than one fundamental frequency in the steady state of a particular frequency is noticed and analyzed. From the analysis the authors have reached to the decision that the consonance of a Swara or a note is very much relevant to the Indian Classical Music. An object oriented approach to programming of Western music analysis is described by T. S. Lande and Arvid O. Vollsnes. Different kinds of structural analysis are shown in their work, ranging from simple melodic analysis, through harmonic analysis to cluster analysis used for analysis of modern music [5]. Music Notation as Objects: An Object-Oriented Analysis of the Common Western Music Notation System is a research work of Kai Lass folk. Again H. G. Kaper, S. Tipie and Jeff M. Wright have also worked on Music composition and sound processing in an object oriented way. They have tried to produce the musical composition and the musical instruments in a same logic, object and procedure. They have introduced a system DISCO for music composition and sound design. Parag Chordia of Georgia Institute of Technology, Department of Music describes the results of the first large-scale raga recognition experiment. Ragas are the central structure of Indian classical music, each consisting of a unique set of complex melodic gestures. They constructed a system to recognize raags based on pitch-class distributions (PCDs) and pitch-class dyad distributions (PCDDs) calculated directly from the audio signal [6]. A large, diverse database consisting of several 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%. Their work demonstrates the effectiveness of PCDs and PCDDs in discriminating raags, eve n when musical differences are subtle [7]. Efforts to have computers mimic human behaviour can only succeed if the computer is using the same types of error criteria as a human. In these researches the ability of a computer to recognize patterns in music the way a human being would are explored. The result of this work is a system for the machine recognition of musical patterns [8-10] is explored. The problem of recognition of musical patterns as a classical pattern recognition problem. The main difference, though, between proposed approach and traditional approaches is that the error criterion used to judge the goodness of a match between a target pattern and a scanned pattern is derived from sophisticated studies of human perception of music that have been carried out in the field of psychology [11-14]. Gaurav Pandey, Chaitanya Mishra, and Paul Ipe have proposed Tansen [15], a system for automatic raga identification using HMM Models. Incorporating research in music perception and cognition, the music recognition system becomes like a human being, using what is known about how humans perceive, memorize, and reproduce music patterns. When a human being attempts to reproduce a memorized piece of music he or she is likely to introduce errors. The error criterion developed takes into account both rhythm and pitch information. The pitch error consists of an objective pitch error and a perceptual error. The latter incorporates algorithms for determining the localized tonal context in a piece of music [19-20]. The key-finding algorithm in [23] provides robust estimates of the tonal context during modulations and when there are deviations of very short duration from the prevailing local tonal context. The system also provides the means for determining the relative weight given to the objective and perceptual pitch [24] errors. This is based on the confidence of the key-finding algorithm. The weight given to the rhythm error is determined by a measure of the rhythm complexity. The algorithm will find the desired piece of music despite typical errors made during the memorization or reproduction phase. This new error criterion could be used in legal cases to determine the perceptual similarity of two pieces of music. Besides these identification and pattern recognition techniques there are also some developments in the formation of artificial intelligence based knowledge sharing atmosphere for the ICM [25]. Where the authors have used the machine training technique with the help of Hidden Markov Models and merged the pattern recognition concept to identify the Indian Ragas. In this paper our proposed work is an innovative addition to the aforesaid research works to explore in a new way of identification and pattern matching procedure.

III. Features Determination

A. *Selection of features*

Indian Music has some predefined characteristics to classify the ragas into thhats. A musical composition has basic three parts- melodious part, lyrical part and rhythm section. But to classify a composition, lyrics and rhythm cannot be the feature to determine its thhat. Therefore we have to depend on the melodious part. For this purpose the concept of notes comes into consideration. The representation of notes with there ratio list is described into the table 1. Some features related to notes can be considered for determination of the class of the ragas. They are Aroha, Aboroha, Jati, Badi Swara, and Sambadi Swara**.** The dependence upon the features is described into the Fig. 1.

