ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

Raga Identification Using Datamining Techniques.

Santhi B*, and Sushmetha N.

Department of Information and Communication Technology, SASTRA University, Thanjavur, Tamil Nadu, India.

ABSTRACT

The Raga system is a method of organizing tunes based on natural principles. Each raga has a distinct image which defines a raga. Raga falls into two categories the Melakarta and the derived ragas. There are 72 Melakarta ragas. The classification of the input signal into raga can be done based on pitch, inharmonicity, energy, pulse clarity, brightness and roll off. In this work the input is classified as Bhairavi or Mohanam. The selection of features and the efficiency of classification techniques has also been discussed.
Keywords: Raga classification, Indian music, Feature extraction, Data mining

**Corresponding author*

INTRODUCTION

The Indian music branches into Hindustani and Classical music. Music in general is characterized by two fundamental elements[1] , namely raga and rhythm known as tala. Raga is the backbone of music.Swara (or notes) is the fundamental attribute of raga. The amalgamation of the seven swaras is the basis of any raga. The sargam is formed by shadja,rishabha,gandhara,madhyama,panchama,dhyavata,nishadha[9].The shadja note is called the base note with respect to which frequency of the other swaras are aligned. The drone[8] swaras include sa and pa. Though, all the notes except the drone have one or more variants. The Parent ragas are formed from the fundamental swaras. The 7 swaras can be divided[7] further that leads to the derived ragas[12].These allied ragas are composed from the same set of notes and might be difficult to distinguish.The raga identification involves feature extraction. There are various application of raga identification such as in development of music recommendation, music indexing etc.The objective of this work is to classify the given signal as bhairavi or mohanam. Both are janaka ragas. While Bhairavi is a sampoorna raga, Mohanam is not .Mohanam does not have the swaras Ma,Ni and hence referred to as symmetric.

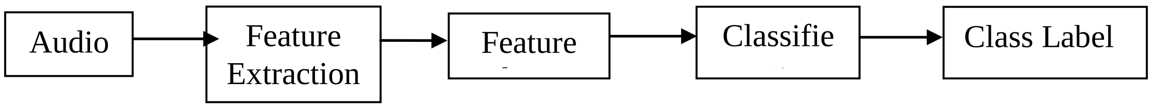

Fig1.Block diagram for the proposed system

RELATED WORKS

R. Pradeep [8]et al. have proposed a discriminating approach for identifying 4(Ahir Bhairav, Bhairavi, Bhupali and Deshkar) of the 51 ragas of the Hindustani music. The preprocessing of the music clips is done to remove noisy regions. The salience function based approach has been adopted for extraction of the main melody. The proposed method has three fold advantage. The ragas with different notes, a change in single note and same notes with change in duration have a higher deviation.

There are five deviation methods that have been employed; average based, median based, minimum value based, maximum value based and global threshold based . An approach similar to KNN classifier has been adopted to build a system to identify the ragas using the 11 features of the music clips.

B.Santhi [2]et al. proposed a robust method for feature selection which is the fundamental of automatic classification of music. The first method involves an experimental setup with fourteen features and the second method involves 8 out of the 14 features which include(RMS, Energy, Roll off, Pitch, Inharmonicity, Flatness, Entropy, Low Energy and Brightness).The two sets of features are applied to various classifiers which include neural network,SVM(support Vector Machine),KStar,J48,Logistic Regression and adaboost envisaging the second method to be successful and proving the right selection of features .

R.K. Nawasalkar[11] et al. proposed a performance analysis of polyphonic inputs of yaman raga.The audio signals take variety of forms like speech, music and environmental sound. The features of music clips from various sources were extracted which form the backbone for the analysis. The MIR Toolbox is used for feature extraction.

MATERIALS AND METHODS

AUDIO FEATURE SELECTION:

It is important to select optimum number of features from a variety of features .In order to achieve better classification , the classifier should be invariant to noise, amplitude scaling or bandwidth of the audio signal. The feature set obtained should also eliminate redundancies and achieve high optimization .Features decide the class and nature of the audio signal. The variety of techniques used for classification can be categorized into temporal and spectral analysis[3]

The features are extracted by breaking the input signal into windows/frames of 100-200ms.The features has been extracted using the MIR toolbox. The feature value is computed for each of the windows. In the first approach all features are taken for analysis and class assignments are done in real-time. This consists of a single texture window and thus it is represented as a single feature vector. Another approach is the texture window in which a number analysis window is created. Only one feature vector is generated for each analysis window. The features are obtained through statistical measures from all the analysis windows .Thus real time classification is not possible. Several class decisions are averaged to obtain a final decision. Fig 1 describes the flow of this work.

