Indian Musical Instruments Recognition Using LPC and LPI

Satish R.Sankaye[1],U.S.Tandon[2], S.C.Mehrotra[3]

Department of CS &IT, Dr.Babasaheb Ambedkar Marathwada University, Aurangabad, India[1]

Department of Physics, College of Natural and Computational Sciences Haramaya University, Dire Dawa, Ethiopia[2]

Ramanujan Chair Professor, Dept. of C.S.& IT, Dr.Babasaheb Ambedkar Marathwada University, Aurangabad India[3]

Abstract: Most Significant Music Industry is today widely supported by Research in the area of Musical information Processing. In the past decade the growth of Digital Music over internet and our day-to-day life has forced researcher to explorethe diverse area of Musical Instrument Identification, Retrieval, Classification and Recognition of large sets of music. In our paper we have explored the ignored area of Indian Musical Instrument Recognition (IMIR) and attempted to build the database for the same. An innovative technique to classify the IMIR using the Modified Linear Predictor Coefficient (LPC) features and Log-Power-Intensity (LPI) is discussed. The proposed method has been tested with experimental results and technique has an accuracy of 96.11% for instrument and 92.45 for individual sound.

Keywords: Indian Musical Instrument Recognition, Linear Predictor Coefficient, Log Power Intensity, Music Instrument Identification, Principle component Analysis, Best First Decision Tree Algorithm

I. INTRODUCTION

Music is said to be universal language, as it is not restricted by caste, community or country neither by cultural nor by civilization. Music is a form of knowledge that links us more closely to the reverberation of the surrounding world. Music information processing [1] [2] has been widely deployed in music Industry and study related towards the sound synthesis on electronic music synthesizers [3] [4], desktop music production based on MP4, Instrument Recognition [5] [6] [7], Musical Information Retrieval [8] [9] are the major zone. The current focus of our research areas is towards the new technologies that target the direct identification of Indian Musical Instrument for those who are not professionals or musicians. The research will be beneficial for the development of musical tutor apps as well as Instrument Tuning applications.

The two most crucial components required to perform the music are the player and the instrument. The note has a dual nature. It is both an independent entity as well as an instrument control signal. The instrument problem refers to the challenge of modeling sound-producing resources (instruments) and their relationship to sound instances (notes).

The paper is planned as follows: Section 2 describes the relative works in the area of musical information processing while section 3 confers about the database initiated by us. Features related to musical instrument recognition and classification techniques are presented in section 4. Section 5 provides an overview of experimental results. Concluding the paper with future work is discussed in Section 6.

II. RESEARCH REVIEW

The universal growth of digitized musical signal has made an impact even in the research area. Most of the researchers have contributed their work in the area of Music Instrument Recognition (MIR) using set of different features and many classification techniques. Some of the prominent research works are discussed below:

Antti Eronen and AnssiKlapuri [5] presented a system for musical instrument recognition that used a wide set of features to model the temporal and spectral characteristics of sounds. The usefulness of the features was validated using test data that consisted of 1498 samples covering the full pitch ranges of 30 orchestral instruments from the string, brass and woodwind families, played with different techniques. The correct instrument family was recognized with 94% accuracy and individual instruments in 80% of cases. They focused on the autocorrelation sequence and then used LPC coefficient calculation with Levinson-Durbin algorithm for instrument identification.

Pierre Leveau, Emmanuel Vincent and et.al [10], proposed a new mid-level representation of music signals that incorporated explicit instrument labels and intends to provide a single front-end for many information retrieval and signal processing tasks. The signal was decomposed into a small number of sound atoms or molecules, where each atom was a sum of windowed harmonic sinusoidal partials and each molecule was a group of atoms spanning successive time windows. They designed an efficient algorithm to extract the most prominent atoms or molecules and investigate several applications of this representation, including polyphonic instrument recognition and music visualization. In this paper, the task of automatic musical instrument recognition has been thoroughly executed using the potential of sound atoms mid-level representations. On monophonic excerpts, the proposed algorithms obtained nearly equivalent performances as that of a standard MFCC feature-based approach.

