Feature Selection For Indian Instrument Recognition Using SVM Classifier

Ms. Ashwini
Department of Electronics and communication,
PES University
Bengaluru,India
ashwinib@pes.edu

Dr. Vijaya Krishna A,
Department of Electronics and communication,
PES University
Bengaluru,India
vijayakrishna@pes.edu

***Abstract*— The Indian Classical Music is extremely varied and is of unique types across the globe. In order to achieve the uniqueness, it uses a wide diversity of Musical Instruments. This paper considers only the Indian musical instruments. Also it provides a study on feature analysis for the identification of Indian musical instruments. ML techniques are utilized for the purpose of choosing and evaluating features that are extracted from various instruments. Efforts are made to enhance the dissimilar parts of the detection system and the majority of the effort is concerned with the features of the musical instruments and classification schemes that makes use of support vector machines (SVMs). The experiment is performed on seven different Indian instruments and the results indicates the capability of the proposed method by providing good discrimination among them.**

Keywords—Classical Indian Music, Feature analysis, Instrument recognition, Machine learning, SVM.

I. Introduction

Classical music in India has greatest continuous tradition consisting of variety of Musical Instruments. The monophonic and polyphonic signals have a simple and straight-forward meaning. In particular, musical genre classification of a given polyphonic audio signal can be analyzed better if the instruments associated in the clip are well-known. This information is used in order to narrow down the set of probable melodic genres. Here, the information related to the musical instruments can be utilized to produce the specific features of instruments in-order to get better quality of musical audio source separation. Majority of the literature provides the details of the work performed on western musical instruments [1][2][3][4]. This paper proposes to use only Indian Musical instruments (excluding percussions) and most of them are melody based. The major limitations in this work are,

I. Fewer instruments
II. Lesser accuracy

In spite of the obvious progressions that are obtained in the past years, the amount of work done on Indian instruments is quite less. This paper proposes a reliable method to identify Indian musical instruments in polyphonic audio tracks which overcomes a few of the major short comes of the past methods. Each instrument can be categorized based on particular set of features, selected from an overall group of all the 34 features. The research is focused on comparing the features that have significant impact on particular musical instrument and generating results consisting of predominant features of 7 Indian musical instruments that are under consideration.

Monophonic signal is a music with single type of instrument or a single melody track, which is of single vocal melody. On the other hand, polyphonic signal is a music having more than single melody. In recent years, identification of Indian musical instruments has gained more attention.

Here is a simple and dependable method proposed which can be used to recognize the Indian musical instruments in polyphonic audio which can overcome the restrictions posed in the past. The proposed research compares the features of significant impact on a musical instrument and producing dominant features for 7 musical instruments of India individually.

II. Feature Extraction

To obtain information from an audio signal, it is required to extract 34 feature set. Standard deviation, autocorrelation, skewness, and mean are considered because of its simplicity. Most of the features are based on Chroma vectors and MFCC co-efficient. Entropy of energy, Zero crossing rate (ZCR), and energy are features which fall under envelope based category and spectral flux, spectral roll-off, spectral spread, spectral entropy, and spectral centroid are considered as spectral features.

By considering the standard deviation and mean of all 34 features over individual frames, a total of 68 parameters are obtained. Based on these, the type of the instrument present in the polyphonic audio is recognized. For the purpose of classification of instruments, KNN and SVM are used. 68 parameters produces 68 dimensions in the classifier. Proper feature set are used to reduce the number of dimensions. For feature selection, 68 parameters are used and is discussed in detail in upcoming sections.

Following are the some of the features extracted.

Zero crossing rate: It is the rate of change of sign of waveform changes within a particular frame.

$$ZCR = \frac{1}{N-1}\sum_{n=1}^{n=N-1} 1_{R<0}(x[n] * x[n-1])$$

$1_{R<0}$ =Indicator Function
$x[n]$= Audio Signal
N = Length of audio signal.

Energy: It is the average of sum of the square of individual audio samples of length N. I can be defined as,

$$Energy = \frac{1}{N}\sum_{n=1}^{n=N-1}(x[n])^2$$

$x[n]$= Audio Signal
N = Length of the audio signal.