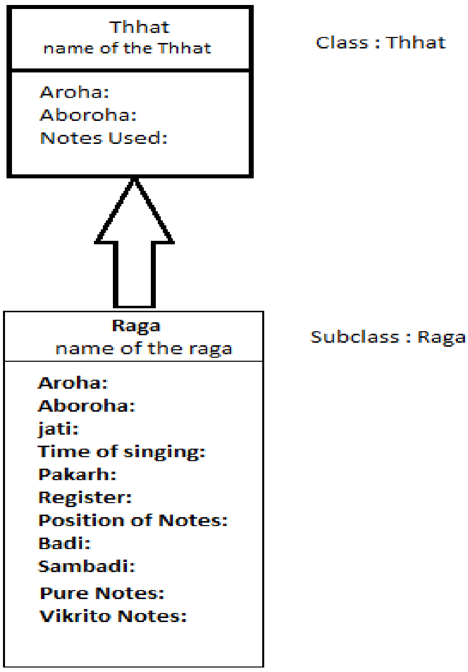

Fig. 1: Class and Subclass representation of Thhat and Raga

B. *Definition of different features*

Definition 1:

Thhat: Different Distributions of notes making different note structures are called Thhats.
A thhat is a musical scale, conceived of as a Western musical scale can have, with the seven notes presented in their order of ascent (aroha).
These thhats are dependent upon the aroho of the raga. In Indian Classical Music there are 10 thhats from each of which many ragas are created. The names of these ten thhats are - Kalyan, Bhairav, Kafi, Asavari, Bilabal, Khamaj, Bhairavi, Purbi and Torhi.

The formations of the thhats are described below:

Kalyan – Sa Re Ga MA Pa Dha Ni SA
Bhairav – Sa re Ga ma Pa dha Ni SA
Kafi – Sa Re ga ma Pa Dha ni SA
Asavari – Sa Re ga ma Pa dha ni SA
Bilabal – Sa Re Ga ma Pa Dha Ni SA
Khamaj – Sa Re Ga ma Pa Dha ni SA
Bhairavi – Sa re ga ma Pa dha ni SA
Purbi – Sa re Ga MA Pa dha Ni SA
Torhi – Sa re ga MA Pa dha Ni SA
Marwa- Sa re Ga MA Pa Dha Ni SA
For instance, Asavari is presented, and notated, as Sa Re ga (flat or komal) ma Pa dha (flat) ni (flat) in ascent, or aroha.

Definition 2:

Aroho & Aboroha**:** The sequence of notes of a particular raga or thhat in ascending order of the frequencies starting from the tonic of the scale of performance is called the Aroha and the sequence of notes starting from the double frequency of the tonic of the scale to the tonic of the scale in descending order of frequencies is called the Aboroha. By these two properties a Raga can be decided to belong to a Thhat. As example, In Raga Multaani –
Aroho – Ni Sa, ga MA Pa, Ni SA.
Aboroha – SA, Ni dha Pa, MA ga, re Sa.
From these two note streams it is clear that the Raga Multaani is composed of the notes used in the Thhat Torhi. So. It can be decided that the Multaani is belonging to the Thhat Torhi.

Definition 3:

Aalap: Aalap of a raga is a rendition of the raga in which part the possible legal combinations of the used notes are performed without any fixed rhythm. Here in the beginning portion the performer starts from the tonic and reaches to higher to higher frequencies according to ability and expertise and then comes downward to reach to the tonic gradually. For example in case of the raga Multaani the Aalap is like:
Ni Sa ga MA Pa, MA ga, MA ga re Sa. Ni Sa ga re Sa, Ni Sa MA ga MA Pa, dha Pa MA Pa MA ga, MA ga re Sa. ga MA Pa Ni SA, MA Pa Ni SA, Ni dha Pa, MA Pa dha Pa, MA Pa MA ga, MA ga re Sa.

Definition 4:

Shrutis**:** The linking between notes is called the Shrutis. To get more pleasant and melodious music the Shrutis are used by the expert musicians. Shrutis are interlinked and merged with the notes. In between the transition from one note to another note these shrutis are present.

The Swaras or notes used in Indian Music and Western Music are:
Sa (Sadaj) = Do, Re (Rishava) = Re, Ga (Gandhara) = Mi, ma (Madhyama) = Fa,
Pa (Panchama) = So, Dha (Dhaivata) = La, Ni (Nishad) = Ti, SA (Sadaj) = Do.

SA has the doubled frequency of Sa. For example if Sa is of 240Hz then SA should be of 480Hz.
The Swaras or notes explicitly and uniquely used in Indian Music are- re (Komal Rishava), ga (Komal Gandhara), MA (Tivra Madhyama), dha (Komal dhaivata) and ni (Komal Nishada). They are called as Vikrita Swaras. The underlined notes indicate the notes in lower octave from the tonic of the scale of performance. Lastly this is to specify that tonic of a scale means the starting note of a scale. The variety of combinations can make a difference in the creation of emotion by the note sequences. Even different ragas created from a single thhat can create different emotions.

