# Characteristic Behaviour of Artist from Raga Performance

Poonam Priyadarshini*
Department of Electronics and Communication
Birla Institute of Technology, Patna, Bihar, India.
*ppbitpatna@gmail.com

## Abstract

Music information retrieval (MIR) is an emerging research area that receives growing attention from both the research community and music industry. It addresses the problem of querying and retrieving certain types of music from large music database. People searched music of their choice from database by song title, composer and performer. Features of interest may include melody, harmony, rhythm, and instrumentation. Classification is a fundamental problem in MIR. Many tasks in MIR can be naturally cast in a classification setting, such as genre classification, mood classification, artist recognition, instrument recognition, etc. The MIR research aims to develop new techniques for processing musical information and searching music databases by content. This paper devises a more practical and efficient approach to MIR by investigating a variety of statistical and signal processing based features, such as Fast Fourier Transform (FFT), Linear Predicative Coding coefficients LPC, Pitch . The feature can be used to make a comparison between different artists singing the same raga. In this paper we are using Indian music Raga as a database and have implemented MFCC algorithms to extract the features of song.

**Keywords** – MFCC, Projection Pursuit, Raga, Gharana, Artist Classification.

## Introduction

Indian classical music is structured using a 12 note scale consisting of 7 basic notes in addition to 5 interspersed half notes. Unlike western classical music however, the base frequency of the scale is not fixed. Additionally, intertonal gaps or the temperament of the scale may also vary. The nucleus of Indian classical music is the raga. A raga may be defined as a melodic structure with fixed notes (minimum five including Sa and at least one of Ma and Pa) and a set of rules that characterizes a particular mood and is conveyed by performance. The ragas have specific ascent (aroha) and descent (avaroha) sequences that may not necessarily be identically reverse. Every raga is assigned rules for pitch transitions.

However, artists are given considerable latitude to improvise within these norms. Although the basic nature of Indian musical tradition is Mathematical analysis is therefore made difficult owing to the fact that the same raga when rendered by artists of two different schools or Gharanas, may well have a very different signature in each rendition. In music, Gharana refers to a biological lineage (father-son etc.) to begin with that manifests into a nice disciple-chain.

Since MFCC is a powerful and popular signal processing tool that can provide a clearer scientific picture of the musical pieces rather than provide just a set of acoustical features, our strategy is to first use MFCC to get a multidimensional point that represents the signature of the artist scientifically and then convert it to two dimension (so that it is visualized), whereupon it becomes a curve, using projection pursuit. The literature on projection pursuit suggests that if two multidimensional points are close, so would be the corresponding two dimensional curves. Projection pursuit is discussed in more details in section 3.

The rest of the paper is organized as follows: Section II describes the relevant literature work for recognition of ragas . In section III, the proposed method is based projection pursuit techniques along with MFCC to analyze the similarity between different artists singing the same raga. Section IV provides analysis of the proposed method and result respectively.

**Related Work**

A large number of distinct feature sets, mostly developed for speech recognition, have been proposed to represent audio signals. Typically Audio signals are based on some form of time-frequency representation. Mel-frequency cepstral coefficients (MFCC), illustrates the shape of the spectrum of the audio signal and are widely used in speech recognition (Lee *et al*., 2006). In Indian classical music mainly the research performed on automatic Indian music information, recognition, and classification of *ragas* and development of retrieval systems. A Hidden Markov Model and Neural Network for recognition of *ragas* and notes with use of *Arohana* and *Awarohana* sequence (Gulati and Rao *et al*., 2016). Gaurav Pandey *et al*., (2003) suggested note transcription system 'TANSEN' and used Hidden Markov Model and string matching technique to get *pakad* notes of *ragas*. Chordia *et al*., (2007) used recognition of annotation, onsets detection, pitch class distribution (PCD) and pitch class dyad distribution (PCCD) using SVM, MVN and Random Forests. In this Chordia also attempted classification of tone profiles and spectral profile, pitch detection, onset detection, PCDs and PCCDs using HMM, MVN, FFNN, KNN, Tree based and Bayesian classifier. Neural Network self organized Maps (SOM) and Bayesian decision rule for recognition of pitch profile, pitch class distribution are described in (Dighe and Karnick *et al*., 2012; Gulati and Rao *et al*., 2016). The signal separation, segmentation and string matching for signal frequency offset and onset in( Sridhar and Geetha *et al*., 2006 and 2009). In (Belle and Joshi *et al*., 2009) statistical framework is used for vocal pitch, folded pitch distribution (FPDs), *Swara* annotation and pitch class distribution (PCDs). Chordia (Chordia *et al*., 2008) used MFCC, GMM and PCD for retrieval of timbre and PCD for artist recognition, instrument recognition and *Thaat* classification. Shetty (Shetty *et al.*, 2009) used ANN for note transcription and *Arohana*, *Awarohana* pattern for *raga* recognition.