CLASSIFICATION

The feature selection is followed by classification. The classification of the preprocessed data can be done in two steps .The first step of the classification process involves training the selected classifier using predetermined data set. This is called the training phase. Class labels are assigned to the audio signal. These labels are used to identify the signal. Depending on whether the class label is known in advance or not, it is classified into supervised and unsupervised learning .The mapping of tuples or the samples to the associated class label is represented either graphically in the form of decision tree or using mathematical rules. The percentage of correctly classified data is determined by the accuracy of the classifier .This is the second step which involves, testing randomly selected tuples.

Classifiers can be classified based on their real time capabilities. Real time classifiers update the results in intervals of milliseconds [3].While non-real time classifiers analyze a longer fragment of the input signal and hence are more accurate than the real time classifiers.

As stated in [3] they can be classified into direct and hierarchical. In direct approach the classes are decided in a single step by analyzing all the features but in case of hierarchical approach the genre dependency is used so that, at each step only one feature is used to distinguish between the sub features. Hence the hierarchical approach accounts for dependency. The addition of a new class would mean to modify the algorithm and re iterate the process. In case of hierarchical only one branch would be modified. Though the direct method is susceptible to errors , it is simple and less expensive than hierarchical. The two methods used here are Naive Bayes and Decision tree.

NAIVE BAYES

The statistical classifier can identify the given tuple belonging to a particular class. The results are accurate and perform better as compared to decision trees and other neural networks. It forms its working based on the Bayes theorem. It effectively assigns class labels to tuples. An attribute of a class does not influence the other attributes. This assumption is called conditional independence.

DECISION TREE

Decision tree is a top down approach for testing the tuples. Classification of the given tuple with unknown class labels using decision tree yields high accuracy. The decision tree is a tree-like flow chart with internal nodes representing the attributes. The decision tree classifies the entire data space into mutually exclusive regions. Rules can be derived from the constructed decision tree; popularly identified as disjunctive normal form (DNF)[10].The decision tree is built using the training data and used to classify the test set. This is the first phase. We begin with the root node and if the examples can be fall under two or more classes then the process has to be carried out recursively leading to a discriminative tree. The process is continued until element have been distinctively classified. The leaf nodes are the solution or the classified data.

EXPERIMENTAL SETUP AND RESULTS

The dataset has been manually annotated. A total of 160 excerpts belonging equally to the two ragas were extracted and applied as input to the MIR toolbox. The features were extracted using MIR Toolbox[6] . The toolbox is dedicated to the Mat lab directory. The training database of the audio is present in the classified form. The music files of Bhairavi and Mohanam are inputted to the toolbox for extraction of various features

ISSN: 0975-8585

like Pitch, Roll-off, inharmonicity, pulse clarity, mode, tempo and brightness. A sample of the dataset has been illustrated in Table 1.

Table 1.Sample dataset of features extracted for Bhairavi and Mohanam

	Low Energy	Tempo	Pulse Clarity	Roll Off	Brightness	Inharmonicity	Mode	Label
1	0.42101	178.5423	0.037728	2791.2099	0.32309	0.56268	-0.19333	y
2	0.45894	117.8247	0.11795	2298.8548	0.29852	0.52974	-0.19172	y
3	0.53105	69.5256	0.14624	2633.6222	0.35372	0.58459	0.19977	y
4	0.54189	77.7911	0.042149	2903.9964	0.27831	0.6603	0.0068507	y
5	0.4827	131.1617	0.17318	5010.1707	0.48515	0.45787	0.015428	y
6	0.44143	112.4976	0.048085	3044.6877	0.3648	0.53643	-0.10544	y
7	0.48677	185.4024	0.15919	4247.8449	0.47292	0.52409	0.26461	n
8	0.47677	130.6208	0.15103	4081.4357	0.7189	0.42643	0.26795	n
9	0.50698	123.1959	0.31656	3645.5986	0.46701	0.46927	0.070659	n
10	0.52719	181.2843	0.18742	5955.2502	0.64555	0.55351	0.11356	n
11	0.52698	150.2223	0.44671	3738.3081	0.5621	0.49625	0.056923	n
12	0.63034	139.6384	0.051391	2459.0592	0.27131	0.56361	0.10634	n

Data Analysis:

Pitch:

Pitch is a perpetual property of sound .The sound produced by the vibration of the vocal chord is categorized into unvoiced and voiced sounds depending on whether they produce vibration of vocal chords or not[3].The voiced sound produces glottal pulse. The frequency of the glottal pulse produced by the vibration of vocal chord is called pitch. Tone is determined by the frequency of vibration and is represented by a letter. The pitch is designated by the first harmonic of a tone whereas absolute pitch is determined by the number of vibrations per second irrespective of the other tones.

