Analytical Approach on Indian Classical *Raga* Measures by Feature Extraction with EM and Naive Bayes

Akhilesh K Sharma
Computer Science & Engineering
Sir Padampat Singhania University,
Udaipur,(Rajasthan), India ,

Avinash Panwar
Computer Science & Engineering
Sir Padampat Singhania University,
Udaipur,(Rajasthan), India

Prasun Chakrabarti
Computer Science & Engineering
Sir Padampat Singhania University,
Udaipur,(Rajasthan), India

ABSTRACT
Music analysis is the main task in the musical information retrieval (MIR) systems. In this paper an analytical study based on these MIR techniques has been carried out to perform analysis of the Indian classical music and Indian *ragas*. The *ragas* are further classified into various *thaats* and their pitch class profiles and statistical measures. This paper demonstrates the strategy by which the various *raga* can be categorized using these statistical measures. The choices of algorithm used are the EM algorithm and the Naive bayes algorithm.

Indian classical music is very popular because of the musical styles and the emotions it can reveal. Thus MIR (musical information retrieval) and its musical analysis is a very good choice for the researchers who have both knowledge of music and computer background.

This paper includes the Matlab programming environment and toolbox for the effective result simulations. The EM and naive bayes algorithm have been utilized and the open source platform has been used for the rest of the work.

Keywords
EM algorithm, naive bayes, Indian classical music, music information retrieval, classification, clustering.

1. INTRODUCTION
In the recent past a lot of work has been done on musical analysis and specially the Indian classical music, generating lot of new insight into this domain. The research related to musical information retrieval is thus attracting the interest of so many researchers.

The history of Indian classical music is very rich. Indian classical music is having two main streams of north Indian and south Indian based music and styles. The music is categorized in different *thaats* based on which the *ragas* are derived. *Ragas* are specific to the time of the day and generate a specific sentiment only when played at the right time.

In this paper an analytical approach was used for estimating the *raga* which was based on feature extraction using Expectation maximization algorithm and the naïve bayes algorithm. It utilized the frequency component, low level features like spectral roll off points, zero cross validation etc and included the variation of the pitch frequency, timber parameters, low frequency components, high frequency components etc[2] for estimation of *Ragas.*

The *Weka* tool developed by *University of Waikato* and the *Jaudio* open source tool for the simulation [1] have been used in this study.

As a continuation to this study the clustering and the classification schemes can be utilized to separate the key details and to arrange them in some orderly fashion [9].

2. RELATED WORK
The Latest research methods and techniques are focusing on carnetic *raga* and its analysis. The *ragas* are classified in two main categories first is the north Indian and the second is the south Indian classical music. The music research and its analysis plays an important role in finding the *raga* patterns on various ways.

To identify their variety the '*thaat*' categorization is available, V.N. bhatkhande proposed system that is very relevant with this type of categorization [1].

He proposed that *thaats* are classified in 10 different ways according to structure provided [8].

The 10 *thaats* are as follows:

Bilawal, kalian, Todi, Bhairavi, Marwa, Kafi, Bhairav, Khamaj, Purvi , Asavari.

These raags possess very different structural patterns so they can be distinguishingly identifiable [1].

Different *ragas* performed on different relevant timings for the largest impact and increased sentiments on humans mind.

The Indian classical music includes many *gharana* and the different style and tradition for those *gharana*. The V.N. bhatkhande[1] describes the culture of these *gharana* and their music forming methods. The *raga* expresses its full impressions only when it is rendered/ performed on specified time (fig1).