**Feature Extraction**

*Timbral Texture Features*

Timbre describes those characteristics of audio signals that allow us to differentiate between audio signals having the same pitch and loudness. It is determined primarily by the harmonic content and the dynamic characteristics of the signal. Following features are used here to determine timbral characteristics Melfrequency cepstral coefficients (MFCC) features are also based on the Short Time Fourier Transform (STFT).The Mel scale is a scale of pitches perceived by listeners to be equally spaced. Mel-Frequency Cepstral Coefficients are coefficients that collectively form a representation of the short-term power spectrum of an audio sample. They are derived from a cepstral representation of the audio clip. The frequency bands are spaced equally along the Mel scale. This gives a closer approximation of the human auditory response as compared to a linear spacing of frequency bands. The basic steps to be followed in deriving MFCCs are as follows:1.pre-emphasis 2. Framing 3. Hamming windowing 3. Fast fourier Transform to obtain power spectrum 4. Log of FFT 5.Mel filter bank 6.DCT for decorrelation 7.delta delta MFCC coefficients.

$$F(Mel) = [2595 * \log 10[1 + f]700] \qquad \ldots(1)$$

The first few coefficients obtained by this method contain most of the signal energy. Hence, usually the first five coefficients are taken as features. Using MFCC a feature set was created consisting of statistical and acoustic features to analyze the audio samples. There are 12 features in all leading to a 12 dimensional point to represent the artist's signature or identity.

*Statistical Features*

1. *Skewness :* Skewness is the measure of the symmetry or asymmetry of a signal.

$$Skewness = \frac{\sum_{i=1}^{N}(Y_i - Y)^3}{(N\text{-}1)_s^3} \qquad \ldots(2)$$

$Y$ is the mean, $s$ is the standard deviation, and $N$ is the number of data points.
2. *Kurtosis:* Kurtosis is a measure of the noisiness of a signal. It shows whether the data is peaked or flat relative to a normal distribution.

$$Kurtosis = \frac{\sum_{i=1}^{N}(Y_i - Y)^4}{(N\text{-}1)_s^4} \qquad \ldots(3)$$

Projection pursuit methods are very useful tools in the area of high-dimensional data analysis. They do however present many limitations as enumerated by Crawford and Fall [19].The technique employed here for dimensional reduction is the Andrews plot method. An Andrews plot or Andrews curve is a method to visualize multi-dimensional data in two dimensions [20]. Each point

x = {x1, x2,…,xn} in the dataset denotes a finite Fourier series.

$$f_{x(t)} = \frac{x_1}{\sqrt{2}} + x_2\sin(t) + x_3\cos(t) + x_4\sin(2t) + x_5\cos(2t) + \text{---...} \qquad \ldots(4)$$

This function is then plotted for $-\pi < t < \pi$

Each point is now a line between $\pi$ and $\pi$. Any structure in the dataset is now visible as curves on a two-dimensional plot. By a property of Andrews plots, any if any two N-dimensional points are close, the corresponding curves in two dimensions are also close accordingly. Hence, we can identify clusters among the curves and tell which artists are close in rendering the same raga.

If two artists train under the same teacher, by comparing the performances of the two artists with that of the teacher's, we can determine which artist more closely resembles the teacher's individual style. Artists from the same gharana, or school of thought, are more likely to produce curves closer to one another when they render the same raga. For artists from different Gharanas, the curves will be further apart. Thus projection pursuit can help in the classification of renderings of ragas according to the Gharana of the artist responsible for the particular rendition.

**Analysis and Results**

We are choosing indian raga eg. Todi Raga and Bihag Raga for doing analysis of music signal. $P_{1\text{-}}$ difference between weighted average note pitch and the pitch of the lowest note of a phrase:

$$P_1 = \frac{\sum_{i=1}^{N} p_i d_i}{T} \; M_{in}(p_i) \qquad \ldots(5)$$

Where T is the phrase duration, $p_i$ denotes the pitch (at the onset) of the i-th note and $d_i$ denotes the duration of the i-th note(departure of the i-th note- onset of the i-th note), N denotes the number of note in a phrase.

the difference between the pitch of the highest and the lowest note of a phrase

$$P_2 = Max(p_i) - Min(p_i) \qquad \ldots(6)$$

$P_3$ = the average absolute difference of the pitches of subsequent note.

$$P_3 = \frac{1}{N-1} \sum_{i=1}^{N-1} | p_i - p_{i+1} | \qquad \ldots(7)$$

$P_4$ = the duration of the longest note of a phrase

$P_4 = Max(d_i)$ …(8)

$P_5$ = the average note duration

$P_5 = \frac{1}{N} S^{N}_{i=1} d_i$ …(9)