978-1-7281-4097-1/20/$31.00 ©2020 IEEE

Spectral Centroid: It is centre of gravity of the spectrum. It is given by,

$$SpectralCentroid = \frac{\sum_{n=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

$x[n]$ *is* Audio Signal track
$f[n]$ *is* Central frequency of the bin

Spectral Entropy: It is the measure of spectral power distribution.
$S(m) = |X(m)|^2$
$X(m)$ is the discrete Fourier transform of $x(n)$

$$P(m) = \frac{S(m)}{\sum_i S(i)}$$

The spectral entropy H is found as follows

$$H = -\sum_{m=1}^{N} . P(m) log2P(m)$$

Spectral Flux: It is the measure of rate of change of power spectrum of signal. It is determined by taking the difference between power spectrum of the neighbouring frames.
Spectral Spread: This is the second central moment of the signal.
Energy Entropy: It refers to the Entropy of normalized spectral energies for a set of sub-frames

$$Energy = E = \frac{1}{N}\sum_{n=1}^{n=N-1}(x[n])^2$$

$$H = -\sum .E \text{ log}2E$$

Where, H represents Energy Entropy of signal.

MFCC: These are the coefficients of cepstral representation. Here the frequency bands are distributed and are not linear according to mel-scale. A total of 13-coefficients are considered in this paper, out of which 2, 3 and 4 falls under excellent category.

Chroma Deviation: It is the standard deviation of twelve vectors. These vectors are known as Chroma Vectors and are the twelve elements representation of the spectral energy and the bins realize 12 equal-tempered pitch classes related to western-type of music.

III THE SUPPORT VECTOR MACHINES

As a well-known fact Support Vector Machine (SVM) is strictly defined by a separating hyperplane that is considered to be a distinct type of discriminative classifier. This algorithm is implemented with the help of various type of kernels. The choice of kernel is application dependent. The hyperplane learning in linear SVM can be achieved by transforming the problem using some form of linear algebra, which is not in scope given here.

Linear kernel SVM

The kernel applied here uses the dot-product and is written as

K(x, xi)=SUM(x*xi)

This linear kernel defines a distance measure which implies the similarity between the support vectors and the new data.

Feature Selection

For feature selection, a set of thirty four features from the trained audio sample is considered. It is trained using Linear kernel SVM and the accuracy of it found to be good. This system can lead to overload and decrease the performance. Even with minimal reduction in accuracy, by selecting a set of features, the performance can be increased. This compromise can be adopted to Indian instruments by selecting dominated feature sets. These features are categorized into poor, average and excellent categories. For demonstration, thousand samples from different instruments are used. It is shown in Figure 1, 2 and 3 respectively.

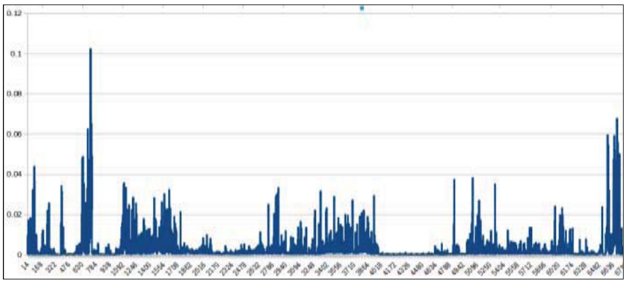

Figure 1. Chroma vector 1 Mean – Poor feature

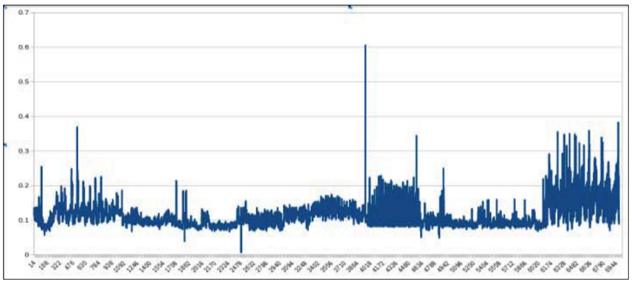

Figure 2. Spectral Flux Mean – Average feature

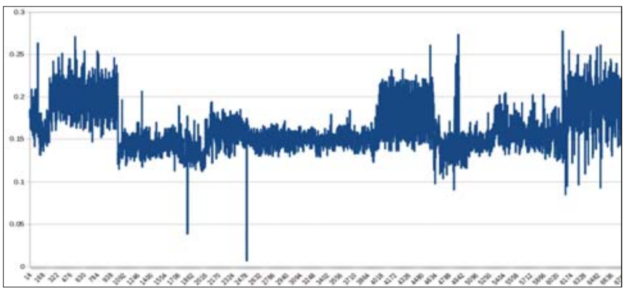

Figure 3. Chroma Deviation Mean – Excellent feature

By observing (Figure 1, 2, 3) the consistency of several feature magnitudes related to various instruments, each instrument can be differentiated. In the same way, for complete set of features, the analysis is performed and led to the characterization of individual feature amongst 34 features as poor, average and excellent. The Table 1 provides detailed listing.