Jun Wu and et.al, [11], discussed an algorithms based on joint modeling of sustained and attack sound for

polyphonic pitch estimation and instrument identification. The proposed algorithm was based on a spectro-temporal Gaussian mixture model of the harmonic partials and the inharmonic attack of each note, whose parameters are later estimated by the Expectation-Maximization algorithm. These parameters were then subjected to a logarithmic transformation and to Principal Component Analysis so as to obtain a low-dimension timbre feature vector. Finally, Support Vector Machine classifiers were trained from the extracted features and used for musical instrument recognition. For the purpose of training, isolated notes of 9 instruments 60 signals of 6 sec duration were collected from the RWC database, McGill University Master Samples CD Library and UIowa database while the testing was performed on synthetic mixtures of isolated notes and on real-world data. The proposed algorithm has outperformed other algorithms based on harmonic modeling alone for multiple pitch estimation and as well as for instrument identification.

Shelke, Akshata, and Abhijit Chitre [12], contributed their work in the area of Indian Musical instrument recognition. They concentrated on pitch, loudness, duration, spatialization and timbre features produced by the Musical sound. The system proposed has been trained and tested with three Indian musical instruments viz flute, harmonium and sitar, which are most commonly used in Indian classical music. As per their experimental result Support Vector Machine (SVM) classifier proved to be an accurate classification technique for Indian Musical Instrument. Using separately recorded notes as test sets, they were able to achieve average accuracy as high as 88.88 % for SVM to decide if a note was played by the sitar or others.

AthanasiaZlatintsi and Petros Maragos [13], proposed the use of a multiscale fractal feature for structured analysis of musical instrument tones. The multiscale fractal dimensional features were successfully used to discriminate several aspects of different music instruments. They performed the experiments using 1331 notes of 7 different instruments viz., Double Bass, Bassoon, Cello, Bb Clarinet, Flute, Horn and Tuba, while the analysis was performed in 30ms frames with 15 ms overlap. The result showed that the MFDs improved the recognition accuracy when fused with the MFCCs, accomplishing an error reduction up to 32%.

ArieLivshin and Xavier Rodet [14] presented a work related to compilation of musical instrument sample databases for elimination of badly recorded samples and validation of sample classification into correct categories. They introduced an algorithm for automatic removal of bad instrument samples and outlier detection techniques using Multiclass Interquantile Range (MCIQR) method. Best evaluation results on a methodically contaminated sound database are achieved using the introduced MCIQR method, which removes 70.1% "bad" samples with 0.9% false-alarm rate and 90.4% with 8.8% false-alarm rate.

III.DATABASEDESIGN

With an intention to give impetus to the research related to Indian Musical Instrument Recognition, we are constructing a copyright-cleared music database. This will help the researcher use this database and also enable them to compare and evaluate various systems and methods against a common standard. The Musical Instrument Recognition leads to the path of initially recording the sound sample from the different sources. It may be recorded directly while playing the instrument using any electronic gadget meant for sound recording in natural environment. Also for the study purpose the musical instrument may be recorded in an anechoic room at a professional studio too. To build a database for the Indian Musical Instrument for the research is considered to be a challenging job. We initialized the task by first finalizing the list of Indian musical instrument as well as the musicians who extended their assistance for sound recordings. For the purpose of this study several recording sessions were performed by different performers for different instruments. The main criterion while setting instruments and microphones was the naturalness. Fig.1, presents a block diagram of the main components of the system implemented for recording.
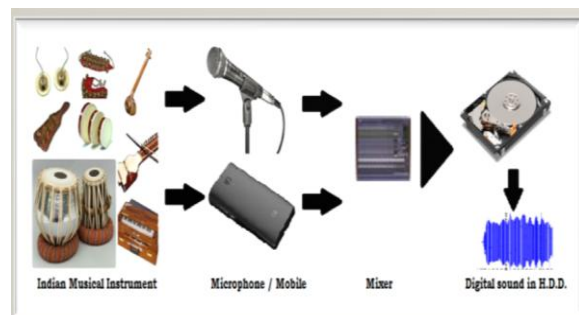

Fig. 1. Block Diagram of Our Recording System Used.