Statistical parameters representing a musical phrase can be divided into two groups: parameter describing melodic quantities of musical phrase ($P_1$, $P_2P_3$) and the parameter describing rhythmical quantities of musical phrase ($P_4$, $P_5$). we are finding the pitch and time duration between notes from praat software, the value of ($P_1$, $P_2P_3$) ($P_4$, $P_5$) are calculated by C++ programming then through matlab simulation Andrew plot is created. when artists from the same Gharana are renderings the same raga then there is closeness in the curve and there is separation in the curve when the same raga is rendered by artists of different Gharanas. The result are shown in the figures below of feature extracted by MFCC of todi segment 1 and 2 and the Andrews plots of Todi Raga and Bihag Raga sung by various Artists .

The data used for result and its simulation graph are shown after the references. Skewness and Kutosis of todi raga segment is calculated.

**References**

J.P.Bello, J. Pickens, 2005. A robust mid-level representation for harmonic content in music signals. *Proc. ISMIR 5*, 304–311.

K. Lee. 2006. Identifying cover songs from audio using harmonic representation. *MIREX (Music Information Retrieval Evaluation exchange) task on Audio Cover Song Identification.*

P.Dighe, H.karnick 2012. Swara histogram based structural analysis and identification of Indian classical music. *ISMIR*.

S.Gulati, P.Rao 2011. A Survey of raag recognition techniques and improvement to the state of the art. *Sound and MusicComputing.*

Pandey G., Mishra C., Paul I. 2003. A system for automatic raga identification. Proc. of the 1st Indian International Conference on *Artificial Intelligence.*

Chordia P., Rae A. 2007. Automatic Raag classification using pitch-class and pitch-class dyad distributions. *Proc. of the 7th International Conference on Music Information Retrieval (ISMIR).*

Chordia P., Rae A. 2007. Modeling and visualizing tonality in North Indian classical music. *Neural Information Processing Systems (NIPS), Music Brain Workshop.*

Chordia P.Rae A. 2008. Raag vidya- real-time Raag recognition for interactive music. *Proc. of the International Conference on New Interfaces for Musical Expression (NIME), Genova, Italy.*

Sridhar R.,Geetha T.V. 2006. Swara identification for South Indian Classical Music. *9th International Conference on Information Technology (ICIT'06)*, 143-144.

Sridhar R., Geetha T.V. 2009. Raga identification of Carnatic music for music information retrieval. *International Journal of Recent Trends in Engineering*,1(1).

Belle S., Joshi R. and Rao P. 2009. Raga identification by using Swara intonation. *Journal of ITC Sangeet Research Academy*, 23.

Chordia P., Godfrey M. Alex Rae, 2008. Extending content-based recommendation- The case of Indian classical music. *Content-Based Retrieval, Categorization and Similarity, ISMIR* .

Shetty S., Achary K.K. 2009. Raga mining of Indian Music by extracting Arohana-Avarohana pattern", *International Journal of Recent Trends in Engineering,1(1).*

Chordia P., Rae A. 2007. Raag recognition using Pitch-class and pitch-class dyad distributions. *Proc. of ISMIR*, 431-436.

Sridhar R., Geetha T. 2009. Raga Identification of carnatic music for music information retrieval. *International Journal of Recent Trends in Engineering*, 1(1), 571-574.

Kirthika P, Chattamvelli R. 2012. A Review of Raga Based Music Classification and Music Information Retrieval (MIR) *Proc. of IEEE*. 1-5

**Todi Raga**

Todi segment 1 ,Time duration: 7.452 ,No. of Notes: 9

Frequencies:79.229681 79.734492 79.283573 79.520801 149.173553 149.555045 155.784859 155.935460 187.822017 187.724698 148.550787 148.098928 157.079779 157.250142 132.151783 132.285104 166.558278 166.029674

Time:3.072506 3.182506 3.222506 3.262506 3.462506 3.482506 3.992506 4.012506 4.222506 4.242506 5.962506 5.982506 6.202506 6.242506 6.642506 6.662506 7.422506 7.452506

p1 = 1006.018790 p2 = 108.371178 p3 = 26.970213 p4 = 7.437506 p5 = 7.496207

Todi segment 2: ,Time duration: 4.8024 ,No. of Notes: 10

Frequencies:89.926760 89.930248 90.207964 90.895479 80.597706 80.323548 75.726120 75.840864

80.119949 80.177241 117.547598 117.547598 95.524046 95.354851 80.271337 80.744238 92.981983 92.101331 79.544591 79.656395

Time:0.562417 0.612417 0.622417 0.662417 1.792417 1.812417 2.882417 2.902417 3.032417

3.052417 3.432417 3.452417 3.962417 4.002417 4.092417 4.142417 4.392417 4.412417 4.772417 4.802417