Timbral features:

Timbre distinguishes sounds that have same pitch and loudness[4].For instance it distinguishes between guitar and piano playing same note at same pitch and loudness[14]. It is often determined by vibrato and tremolo which are the dynamic characteristics of sound and the harmonic content of sound. The musical note is made up of distinct frequencies. Each musical note is measured in terms of Hertz(Hz). The pitch of the lowest frequency defines the note. The sound is a combination of fundamental (lowest) frequency and harmonic. The harmonic is obtained by incrementing in terms of integral multiples of the fundamental frequency -x2,x3,x4 etc[3].The inharmonicity of tones is caused by the deviation from whole multiples of fundamental frequency. There are many instruments containing overtones and demonstrate inharmonicity. Timbral features unlike rhythmic express short term spectral information[13].

The periodic changes in pitch are defined as vibrato while the periodic of amplitude of tone is defined as tremolo. Thus vibrato, a desirable characteristic of human could be called frequency modulation (FM) and tremolo could be called amplitude modulation (AM).

Roll off:

It is feature which is used to approximate the high frequency in the signal. Nearly 85 percent magnitude of the spectrum is concentrated below roll-off [3].The spectral shape is determined using this feature. It is directly proportional to the frequency. The equation for roll off is

$$\sum_{k=1}^{M} X_r \ ([k]) = 0.85\sum_{k=1}^{N/2} X_r\,[k]$$

If M is the largest value of k for which this equation is satisfied, then this frequency M is the roll off.

Inharmonicity:

It is the estimation of partials [11] that are not multiples of fundamental frequency is inharmonicity. The value lies between 0 and 1[2, 5].

Other features:

Low energy rate:

Energy is the measure of percentage of frames with RMS lesser the mean. It determines if the signal remains constant throughout, or the frames are constrained [2].It is computed on texture window basis [3].

Pulse Clarity:

It estimates the rhythmic clarity, indicating the strength of the beats. It plays an important role in genre classification. It provides subjective characterization.

Brightness:

It sets the cutoff frequency and also measures the amount of energy above that[5].

Tempo:

It demonstrates the rhythmic content of the song. It is calculated as the speed of the song in beats per minute (BPM)[15].

Table2:Comparison of the ability to predict true class

	Naive Bayes	ADTree
Correctly Classified Instances	151	156
Incorrectly Classified Instances	8	3
Kappa statistic	0.7712	0.9123
Mean absolute error	0.0487	0.0745
Root mean squared error	0.1861	0.1468
Relative absolute error (in percent)	11.0183	16.847
Root relative squared error (in percent)	41.7809	54.7485

The kappa statistics measures the agreement of prediction with the true class where 1.0 signifies complete agreement. The error values are more meaningful for regression tasks. The Mean Absolute Error(MAE) is the average of the absolute value of the forecast and the observation of the verification set . The Root Mean Squared Error is the average of the square between the forecast and the corresponding observed value. Table 2 shows a detailed view of various performance measures .

RESULTS AND DISCUSSION

The analysis of features of the training set of data for the Bhairavi raga has been discussed in Table 3. The classifier is trained with few elements selected from the dataset. This is the learning phase. The features can be chosen according to the user's discretion.

Table 3:Representation of Naive Bayes classifier on the training set for class='Bhairavi'

	Mean	Std. Deviation	Weight Sum	Precision
Low Energy	0.492	0.0558	14	0.0107
Brightness	0.3604	0.0987	14	0.0187
Pulse Clarity	0.1423	0.0955	14	0.0178
Tempo	127.6375	38.6286	14	5.7273
Roll off	3554.239	1913.33	14	365.877
Inharmonicity	0.5303	0.0593	14	0.0114
Mode	-0.1012	0.0945	14	0.0205

The True Positive(TP) rate is the fraction of samples identified as belonging to a class by the classifier of all those which belong the class.

The False Positive (FP) rate is the fraction of instances identified wrongly as belonging to a particular class among those which belong to other class.

Precision is the fraction of samples identified as belonging to a class among those samples which actually belong .