The Recording was performed in studio and alsoin the open. The recording wasconducted in an anechoic room at "Silver Oak Studio" 23, Aditya Nagar, Aurangabad. The microphone used was Dynamic type make SM-58 Shure (The Shure SM58 is a unidirectional (cardioid) dynamic vocal microphone). The distance between the microphone capsules and instruments was kept at 15-20 cm. Sounds were recorded using the Channel Mixer Yamaha MG Series and MAudio Sound Card. Before the recording started, instrumentalists tuned their instrument according to the frequency of 440 Hz, and then played sounds from the whole chromatic scale using differentiated articulation and dynamics. The samples were also recorded during the time of instrument tuning and as well as the performance for the instruments.

In addition to the recording in the studio, we also conducted the recording with the help of recorder either with the help of Computer or with the help of mobile in open as directly played by performer to obtain the genuineness in sound samples. Although most of the Indian Musical Instruments could be arranged, there were practical constraint of the simultaneous availability of instrument availabilities and the artistplaying them. We recorded by playing the instruments belonging to all the four groups, namely: *AvanaddhaVadya* (Percussion), *Ghana Vadya* (Idiophone), *SushirVadya* (Aerophones) and *TantuVadya* (Chordophones) as listed in the table 1.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

TABLE I: LIST OF MUSICAL INSTRUMENT IN OUR DATABASE

Sr. No.	Database of Musical Instrument		
	Family of the Instrument	Name of instrument	No. of Sound excerpt
1.	Avanaddha Vadya	Dhaf	92
2.		Dholki	107
3.		Tabla(Dhin,Ghae ,Ke, Ta,Tee& Tit)	814
4.		Pakhawaj	91
5.	Ghana Vadya	Ghatam	107
6.		Ghungaroo	82
7.		Manjira	65
8.		Morsing	75
9.		Triangle	102
10.	Sushira Vadya	Bansuri / Flute	130
11.		Harmonium	65
12.		Shehnai	65
13.		Sambal	50
14.	Tantu Vadya	Tanpura	65
15.		Clarinet	50
16.		Mandolin	50
17.		Sarangi	60
18.		Santoor	60
19.		Sarod	60
20.		Sitar	60

Recorded sounds were transferred through Channel Audio Mixer via the sound recording software Cool Edit Pro into the computer hard disc as digital format. The next step was the preparation of the database recorded sound. The sound samples were recorded at sampling rate of 44100Hz to get CD quality, 16-bit, stereo type PCM WAV file format for all samples.

IV. FEATURES OF MUSICAL INSTRUMENTS

In this research paper we proposethe Indian Musical Instrument Recognition (IMIR) technique using the Modified Linear Predictor Coefficient [15] and Log Power Intensity Feature vector with the help of Principle Component Analysis Classification Method.The Details about the feature vectors are as discussed below:

A. Linear Predictor Coefficient:

Linear Predictor Coeffficient (LPC) [16] is generally used for speech analysis and re-synthesis. It transmits spectral envelope information and is used to construct vocoders where musical instruments are used as excitation signal to the time-varying filter estimated from a singer's speech [17]. LPC uses the Levinson-Durbin recursion to solve the normal equations that arise from the least-squares formulation. This computation of the linear prediction coefficients is often referred to as the autocorrelation method.We tried to use the LPC feature vector. The Waveform of sound samples is collected in the feature Vector Xi. The feature Vector Xi has been smoothened by computing an LPC feature set, A = [1 A(2) . …. A(N+1)] of Nth- order linear predictor using,

$$n = [2+fs/10000] \approx 6.4 = 7$$

The digital filter method has been used to complete the filter response by using the following equation:

$$H(e^{jw}) = \frac{1}{A(e^{jw})} = \frac{1}{a(1) + a(2)e^{-jw} + \ldots + a(n+1)e^{-jnw}}$$

The Modified LPC Feature vector has been converted to the logarithmic form as per the human perception and it was normalized. For the Final Feature set, we discarded the values close to 0 which werenot effective,later we selected the odd features from the remaining P_i, minimizing the number of feature vector to the final size of P_i where i= 1 to 15.