Excellent	Good	Poor
Zero Crossing Mean Zero Crossing SD Energy SD Energy Mean Spectral centroid Mean Spectral centroid SD Spectral entropy Mean Spectral entropy SD Spectral spread SD Spectral flux SD Energy Entropy Mean Chroma deviation Mean Chroma deviation SD MFCC (2,3,4) Mean	MFCC (1,2,7,9,12) SD MFCC(1,6,9,10) Mean Spectral Spread Mean Spectral flux Mean Spectral Roll-Off Mean	Rest of the features.

Table 1. Category of features

IV Results and Comparison

Once completed the feature extraction and training them from the sample data using SVM, it is required to select c parameter from SVM. To select c parameter, it is required to calculate F1 measure for every related set of c. Table 2 provides the result obtained by considering these parameters.

For testing purpose, seven Indian instruments are considered, namely veena, sitar, flute, sarod, harmonium, shankha and violin. Overall F1 measure of SVM for 34 features is 97.2%. By making feature extraction from average and excellent categories, the Ff1 measure is found to be 93% SVM respectively. Further selecting features from excellent category alone, the f1 measure is 93% for SVM respectively.

c-parameter	accuracy	F1 measure
0.001	93.4	93.3
0.01	96.3	96.3
0.5	**97.1**	**97.2**
1	96.9	96.8
5	96.2	96.2
10	96.2	96.1
20	96.3	96.2

Table 2. c-parameter for all 34 features

svm	flu	har	Vee	sit	sha	sar	Vio
flute	94.15	0.73	0	0	0.73	0.61	3.78
har	0.2	97.78	1.01	0.3	0.51	0.1	0.1
vee	0	0.92	98.47	0.31	0	0.31	0
sit	0	0.81	0.61	98.38	0	0.1	0.1
sha	0.2	1.3	0.2	0.2	97.7	0.2	0.2
sar	0.1	0.91	0.4	0.3	0.1	98.18	0
vio	3.03	0.91	0.2	0.51	0.3	0.71	94.34

Table 3. Confusion matrix for all 34 features

The following results are derived by preferring only excellent features.

c-parameter	Accuracy	F1 measure
0.001	70.7	70.7
0.01	81.2	81.1
0.5	86	85.9
1	85.8	85.8
5	**86.2**	**86.1**
10	86	86
20	85.9	85.8

Table 4. c-parameter for excellent category feature

svm	flu	har	Vee	sit	sha	sar	vio
flute	85.4	1.1	3.8	0.2	3.3	1.8	4.4
har	0.61	86.77	2.63	2.63	4.95	1.21	1.21
vee	1.33	5.31	81.12	8.47	0	3.06	0.71
sit	0.4	4.34	5.05	87.68	0.1	1.92	0.51
sha	1.5	4.7	1.2	1	91.1	0.5	0
sar	1.92	1.61	1.92	0.71	0.51	92.12	1.21
vio	11.7	3.6	2.9	0.6	0.8	1.5	78.9

Table 5. Confusion matrix by considering excellent category feature

By preferring all average and excellent features, the following results are obtained.

c-parameter	Accuracy	F1 measure
0.001	82.3	82.3
0.01	90.3	90.2
0.5	92.4	92.4
1	92.1	92
5	**93**	**93**
10	92.6	92.6
20	91.6	91.6

Table 6. c-parameter for average category features.

svm	flu	har	vee	sit	sha	sar	vio
flute	93.9	0.2	0.8	0	0.7	0.6	3.8
har	0.3	92.2	1.41	2.22	2.63	0.3	0.91
vee	0.31	2.35	89.29	5.92	0.2	1.84	0.2
sit	0.1	1.92	2.32	94.95	0	0.4	0.71
sha	0.6	2.3	0.6	0	95.9	0.3	0.3
sar	0.7	0.3	1.52	0.2	0	97.07	0.2
vio	7.1	2.4	0.8	0.8	0.4	0.4	88.1

Table 7. Confusion matrix for average category features.