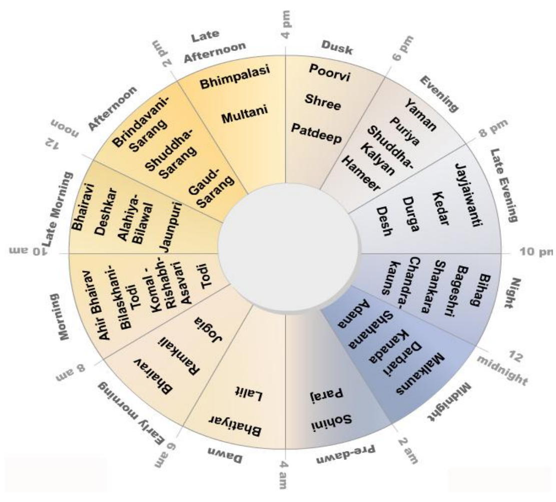


Figure 1. Samay for the *raga* to be performed[10][8].

The time regions are provided in the fig 1.

The Indian classical music divides the music *raga* according to its time when it needs to be rendered and it will be having a maximum effect on the listener [2].

In fig 2 the key representations are shown as on device so that the particular *raga* can also be framed.

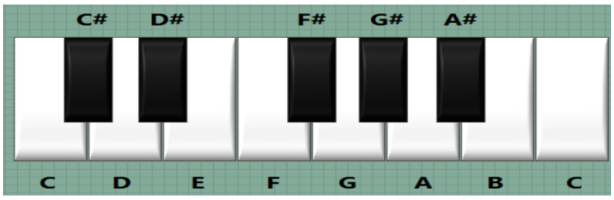


Figure 2. Key representation of ensemble device[6].

As per the combination and the rendering of keys, the *raga* may be obtained (as shown in the fig 3). These *raga* representations are to be utilised for the rendering while the artist is performing live.

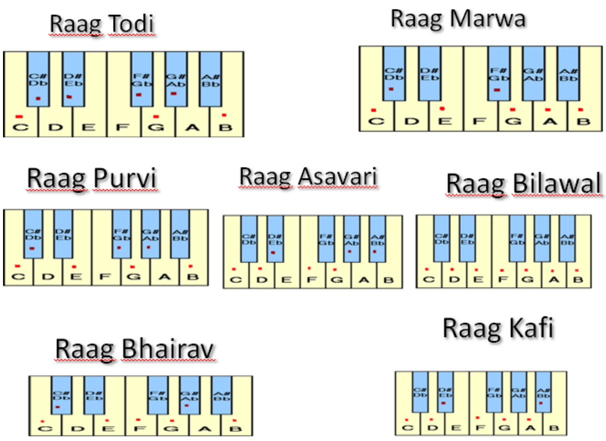


Figure 3. Different key formations of the various raag on ensemble device [9].

2.1 Characteristics of Raaga:

Jaati: Jaati *raga* express the no. of notes in the increasing and decreasing order.it is called aaroh (increasing order) and avaroh (decreasing order). These can be classified further as adava (5 ascending , 5 descending) , shadava (6 ascending ,6 descending), sampoorna (7 ascending ,7 descending) [3].

Vaadi and samvadi: Vaadi is the main note of the *raga*. The samvadi is the second most so popular note in the *raga*[2].

Chalan samay and rasa: Chalan is responsible to identify the octave. Samay is responsible for identifying the corresponding time at which the *raga* has its highest impact on the listener [3].

Rasa: It indicates the emotions attached with the particular *raga*. Every *raga* has the different emotions attached with it like veer rasa, shrangar rasa, bhakti rasa, hasya rasa etc[2].

Meend : It is also known as the glide from one note to the other note or on different octaves. It is performed by the vocalists while performing and by instrumentalists as well[2].

Andolan: It is called as a very gentle oscillations or swings that starts from fixed notes and touches the outer frequency or periphery of a different note. Thus it resembles different tones in between the note[9].

Frequency plays an important role in musical signal[10] analysis so the formula gives an important aspect of frequency as[7]:

$\Delta T = 1/\text{Frequency}$

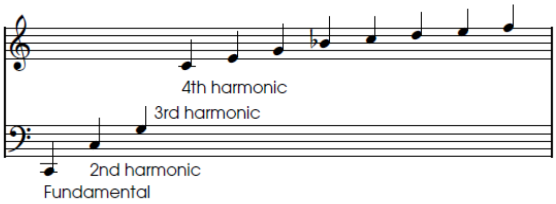


Figure 4. The musical overtone series[8].