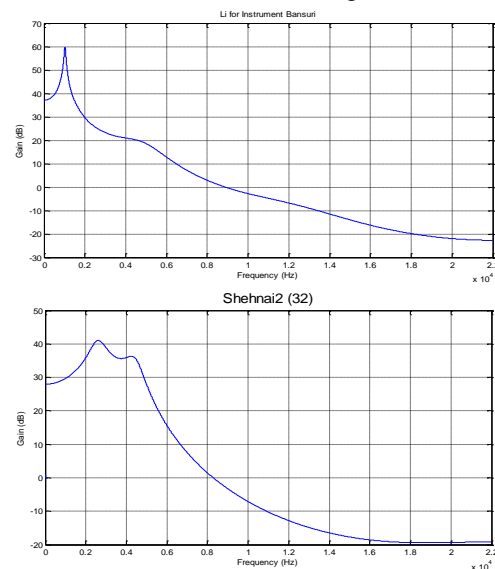The Modified LPC feature vector for the instrument *Bansuri* and *Shehnai* is as shown in fig 2.

Fig. 2Modified LPC Feature Vector for *Bansuri* and *Shehnai*

B. Log Power Intensity:

The way of expressing sound wave amplitude is referred as sound power Intensity (PI).It is represented in decibel rather than watts/meter^2, primarily because the ears doesn't hear linearly. As the decibel scale is nearly linked with our ear's perception of loudness, sound intensity level can becomes one of another important feature to be studied for musical instrument study too. The Log Power intensity (LPI) level is computed using the formula given below:

$$\text{intensity Level (LPI)} = 10 \log (I_2 / I_1),$$

Where LPI = the number of decibels by which I_2 is greater than I_1. I_2 and I_1 are the highest and lowest sound power intensities being compared.

After studying these feature set,different model are being implemented on the feature set for the Identification of the musical instrument or classifying the excerpt as a member of particular family.

C. Principle Component Analysis:

Principle Component Analysis (PCA) [18][19][20] is a useful statistical technique that has found application in fields such as face recognition, handprint recognition, human-made object recognition, industrial robotics, mobile robotics, speech recognition, music information retrieval and image compression. It is a representative of

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 12, December 2015

the unsupervised learning method which yields the linear projection.

Let $T_x = \{x_1, x_2, x_3, \ldots, x_l\}$ be a set of training vectors from the n-dimensional input space Rn. The set of vectors $T_z = \{z_1, z_2, z_3, \ldots, z_l\}$ is a lower dimensional representation of the input training vectors T_X in the m-dimensional space R_m.

The vectors T_Z are obtained by the linear orthonormal projection

$$z = W^T x + b \qquad (1)$$

where the matrix $\mathbf{W}$ [$n \times m$] and the vector $\boldsymbol{b}$ [$m \times 1$] are parameters of the projection. The reconstructed vectors $T_X = \{\mathbf{x'}_1, \mathbf{x'}_2, \ldots, \mathbf{x'}_1\}$ are computed by the linear backprojection

$$x' = W(z - b) \qquad (2)$$

obtained by inverting (1). The mean square reconstruction error is computed as

$$\varepsilon_{MS}(\mathbf{W}, \boldsymbol{b}) = \frac{1}{l}\sum_{i=1}^{l}\|x_i - \bar{x}_i\|^2, \qquad (3)$$

is a function of the parameters of the linear projections (1) and (2). The Principal Component Analysis (PCA) is the linear orthonormal projection (1) which allows for the minimal mean square reconstruction error (3) of the training data TX. The parameters (W, b) of the linear projection are the solution of the optimization task

$$(\mathbf{W}, \boldsymbol{b}) = \operatorname*{argmin}_{W', b'} \varepsilon_{MS}(\mathbf{W'}, \boldsymbol{b'}), \qquad (4)$$

subject to

$$(\mathbf{w}_i \cdot \mathbf{w}_j) = \delta(i, j), \forall i, j,$$

where w_i, $i = 1, \ldots, m$ are column vectors of the matrix $\mathbf{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ and $\delta(i, j)$ is the Kronecker delta function. The solution of the task (4) is the matrix $\mathbf{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m]$ containing the m eigenvectors of the sample covariance matrix which have the largest eigen values. The vector $\boldsymbol{b}$ equals to $\mathbf{W}^T\boldsymbol{\mu}$, where μ is the sample mean of the training data.