Sa	Re	re	Ga	Ga	ma	MA	Pa	dha	Dha	ni	Ni
1	1.054	1.125	1.186	1.253	1.333	1.415	1.5	1.158	1.691	1.777	1.893

TABLE 1: REPRESENTATION OF DIFFERENT NOTES WITH THE RATIOS

On the other hand, in case of the any traditional and contemporary music, the songs are also the composition of notes in different manners of melodious combination. The way of combination of notes plays the main role to create the emotion in contemporary songs as in the Indian Classical Music.

IV. METHOD OF ANALYSIS

A. *Pitch period extraction from signal*

A method based on Phase-Space Analysis (PSA) is used for extracting pitch periods. PSA [3] uses the basic fact that, for a repetitive function, two points of the signal having a phase difference of 2π have the same displacement and, therefore, if plotted in a 2-D phase diagram would be a straight line with a slope of $\pi/4$. For a quasi-periodic signal this diagram would be very close to this straight line. For all other phase angles the curves would be wider loops. In a phase-space diagram the points representing such pairs would be lying on a straight line with a slope of $\pi/4$ to the axis. In case of a quasi-periodic signal, viz., speech or music these points would lie in close very flattened loop with the same axis. As the phase between two such points increases from 0 to 2π the corresponding phase-space diagram broadens from this flat loop and forms close broader loops. It can also be shown that the root mean square value of the deviation is same as the absolute value of the difference between displacements of the two points. For quasi-periodic signals the minima in the sum of deviations occurring at a phase difference of $2n\pi$, where n is a small integer, shows comparable small minima values, usually a little larger than that for n=1. A study of a large volume of signal reveals that (a) the sum of deviations last at phase differences other than $2n\pi$, are always large and never fall below minimum deviations plus 0.2(max – min) of deviations and (b) that deviations for phase differences $2n\pi$, $n\neq1$, never falls below that for 2π when this value exceeds the aforesaid threshold. This allows a clear threshold and algorithmic logic for pitch extraction. Henceforth the pitch pattern files, extracted using the above method, will be referred to as 'cep file'.

B. *Smoothing*

The .Cep file contains pitch extracted only in the quasi-periodic region of the signal. The PD algorithm above uses a predefined specific range for possible pitch values e.g., in the present case the default range is fixed between 70 Hz to 700 Hz. Three types of error were observed. One is that the determined pitch value is approximately equal to half or double of the actual pitch value. This error is typical for PSA and is usually a rare and isolate occurrence. The other is pitch occurring out side the range. Third one is spike in pitch sequence. The smoothing operation for the first type of error is simply detection of it and doubling or halving the value as required. For the two other types of errors are detected through an examination of local pitch values and the erroneous values are replaced using linear interpolation from the neighboring valid data.

C. *Stady state detection*

Even in a perfectly perceptible steady note all the pitch values are never identical. There are always some involuntary small variations in pitch, which has no relevance in the perception of the corresponding value of the note. The definition of steady state is not a trivial problem. A steady state is defined in the pitch files as that subsequence of pitch data where all the pitch values lie within a predetermined band around the mean value. Furthermore as the signal considered in the present study is taken from the total singing the steady states are restricted to minimum duration of 60 milliseconds.

D. *Tonic identification*

An error minimizing technique has been proposed for the purpose. The basic approach is to try out a large number of pitch values in a given range as a possible tonic and to find the deviations of the actual steady states from the closest note predicted by the selected tonic. It is reasonable to assume that for an incorrect trial tonic the sum of these deviations would be large. In fact, the sum of these deviations for each trial tonic is a function of the value of the trial tonic. It is assumed that the minimum in this function will indicate the actual tonic. In Fig. 2 the pitch values with their corresponding errors that are occurred are plotted.