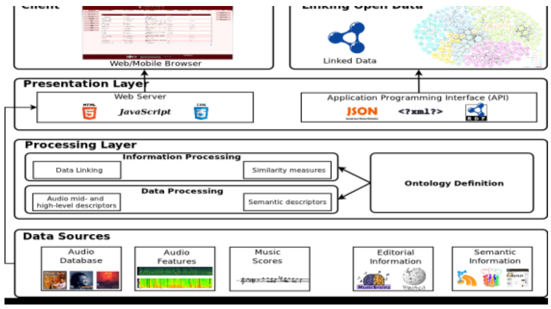


Figure 5. Musically aware system for music collections [4]

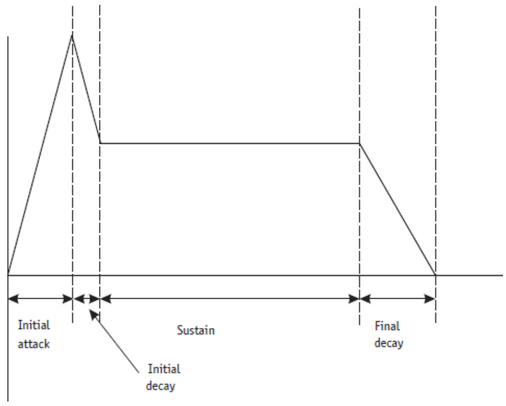


Figure 6. The envelope of an audio signal[10].

Re-arranging the above formula:

$\text{Frequency} = 1/\Delta T$

This means that if an instrument has an attack time of 1 millisecond, the equivalent frequency is 1 kilohertz (kHz)[9].

2.2 Musical Notation and Frequency

The skill is important to become a competent recording engineer, and the producer is strongly advised to pursue the study of music. To explain the concept note A has a frequency of 440Hz (this is the note occupying the second space of the treble-clef staff) [5]. The A that is located on the top line of the bass clef staff is an octave below 440Hz and has a frequency of 220Hz [10][11]. An octave relationship is a doubling or halving of frequency. Figure 7 shows a musical scale with the corresponding frequencies[11][21][22].

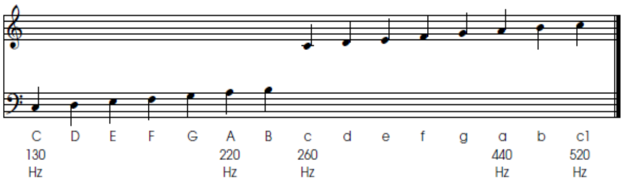


Figure 7. A musical scale with the corresponding frequencies [10][11].

3. THE PROPOSED STRATEGY: METHODS ADOPTED

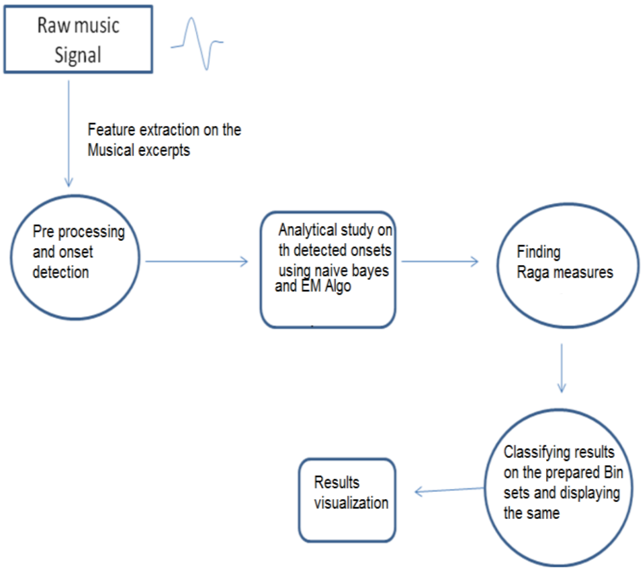


Figure 8. The proposed Strategy/Model

The proposed strategy successfully exploits the features of the musical excerpts and it enables comparison of the naïve bayes and the expectation maximization algorithm. The tools like *Matlab* and *Weka* have been used for the result formation.