V. RESULT

The proposed method is based on above 15 Modified LPC Feature Vector and 9 Log Power Intensity Vector. The recognition rate was obtained by 10-fold cross validation and PCA Classification method. Overall 412 files were extracted belonging to eight different instruments: *Benjo, Dhaf, Harmonium, Triangle, Tabla, Ghungaroo, Ghatam and Morsing*. The result has shown significant performance as 396 have been correctly classified whereas the number of incorrectly classified samples was only 16 as seen in fig. 3 and table II below:
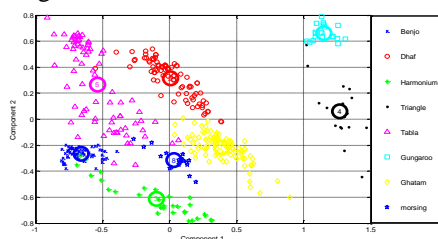

Fig.1. PRINCIPAL COMPONENT OF MUSICAL INSTRUMENT

TABLE III : CONFUSION MATRIX OF MUSICAL INSTRUMENT

a	b	c	d	e	f	g	h	
58	0	0	0	1	0	0	0	a
0	88	1	0	0	0	0	0	b
0	2	26	0	3	0	0	0	c
0	0	0	16	0	1	1	0	d
0	2	2	0	74	0	0	0	e
0	0	0	0	0	19	0	0	f
0	1	0	0	0	0	99	0	g
0	0	1	0	1	0	0	16	h

a) *Benjo*, b) *Dhaf*, c) *Harmonium*, d) *Triangle*, e) *Tabla*, f) *Ghungaroo*, g) *Ghatam*, h) *Morsing*

Using the Log Power Intensity, experimental results related to the different sound of *tabla* instrument for recognition of individual sound (bol) can be seen in fig 4 and Table III. The Experimental results has outperformed with the Log Power Intensity for the individual bols of *tabla* instruments. We have achieved the result of 92.45%.
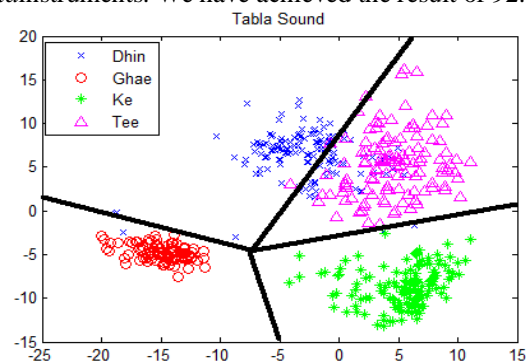

Fig. 4. Log Power Intensity Features Classification of *Tabla* Bols

VI. CONCLUSION

In this paper, we proposed exclusive features of Modified Linear Predictor Coefficient and Log Power Intensity for the recognition of Indian Musical Instrument as well as individual sound of the instruments (*bols*) signals. The result shows that the features are quite effective and it needs to be worked out on more number of musical instruments and number of samples. Future work will focus on increasing the number of instrument within the class as well as instruments belonging to other classes.

Dr. P.D. Deshmukh, Principal MGM **Dr.G.Y.Pathrikar** College of CS & IT, who has always motivated to complete the task.

REFERENCES

[1] GotoMasataka, et al, "RWC music database: Popular, Classical and Jazz MusicDatabases," in ISMIR Vol 2,2002 pp. 287–288.