The F-Measure is the harmonic mean of precision and recall. It can be calculated as stated

F-Measure=2*(Precision *Recall)/(Precision + Recall)

The accuracy estimation for decision tree is as follows

Table4: Comparison of accuracy and other factors between the two classifiers

Classifier Name	Accuracy	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area
Naive Bayes	94.96%	0.95	0.178	0.95	0.95	0.95	0.987
AD Tree	98.113%	0.981	0.088	0.981	0.981	0.981	0.922

The accuracy is around 94.96 % for Naive Bayes while it is 98.113% in case of AD Tree. Table 4 gives a detailed comparison of the two classifiers. It shows that the accuracy of the Naive Bayes is 94.96% and that of AD Tree is 98.113%.

CONCLUSION

In this work, an effort has been made to identifies if the given song is Bhairavi or Mohanam. The various features have been used for effective classification. Naive Bayesian and Decision Tree have been employed in classification. The accuracy has been estimated for both the classifiers. The efficiency of the classifiers are determined using the confusion matrix. The decision tree exhibits higher accuracy as compared to Naive Bayes.

As a future work, the same technique can be used to classify any given song into a particular raga. This can be done with the help of the arohana-avarohana pattern which is unique to each raga.

REFERENCES

[1] Prithvi Upadhyaya, Suma S. M., and Shashidhar G. Koolagudi(2015).”Identification of allied ragas in carnatic music”.IC3 '15 Proceedings of the Eighth International Conference on Contemporary Computing (IC3), IEEE Computer Society Washington, DC, USA,pp 127-131

[2] P.Ramesh Kumar,M.Monisha ,B.Santhi, T.Vigneshwaran(2014).”Robust Feature Selection Method for Music Classification”. International Conference on Computer Communication and Informatics (ICCCI),IEEE.

[3] Hariharan Subramanian(2004) .”Audio Signal Classification”. Credit Seminar Report, Electronic Systems Group, EE. Dept, IIT Bombay,.

[4] YF Huang, SM Lin, HY Wu, YS Li(2004) ,”Music Genre Classification Based on Local Feature Selection using a Self Adaptive Harmony Search Algorithm”,Data Knowledge and Engineering,Elsevier ,Vol.92,pp 60-76.

[5] P.Ramesh Kumar,M.Monisha ,B.Santhi(2014) ,”Enhancement of information hiding in audio signals with efficient LSB based methods”,Vol.7,pp 80-85.

[6] Ms.P.Kirthika, Dr. Rajan Chattamvelli(2012) ,”A Review of Raga Based Music Classification and Music Information Retrieval (MIR)”,IEEE Conference on Innovative Practices and Future trends, In proc. of IEEE.

[7] B.Santhi and N.Sairam(2011),”Melakartha raga generation through breadth first search algorithm”,Journal of Theoretical and Applied Information Technology,Vol.31No.2.

[8] R. Pradeep, Prasenjit Dhara, K. S. Rao and Pallab Dasgupta(2015),”Raga Identification Based on Normalized Note Histogram Features”,International Conference on Advances in Computing,Communication and Informatics(ICACCI),IEEE.

[9] Vijay Kumar*, Harit Pandya* and C.V. Jawahar(2014),”Identifying Ragas in Indian Music”,International Conference on Pattern Recognition,IEEE,pp 767-772.

[10] [Chidanand Apte and Sholom Weiss (1997),”Data mining with decision trees and decision rules “,Data Knowledge and Engineering,Elsevier ,Vol.13,Issues2-3,pp 197-210.

[11] R.K. Nawasalkar, V.M. Thakare , N.D. Jambhekar and P.K. Butey(2015),”Performance Analysis of Different Audio with Raga Yaman”,First International Conference on Next Generation Computing Technologies,Dehradun.

[12] Trupti Katte and Prof. B. S. Tiple(2014),”Techniques for Indian Classical Raga Identification– A Survey “,Annual IEEE Indian Conference(INDICON)

[13] Babu Kaji Baniya*, Deepak Ghimire and Joonwhoan Lee (2014),”Automatic Music Genre Classification Using Timbral Texture and Rhythmic Content Features “,ICACT Transactions on Advanced Communications Technology (TACT) Vol. 3, Issue 3.

[14] Sujeet Kini, Sankalp Gulati and Preeti Rao(2011) ,”Automatic Genre Classification of North Indian Devotional Music,”National Conference on Computing”,Bangalore,India, pp 1-5.

[15] Babu Kaji Baniya, Joonwhoan Lee and Ze-Nian Li,”Audio Feature Reduction and Analysis for Automatic Music Genre Classification”(2014),IEEE International Conference on Systems, Man, and Cybernetics ,San Diego, CA, USA,pp 457-462.

[16] http://www.cs.usfca.edu/~pfrancislyon/courses/640fall2015/WekaDataAnalysis.pdf.