4. CONDUCTED EXPERIMENT AND RESULTS

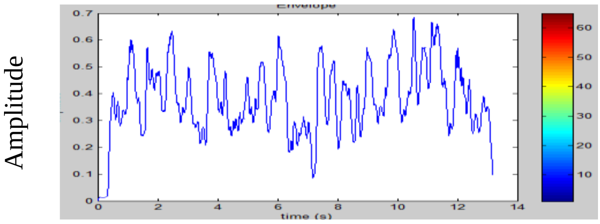


Figure 9 . Feature extraction preprocessing step and envelop detection of musical excerpt.

Thaat

KALYAN	1.9814	1.0185	1.0001
BILAWAL	1.997	1.0021	1.0009
BHAIRAVI	1.0001	1.0047	1.9952
TODI	1.0001	1.9978	1.0021
ASAVARI	1.0009	1.0001	1.999
KAFI	1.0303	1	1.9697
PURVI	1.0005	1.9994	1
MARWA	1.0321	1.9678	1.0001
BHAIRAV	1.0046	1.9933	1.0021
KHAMAJ	1.9695	1.0002	1.0303

[total]	13.0167	13.984	12.9994
C 0	1	1	1
1	4.0167	4.984	3.9994
[total]	5.0167	5.984	4.9994
Db 0	3.9791	1.0209	3
1	1.0375	4.9631	1.9994
[total]	5.0167	5.984	4.9994
D 0	1.0375	4.9631	1.9994
1	3.9791	1.0209	3
[total]	5.0167	5.984	4.9994
Eb 0	3.9852	3.9813	1.0334
1	1.0314	2.0026	3.966
[total]	5.0167	5.984	4.9994
E 0	1.0314	2.0026	3.966
1	3.9852	3.9813	1.0334
[total]	5.0167	5.984	4.9994
F 0	2.0142	3.9836	1.0023
1	3.0025	2.0004	3.9971
[total]	5.0167	5.984	4.9994
Gb 0	3.0025	2.0004	3.9971
1	2.0142	3.9836	1.0023
[total]	5.0167	5.984	4.9994
G 0	1	1	1
1	4.0167	4.984	3.9994
[total]	5.0167	5.984	4.9994
Ab 0	4.0104	1.9886	2.001
1	1.0063	3.9953	2.9984
[total]	5.0167	5.984	4.9994
A 0	1.0063	3.9953	2.9984
1	4.0104	1.9886	2.001
[total]	5.0167	5.984	4.9994
Bb 0	3.0158	4.979	1.0052
1	2.0008	1.005	3.9942
[total]	5.0167	5.984	4.9994
B 0	2.0008	1.005	3.9942
1	3.0158	4.979	1.0052
[total]	5.0167	5.984	4.9994

In this research paper the EM algorithm and the Naïve bayes algorithm are used to show the results. The algorithm are implemented in java and weka tool is used for showing the results based on the analysis performed on *raga* dataset.