[2] M. Goto, H. Hashiguchi, T. Nishimura, and R. Oka, "RWC music database: Music genre database and musical instrument sound database," in Proc. of the 4th Int.Conf. on Music Information Retrieval (ISMIR), 2003, pp. 229–230

[3] Starr, Harvey, "For use with a music synthesizer" U.S.Patent No.5,398585 Mar.1995.

[4] Olson, Harry F., and Herbert Belar, "Electronic music synthesizer" The journal of the Acoustical Society of America 27.3 (1955):595-612.

[5] Eronen, Antti, and AnssiKlapuri. "Musical instrument recognition using cepstral coefficients and temporal features." Acoustics, Speech, and Signal Processing, 2000. ICASSP'00. Proceedings. 2000 IEEE International Conference on. Vol. 2. IEEE, 2000

[6] Essid, Slim, Gaël Richard, and Bertrand David. "Instrument recognition in polyphonic music based on automatic taxonomies." Audio, Speech, and Language Processing, IEEE Transactions on 14.1 (2006): 68-80.

[7] Eggink, Jana, and Guy J. Brown. "Instrument recognition in accompanied sonatas and concertos", Accoustic, Speech, and Signal Processing, 2004. Proceeding, (ICASSP'04). IEEE International Conference on Vol. 4. IEEE, 2004.

[8] Ghias, Asif, et al., "Query by humming: musical information retrieval in an audio database." Proceeding of the third ACM international conference on Multimedia. ACM, 1995.

[9] Lemstrom, Kjell and Pauli Laine, "Musical information retrieval using musical parameters.' Proceeding of the 1998 international Computer Music Conference,1998.

[10] Leveau, Pierre, et al., "Instrument-specific harmonic atoms for mid-level music representation", Audio, Speech, and Language Processing, IEEE Transactions on 16.1.2008: 116-128.

[11] Wu, Jun, et al., "Polyphonic pitch estimation and instrument identification by joint modeling of sustained and attack sounds". Selected Topics in Signal Processing, IEEE Journal of 5.6(2011): 1124-1132.

[12] Shelke, Akshata, and Abhijit Chitre. "An Effective Feature Calculation For Analysis & Classification of Indian Musical Instruments Using Timbre Measurement." Proceedings of the Sixth International Conference on Computer and Communication Technology 2015. ACM, 2015.

[13] AthanasiaZlatintsi and Petros Maragos "Multiscale Fractal Analysis of Musical Instrument Signals With Application to Recognition". Audio, Speech, and Language Processing, IEEE Transactions on 21.4.(2013):737-748.

[14] ArieLivshin and Xavier Rodet,"Purging Musical Instrument Sample Databases Using Automatic Musical Instrument Recognition Methods".Audio, Speech, and Language Processing, IEEE Transactions,Vol. 17, No. 5, July 2009.

[15] Satish RamlingSankaye, S.C. Mehrotra, U.S. Tandon, "Indian Musical Instrument Recognition Using Modified LPC Feature", International Journal of Computer Applications (0975–8887) Volume 122, No 13, 6-10, July 2015.

[16] http://en.wikipedia.org/wiki/Linear_predictive_coding

[17] Kim, Youngmoo E., and Brian Whitman. "Singer identification in popular music recordings using voice coding features." *Proceedings of the 3rd International Conference on Music Information Retrieval*. Vol. 13. 2002.

[18] Aleix M. Martinez and Avinash C. Kak. "PCA versus LDA", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, N0. 2. February 2001.

[19] Delac, Kresimir, MislavGrgic, and Sonja Grgic. "Independent comparative study of PCA, ICA, and LDA on the FERET data set." International Journal of Imaging Systems and Technology 15.5 (2005): 252-260.

[20] Toygar, Önsen, and Adnan Acan. "Face recognition using PCA, LDA and ICA approaches on colored images." Journal Of Electrical & Electronics Engineering 3.1 (2003): 735-743.

[21] Anissa Bouzalmat and et,al. "Comparative Study of PCA, ICA, LDA using SVM Classifier", Journal of Emerging Technologies in Web Intelligence, vol 6, No.1 , February 2014