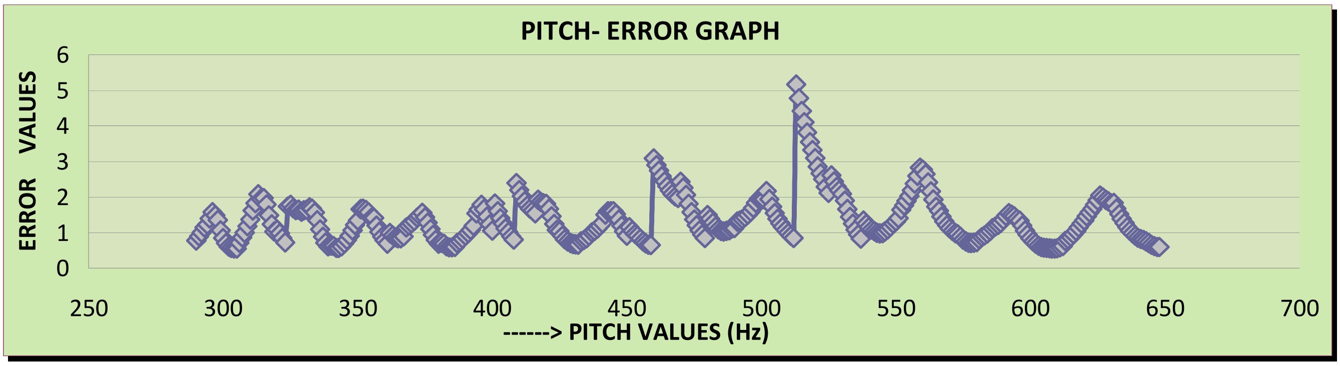

Fig.2. Plotting of Pitch values with Errors

V. Experimental Details And Result Discussion

In this experiment the Wave Surfer software is used to get the pitch file of a song. At first the .wav file is opened in the software. The file should be mono, and sample encoding will be set to line 16 and sample rate is at 22050. It is opened in a waveform structure. From the Structure a pitch contour of the song is generated. In this circumstance the pitch form will give all the pitches used in the song, and the pitch data are to be saved in an excel sheet. The pitch data are to be sorted in ascending order but there occur some 0 values that are to be deleted. This .f0 file i.e. the file obtained from saving the data file of the .wav file is passed through the error minimization programming. It gives the pitches with their corresponding error. In table 2 the detected tonics, notes and the total error of identification is listed down. By applying the algorithm the identification of the tonic is performed with up to 94% accuracy. The described procedure can be followed by an object oriented approach to Indian Classical Music.

VI. Conclusions

In this paper an innovative method is proposed for classification of different melodious audio stream into some specific featured classes based on object oriented modeling. The Indian Music has a subjective approach i. e. Guru-Shishya Parampara i.e. Teacher - Student Interaction, in which there is no defined structure to follow without the training of the Guru or the teacher. From the proposed procedure overall 94 % accuracy to identify the tonic notes and other notes can be obtained. For this purpose an object oriented representation of the Indian Classical Music is essential. This object oriented methodology for ICM will lead to develop the inheritance and polymorphism model for musical pattern recognition and pattern analysis.

Table 2: Identified Notes with corresponding error in identification

SL. NO.	Name of Thhat	Tonic Identified at(Hz)	Identified Notes (with wrongly identified notes)	Accuracy Achieved (%)	Error Occurred (Ratio)
1	Kalyan	**272**	Sa Re Ga MA Pa Dha Ni	100	**1.686962**
2	Bhairav	**274**	Sa re Ga ma Pa dha Ni	100	**1.59536**
3	Kafi	**272**	Sa Re ga ma Pa **dha** Dha ni	85	**1.587548**
4	Asavari	**272**	Sa Re ga ma Pa dha ni **Ni**	85	**1.548972**
5	Bilabal	**276**	Sa Re **ga** ma Pa Dha Ni	85	**1.056914**
6	Khamaj	**282**	Sa Re Ga ma Pa Dha ni	100	**1.850798**
7	Bhairavi	**268**	Sa re ga ma Pa dha ni	100	**2.56441**

8	Purbi	**276**	Sa re Ga MA Pa dha Ni	100	**1.962379**
9	Torhi	**272**	Sa re ga **ma** MA Pa dha Ni	85	**1.831847**
10	Marwa	**274**	Sa re Ga MA Pa Dha Ni	100	**1.554882**

ACKNOWLEDGMENT

Authors are grateful to Department of Science and Technology (DST) for sanctioning a research Project under Fast Track Young Scientist scheme reference no.: SERC/ET-0213/2011 under which this paper has been completed.