=== Stratified cross-validation ===

=== Summary ===

Correctly Classified Instances	6	60	%
Incorrectly Classified Instances	4	40	%
Kappa statistic	0		
Mean absolute error	0.5273		
Root mean squared error	0.5347		
Relative absolute error	100	%	
Root relative squared error	100	%	
Total Number of Instances	10		

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0 0 0 0 0 0 0

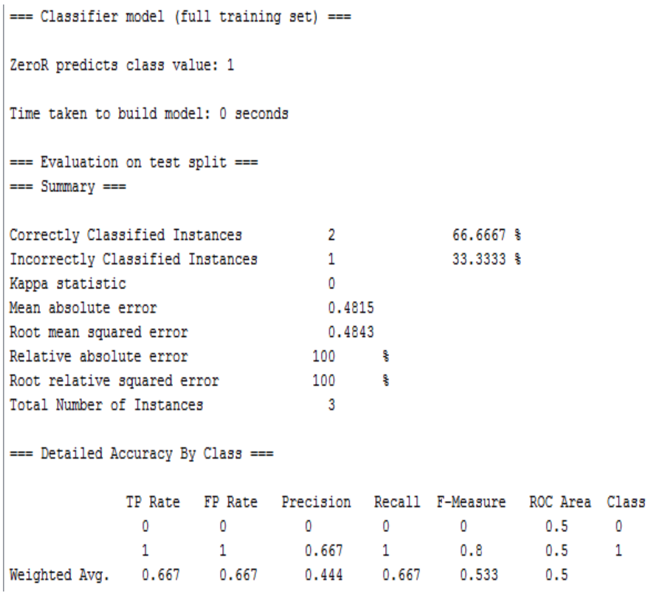
1 1 0.6 1 0.75 0 1

Weighted Avg. 0.6 0.6 0.36 0.6 0.45 0

Based on the above the confusion matrix obtained is as follows:

===Confusion Matrix =====

Bhairavi	Yaman	←classified as
43	74	| Bhairavi
35	87	| Yaman



```
=== Classifier model (full training set) ===

ZeroR predicts class value: 1

Time taken to build model: 0 seconds

=== Evaluation on test split ===
=== Summary ===

Correctly Classified Instances          2               66.6667 %
Incorrectly Classified Instances        1               33.3333 %
Kappa statistic                         0
Mean absolute error                     0.4815
Root mean squared error                 0.4843
Relative absolute error               100      %
Root relative squared error           100      %
Total Number of Instances               3

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0         0         0          0         0          0.5      0
                 1         1         0.667      1         0.8        0.5      1
Weighted Avg.    0.667     0.667     0.444      0.667     0.533      0.5
```

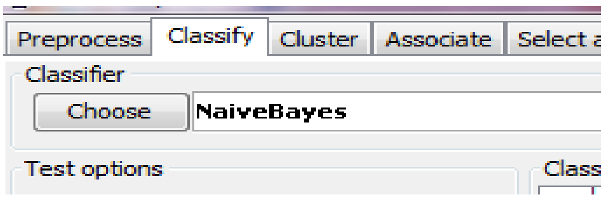
Figure 10: Results on classifier with accuracy measure.

The confusion matrix shows the classified segment. In this confusion matrix the musical data was supplied as input sample and the confusion matrix was plotted. The following fig shows the same:

===Confusion Matrix =====

Bhairavi	Yaman	←classified as	
0	1	| Bhairavi	=0
0	2	| Yaman	=1

The naïve bayes algorithm has been used since the parameters used in the *Ragas* are independent for different *ragas*.



Fjgure 11. Weka option for Naïve bayes algorithm tab.

Figure 11 shows the Feature extraction (onset detected bins) for a given musical excerpt in the form of onsets. The frequency estimation has been used to plot the particular *raga*.

Matlab and MIR toolbox have been used to perform onset detection and for extracting the relevant features of the musical excerpt.

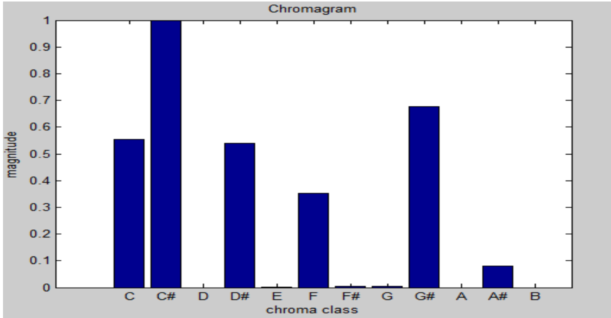


Figure 12. Frequency representation on detected onsets

While processing the musical excerpted portion the focus is on the bins and onset frequency of the music segment.