OUTPUT:

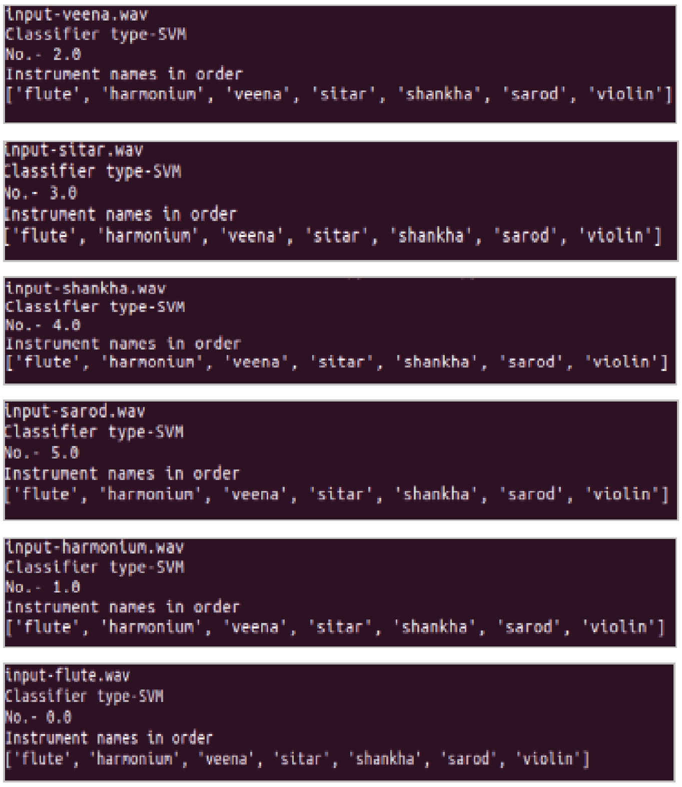

Figure 4. Testing results

From the above observation we can see that, when we use the features from excellent and average category instead of 34 features, the efficiency of SVM decreases from 97.2% to 93% which is much better result compared to other papers[2].

REFERENCES

[1] S.Essid, G.Richard and B.David,“Musical instrument recognition by pairwise classification strategies,” IEEE Trans. Audio, Speech, Lang. Process., vol. 14, no. 4, pp. 1401–1412, Jul. 2006.

[2] Jayme Garcia Arnal Barbedo and George Tzanetakis, “Musical Instrument Classification Using Individual Partials” Member, IEEE Transactions on audio, speech, and language processing, vol. 19, no. 1, January 2011.

[3] A. Eronen, “Automatic musical instrument recognition,” Ph.D. dissertation, Tampere Univ. of Technol., Tampere, Finland, 2001

[4] Giulio Agostini, Maurizio Longari and Emanuele Pollastri, “Musical Instrument Timbres Classification with Spectral Features” EURASIP Journal on Applied Signal Processing 2003.

[5] Satish R.Sankaye , Suresh C.Mehrotra and U.S. Tandon, “Indian Musical Instrument Recognition using Modified LPC Features”, International Journal of Computer Applications (0975 – 8887) Volume 122 – No.13, July 2015

[6] Surendra Shetty and Sarika hedge, “Clustering of Instruments in Carnatic Music for Content Based Information Retrieval” Advanced Computing (IACC), 2016 IEEE 6th International Conference on 27-28 Feb. 2016

[7] G. Agostini, M. Longari, and E. Pollastri, “Musical instrument timbres classification with spectral features,” *EURASIP J. Appl. Signal Process.*, vol. 2003, pp. 5–14, 2003.

[8] E. Benetos, M. Kotti, and C. Kotropoulos, “Musical instrument classification using non-negative matrix factorization algorithms and subset feature selection,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2006, pp. 221–224.

[9] J. J. Burred, A. Robel, and T. Sikora, “Polyphonic musical instrument recognition based on a dynamic model of the spectral envelope,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, 2009, pp. 173–176.

[10] J. C. Brown, “Computer identification of musical instruments using pattern recognition with cepstral coefficients as features,” *J. Acoust. Soc. Amer.*, vol. 105, pp. 1933–1941, 1999.

[11] A. Eronen, “Comparison of features for music instrument recognition,” in *Proc. IEEE Workshop Applicat. Signal Process Audio Acoust.*, 2001, pp. 19–22

[12] S. V. Chandan, M. R. Naik, Ashwini and A. V. Krishna, "Indian Instrument Identification from Polyphonic Audio using KNN Classifier," 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2019, pp. 135-139.