REFERENCES

[1] A. K. Datta, et.al, "Pitch Analysis of Recorded Vocal Performances in Hindustani Music: Evidence of a Personal Scale", J. Acoust. Soc. Ind., XXV, VI-1.1, 1997.

[2] A. K. Datta, et.al, "Multiple States in a Note in Hindustani Music and their Relevance to Consonance", J. Acoust. Soc. Ind., XXVI, 394, 1998.

[3] R. Sengupta, "Study on some Aspects of the Singer's Formant in North Indian Classical Singing", J. Voice, 129, Raven Press, 1990.

[4] [4] A. K. Datta et.al, "Relevance of Consonance in Indian Musical Scale: Theory and Practice", J.Acoust. Soc. Ind., XXIII, 110, 1995.

[5] T.S.Lande, A.O. Vollsnes, "Object Oriented Music Analysis", Computers and the Humanities 28: pp. 253-257, 1995.

[6] Parag Chordia. Automatic raga classification using spectrally derived tone profiles. In Proceedings of the International Computer Music Conference, 2004.

[7] Parag Chordia. Automatic raag classification of pitchtracked performances using pitch-class and pitch-class dyad distributions. In Proceedings of International Computer Music Conference, 2006.

[8] Ching-Hua Chuan and Elaine Chew. Audio keyfinding using the spiral array ceg algorithm. In Proceedings of International Conference on Multimedia and Expo, 2005.

[9] Patricio de la Cuadra, Aaron Master, and Craig Sapp. Efficient pitch detection techniques for interactive music. In Proceedings of the International Computer Music Conference, pp. 403–406, 2001.

[10] C. Duxbury, J. P. Bello, M. Davies, and M. Sandler. A combined phase and amplitude based approach to onset detection for audio segmentation. In Proc. of the 4th European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), pp. 275–280, 2003.

[11] E. Gomez and P. Herrera. Estimating the tonality of polyphonic audio files: Cognitive versus machine learning modelling strategies. In Proceedings of International Conference on Music Information Retrieval, 2004.

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

[13] C. Krumhansl and R. Shepard. Quantification of the hierarchy of tonal functions within a diatonic context. Journal of Experimental Psychology: Human Perception and Performance, 5(4), pp. 579–594, 1979.

[14] Carol Krumhansl. Cognitive Foundations of Musical Pitch. Oxford University Press, 1990.

[15] 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.

[16] Sun, X.: A pitch determination algorithm based on subharmonic-to-harmonic ratio. In: Proc. of International Conference of Speech and Language Processing , 2000.

[17] Duxbury, C., Bello, J.P., Davies, M., Sandler, M.: A combined phase and amplitude based approach to onset detection for audio segmentation. In: Proc. of the 4th European Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), pp. 275–280, 2003.

[18] Kameoka, A., Kuriyagawa, M.: Consonance theory, part I: Consonance of dyads. Journal of the Acoustical Society of America 45(6), pp. 1451–1459, 1969.

[19] I. Shmulevich, E. J. Coyle, "Establishing the Tonal Context for Musical Pattern Recognition," Proceedings of the 1997 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 1997.

[20] I. Shmulevich, E. J. Coyle, "The Use of Recursive Median Filters for Establishing the Tonal Context in Music," Proceedings of the 1997 IEEE Workshop on Nonlinear Signal and Image Processing, 1997.

[21] E. J. Coyle, I. Shmulevich, "A System for Machine Recognition of Music Patterns," Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Seattle, WA, 1998.

[22] I. Shmulevich, D. Povel, "Rhythm Complexity Measures for Music Pattern Recognition," Proceedings of IEEE Workshop on Multimedia Signal Processing, 1998.

[23] O. Yli-Harja, I. Shmulevich, K. Lemstrm, "Graph-based Smoothing of Class Data with Applications in Musical Key Finding," Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, 1999.

[24] I. Shmulevich, O. Yli-Harja, E. J. Coyle, D. Povel, K. Lemstrm, "Perceptual Issues in Music Pattern Recognition Complexity of Rhythm and Key Finding," Computers and the Humanities, Kluwer Academic Publishers, 2001.

[25] S.Chakraborty, D.De, K. Roy, "A Knowledge Sharing Virtual Community for identification of different Indian Classical Ragas", NCICT, Allied Publishers, pp. 81-84, 2011.