While detecting the onset, the N-bins are to be framed and mapped with the specified *raga* for which it has been created.

Naïve bayes used for this is like :

P(x/y)=P(y/x) P(x)/P(y)

P(x) : probability of event x before evidence is seen.

P(y) : Probability of event y after evidence is seen.

And thus for the classification and learning the probability of the class is given by an instance :

Evidence y=instance

Evidence x=class value for instance,

The naïve assumption derived from this is that evidence splits into parts also known as attributes that are independent,

P(x/y)=P(y1/x) P(y2/x) ….. P(yn/x) P(x)/P(y)

C	Db	D	Eb	E	F	F#	G	Ab	A	Bb	B
1	0	1	0	1	1	0	1	0	1	0	1
Thaat/Raga Bilawal		←		Evidence							
?											

The values shown as Boolean one are the belongings of the *raga* bilawal *thaat*, and their values may be used in order to use it in naive bayes.

P(raag/E)=P(c=1|bilawal)* P(D=1|bilawal)* P(E=1|bilawal)* P(F=1|bilawal)* P(G=1|bilawal)* P(A=1|bilawal)* P(B=1|bilawal)*P(bilawal)/P(E)

C	Db	D	Eb	E	F	F#	G	Ab	A	Bb	B
1	0	1	0	1	0	1	1	0	1	0	1
Thaat/Raga Yaman		←		Evidence							

P(raag/E)=P(c=1| Yaman)* P(D=1| Yaman)* P(E=1| Yaman)* P(F=1| Yaman)* P(G=1| Yaman)* P(A=1| Yaman)* P(B=1| Yaman)*P(Yaman)/P(E)

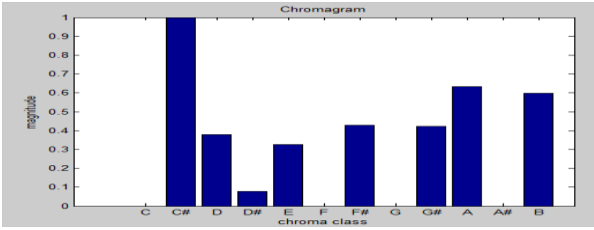


Figure 13. Basic sargam onset note detection in bhairavi

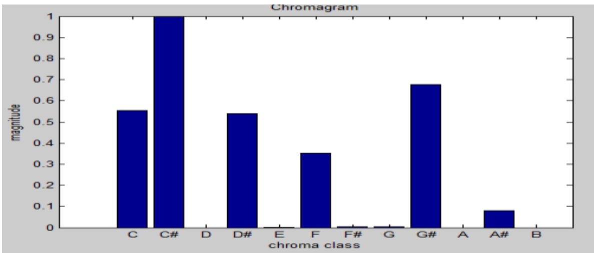


Figure 14. *Raga* hamsadhwani onset detection for notes.

Following table-1 shows the comparisons between the various models used in the experiment.

Table 1. Computation table on various Models.

Models	Normal Approach	Expectation-Maximization	Naïve bayes
Precision	0.8123	0.8043	0.8742
Accuracy	56%	70%	82%

Here in the above diagrams for the chromagram, the *raga* onsets are shown, and the specific structure that is followed is also shown into it. The 44.1 Khz frequency i.e. the standard frequency for the music CD quality is taken for the musical measurements.

While collecting bins for the total of 365 different *ragas*, the frequent onsets over these collections of the *raga* data set have also been identified.

Upon implementing the algorithms on these collected data sets it was observed that the accuracy of estimation of *ragas* was 82% which is a very promising outcome.