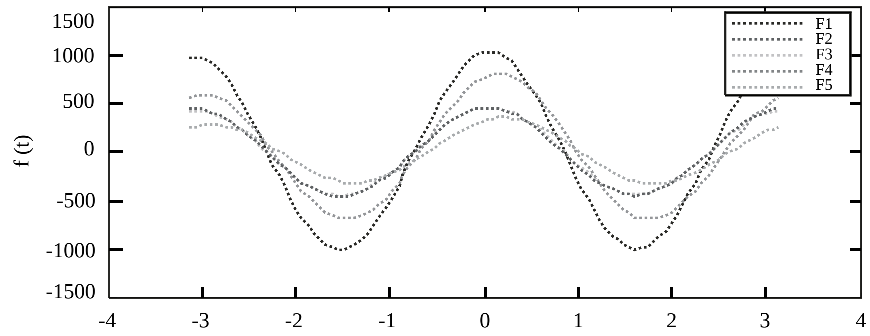
p1 = 445.011141 p2 = 41.764107 p3 = 13.241155 p4 = 4.787417 p5 = 5.895031

Todi segment3: ,Time duration: 1.795 ,No. of Notes: 10

Frequencies: 108.008 108.171 106.964 106.255 100.118 100.354 97.988 97.263 97.006 97.938 98.601 98.153 95.924 95.762 96.001 96.673 100.117 100.747 134.130 134.772

Time: 0.265 0.305 0.335 0.375 0.515 0.555 0.625 0.665 0.685 0.725 0.835 0.875 1.285 1.325 1.345 1.385 1.475 1.515 1.755 1.795

p1 = 452.039004 p2 = 38.608002 p3 = 5.851611 p4 = 1.775000 p5 = 5.192201 similarly Todi segment 4 and 5 are analysed and graph is plotted below.




Andrews plots of todi raga by various Artists

**Bihag Raga**

Bihag segment 1 , Time duration: 6.560 ,No. of Notes: 17

Frequencies:178.045389 178.067861 167.727867 167.127692 166.839979 166.452918 168.516458 168.012153 167.644395 167.896816 166.299863 166.561861 169.441572 169.152444 134.001534 134.422676 133.595738 133.316315 225.971835 225.054906 200.654598 200.080248 200.307598 200.281928 198.861871 198.261671 196.391755 196.733337 134.034967 134.046448 133.824122 133.937438 178.378562 178.977435

Time: 0.340351 0.380351 0.520351 0.550351 0.560351 0.680351 0.730351 0.760351 0.770351 0.800351 0.900351 0.930351 1.020351 1.060351 1.190351 1.300351 1.730351 1.770351 1.930351 1.970351 2.650351 2.700351 3.890351 3.960351 4.020351 4.070351 4.130351 4.220351 4.360351 4.430351 4.620351 4.700351 6.500351 6.560351

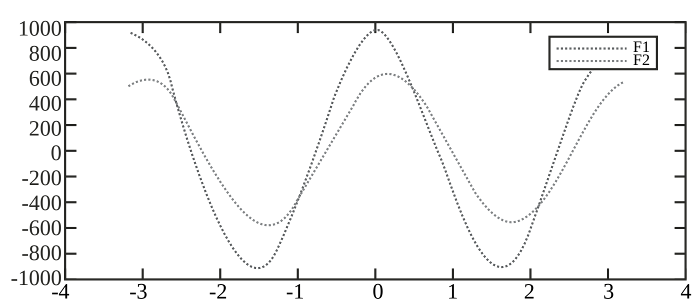
p1 = 913.371085 p2 = 92.057343 p3 = 17.628485 p4 = 6.530351 p5 = 6.026877

Bihag segment 2:

Time duration: 8.834

No. of Notes: 9
Frequencies: 273.252 273.171 288.234 283.983 139.639 139.987 141.052 141.113 142.496 142.248 14 1.688 141.863 112.952 112.687 93.790 93.330 92.626 92.726

Time: 1.094 1.134 1.214 1.254 3.294 3.324 3.704 3.734 4.094 4.124 4.794 4.834 7.344 7.374 8.704 8.730 8.774 8.834
p1 = 533.223948 p2 = 193.432526 p3 = 26.430946 p4 = 8.804000 p5 = 5.029001




Andrews plots of Bihag raga by two Artists

Skewness

Todi raga 1st segment. The Spectral skewness related = 4.9419
Todi raga 2nd segment. The Spectral skewness related = 5.9917
Todi raga 3rd segment. The Spectral skewness related = 5.0561

Kutosis

The Spectral kurtosis todi seg1 = 31.833
The Spectral kurtosis todi seg2 = 56.2213
The Spectral kurtosis todi seg3= 32.0919