RESULTS

Results are shown as follows for the three approaches after the simulation of EM algorithm, Naïve bayes algorithm and the normal methods:

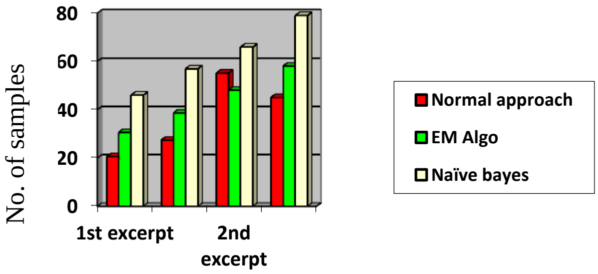


Figure 15 . Result for the excerpted portion verified by the users.

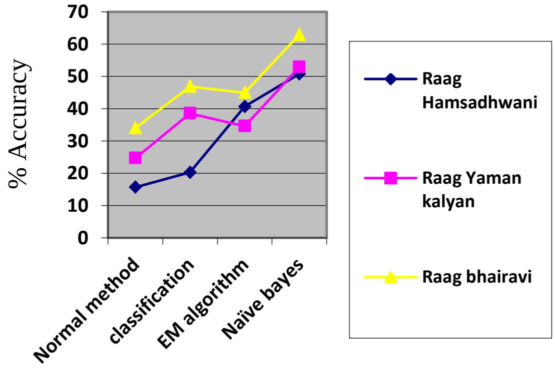


Figure 16 . Diagram for % Accuracy of supplied data.

The following Figure 17 shows the time and frequency domain for analyzing HPCP and the music segment. There is a need to convert the signal from time to frequency domain.

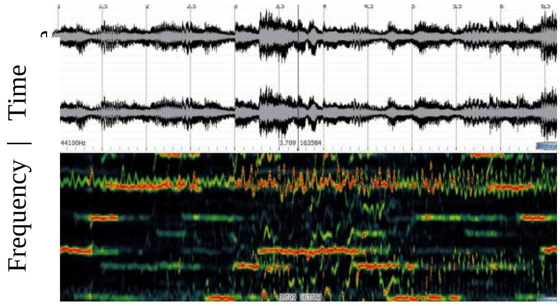
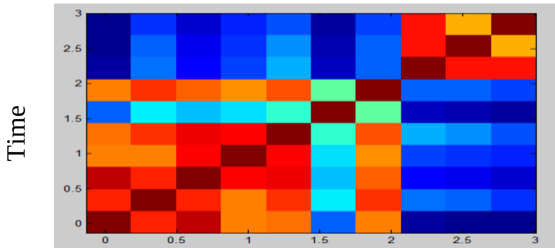


Figure 17 . Diagram of Musical HPCP profile calculation, In time and frequency domain.

The figure 18 depicts the chroma variations of the spectrogram in frequency domain, it is used to analyze the musical excerps.



x-Frequency axis (Using Chromagram)

Figure 19 . Result of segmented similarity matrix in spectrogram measure of the musical excerpt.

5. CONCLUSION AND FURTHER RESEARCH

It should be noticed that the *raga* are very much interrelated, and to classify them is a very challenging task. In this study along with various other algorithms the Naïve Bayes algorithm was used and it demonstrated high percentage of accuracy in predicting the *ragas*. In this paper, the visualization and identification of different aspect of raga characteristics has been obtained. The outcome is very optimized and promising. The future scope of this work includes the raga and their classification in different styles, *jatis* and based on the changing pitch, timbre of the artists while performing the *raga*. Further studies can be carried out for detection of beats for classification of various *taalas*.

As a further enhancement in this study, self organized map based approach can also be used in order to improve the accuracy and the segmentation of the various *ragas.*

6. ACKNOWLEDGMENT

The acknowledgement includes the suggestions from the reviewers and fellow researchers, and ideas incorporated to increase the quality of work. With deep appreciation, like to thank the advisory committee members, IR group fellows, and also management for providing MIR(Music Information Retrieval) lab for experiments.

7. REFERENCES

[1] Bhatkhande, V. (1934). Raga parichay. Hindusthani sangeet paddhati. Sangeet Karyalaya , (pp. 50-212).

[2] Khan, A., & Ruckert, G. (2004). The Classical Music of North India. New Dehli: Munshiram Manoharlal.

[3] Krishna, T. M., & Ishwar, V. (2012). Proc. of the 2nd CompMusic Workshop, Istanbul, Turkey,(pp. 12-18).

[4] Sordo, M., & Gulati, S. (2012). Proc. of the 2nd CompMusic Workshop,Istanbul, Turkey, (pp. 20-24).

[5] Hainsworth, Stephen, W., Malcom, D. M., & Patrick, J. W. (2001). Analysis of reassigned spectrograms for musical transcription. Applications of Signal Processing to Audio and Acoustics, (pp. 32-51).

[6] NI national instruments. (n.d.). Labview 7.1 . Retrieved 1 26, 2014

[7] Klapuri, A., & Davy M., (2006). Signal Processing Methods for Music Transcription. Springer , (pp. 122-131).

[8] (2014, Jan 26). Retrieved from www.swarganga.org: https://www.swarganga.org/articles/icmconcepts/icm16.php

[9] Chordia, P. & Rae, A., (2008). Understanding emotion in raag: an empirical study of listener responses. Springer-Verlag , (pp. 110-124).

[10] Pandey, G., Mishra, C., & Ipe, P. (2003). Tansen : A system for automatic raaga identification. Proceeding of the 1st Indian International Conference on Artificial Intelligence , (pp. 1350-1363).

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

[12] Sharma, A.K., Panwar A. (2014). "An efficient approach using LPFT for the karaoke formation of musical song," IEEE International Advance Computing Conference (IACC), vol., no., (pp.601-605).

[13] Sharma, A., and Lakhtaria, K.I.(2013). Data mining based predictions for employees skill enhancement using pro- skill-improvement program and performance using classifier scheme algorithm, (IJARCS), Volume 4, No. 3, (Special Issue), ISSN 0976 – 5697.

[14] Sharma, Akhilesh K, Panwar, Avinash, Lakhtaria, Kamaljit I; Vishwakarma, Santosh (2014). "An Analytical approach based on self organized maps (SOM) in Indian classical music raga clustering," Seventh International Conference on Contemporary Computing (IC3), vol., no., pp.449,453, 7-9 Aug. 2014.

[15] Santosh K. Vishwakarma, Kamaljit I Lakhtaria, Divya Bhatnagar, Akhilesh Sharma (2014). "An efficient approach for inverted index pruning based on document relevance", Conference Proceeding of Fourth International Conference on Communication Systems and Network Technologies, Page No. 487-490. DOI 10.1109/CSNT.2014.103.

[16] S. K. Vishwakarma, C. S. Jangid, K. I. Lakhtaria, (2014). "Ad-hoc Retrieval on FIRE Data Set with TF-IDF and Probabilistic Models". International Journal Computer Application, Page No.22-25.

[17] Hamel, P., & Eck, D. (2010, August). Learning Features from Music Audio with Deep Belief Networks. In ISMIR (pp. 339-344).

[18] Vigliensoni, G. JWEBMINER REFINEMENT FINAL PROJECT REPORT.

[19] McKay, C., & Fujinaga, I. (2010, March). Improving automatic music classification performance by extracting features from different types of data. In Proceedings of the international conference on Multimedia information retrieval (pp. 257-266). ACM.

[20] Tzanetakis, G. (2014). Music Information Retrieval.

[21] Friberg, A., Schoonderwaldt, E., Hedblad, A., Fabiani, M., & Elowsson, A. (2014). Using perceptually defined music features in music information retrieval. arXiv preprint arXiv:1403.7923.

[22] Pesek, M., Leonardis, A., & Marolt, M. (2014). A compositional hierarchical model for music information retrieval. In Proceedings of the International Conference on Music Information Retrieval (ISMIR), Taipei.