An Effective Feature Calculation For Analysis & Classification Of Indian Musical Instruments Using Timbre Measurement

Akshata Shelke
Student, Dept. of Electronics and Telecommunication Engineering,
Vishwakarma Institute of Information Technology, Pune, India.
akshata.shelke@gmail.com

Abhijit Chitre
Associate Professor, Dept. of Electronics and Telecommunication Engineering
Vishwakarma Institute of Information Technology, Pune, India.
chitreabhijit77@gmail.com

ABSTRACT
Musical instrument recognition is significant field in the research of computer music which is related to the modelling of sounds. Analysing & synthesing the structure of musical note is of importance both for modelling music signals and their automatic computer-based recognition.

Musical sound is produce by five dimensions: pitch, loudness, duration, spatialization, and timbre. First four parameter can be controlled but timbre remains difficult. Timbre then naturally became the main subject of this work. It is important property of sound that separate one music instrument from another and independent of pitch and volume. This work presents a system for identifying a specific musical instrument from monophonic recordings .The system proposed in this paper has been trained and tested with three Indian musical instruments samples. Instruments include flute, harmonium and sitar, which are most commonly used in Indian classical music.The Statistical and spectral parameter are used for the classification of the sounds in Indian Musical Instruments. The SVM classifier proves its ability in automatic and accurate classification of Indian Musical Instrument. Using separately recorded notes as test sets, we were able to achieve average accuracy as high as 88.88 % for SVM that deciding if a note is played by the sitar or others**.**

Keywords
Musical signal processing; Musical Instrument Identification ;Timbre; timbre time-domain; timbre spectral domain ; Support Vector Machine (SVM).

1. INTRODUCTION
Most of the musical instrument identification systems that are available are operated on isolated notes of western musical instruments. But significant work is not yet done on Indian musical instruments. There is need of extensive study of Indian musical instruments for various applications in different areas.

The recognition of sound sample automatically is forthcoming research area as it may helpful to monitor large fields, and identify Musical Instrument automatically based on their features. Musical instrument identification is a subtask which capably play as key feature of Multimedia content analysis systems. Musical instrument identification is an important attribute of music information retrieval. Knowing the instruments implicated in a given musical piece very useful information.

Human ears possess an ability to distinguish musical colours. We can easily separate the sound of piano from the sound of guitar because they have different feeling or colour in their sound. When required training is processed, human ears can extract a certain instrument using the knowledge of timbre of the sound of the specific instrument. Various audio features are used in recent researches in order to achieve automatic timbre recognition system using machine learning algorithms. Recognizing audio features can be a challenging problem since it is a continuous task unlike descrete data, image or text recognition. Acquisition and recognition need to be in sync of time frame in order to achieve correct predictions.[2] In order to achieve a correct set up for the recognition, we need to explore more about the musical charactersitics, timber.

The Musical sound is produce by five dimensions: pitch, loudness, duration, spatialization, and timbre. First four parameter can be controlled but timbre remains difficult. Timbre then naturally became the main subject of this work. Two approaches were tested to understand the dimensions of timbre, the first by examining the physical gestures associated with playing an instrument and the other by looking at the perception and psychoacoustic literature. This can be seen as a global approach, encompassing both the performer of a musical instrument and the auditor of the sounds produced. The conclusions of the two approaches were then used in the analysis and modelling of musical instrument sounds.

Timbre is also known as color or tone of sounds, defined as the characteristic feature of a sound that distinguish one music instrument from another & independent of pitch and volume.

Timbre is mainly determined by the vocal content of a sound and the vibrant characteristics of the sound such as vibrato and ADSR envelope of the sound .it is important factor for classifying Music.

Due to many attributes of sound are required to recognize timbre, it is necessary to acquire features that are objective rather than subjective.

ICCCT '15, September 25-27, 2015, Allahabad, India
© 2015 ACM. ISBN 978-1-4503-3552-2/15/09…$15.00
DOI: http://dx.doi.org/10.1145/2818567.2818586

The following five acoustic parameters to be considered related to timbre.

1) The range between tonal (usually pitched) and noise like character
2) The spectral envelope
3) The Amplitude envelope in terms of rise, duration, and decay (ADSR-attack, decay, sustain, release)
4) The small changes both of spectral envelope and fundamental frequency .
5) The prefix, or onset of a sound.

In this study, we propose a novel framework that focuses on the recognition of Indian Musical Instrument (Flute,Harmonium,Sitar). Through this work, Indian Musical Instrument are classified based upon the extracted features. The SVM classifier proves its ability in automatic and accurate classification of Indian Musical Instrument. Finally, it can be concluded from the experimental results that this approach can classify the Sound sample with an average accuracy of 88.88%. The purpose of this research work is to provide distinguish sounds from mixture of Musical instrument.

The proposed detection technique involves four phases which are Envelope analysis, Statistical analysis,extraction of timbre feature and classification. The challenge for the musical instrument identification systems are Spectrum analysis of sample Instruments, capture their silent properties of sound and defining our Features for Timbre Recognition are also addressed through the proposed work.

The structure of the paper is organized as follows: Section 2 describes the literature survey. The proposed approach is illustrated in section 3. Feature Extraction describe in section 4. Section 5 describes the details regarding classification. The experimentation and results are projected in section 6 and section 7 concludes the paper.

2. LITERATURE SURVEY

Extensive research has been conducted to explore various methods for automatic recognition of musical instrument . The research done on Western Musical instrument such as piano, clarinet, Trumpet, Guitar, Violin. As stated before, this work concentrates particularly recognizing on Indian Musical Instruments such as Flute, Harmonium, Sitar.

Swe Zin Kalayar Khine, et all [10] discussed framework that describe recognition of singer based on timbre quality.The proposed work is composed of vocal detection process by means of extracting vocal parameter from each song and cepstral coefficient both feature are important for timbre recognition of singers. This parameter are classified using GMM. The classification showed an typical error rate of 12.2%for Singer detection. The authors Poli[11], developed an signal processing based method of spectral envelope of MFCC features that useful to identify singers. The accuracy of the algorithm was tested by comparing the manually detected sound sample with those detected automatically.In [18], timber is characterized by the Dynamic characteristics such as harmonic lines of the harmonic sound.

The study by Daniel Mintz [5], presents a novel approach to sound synthesis based on timbre. The proposed method describing the timbre quality is greatly affected by spectral shape, amplitude envelope, and spectra-temporal variation, and the timbral separation converts these parameter values into direct envelopes. The thesis consists of a complete theory behind timbre Analysis & Synthesis. The James Beauchamp.[4]presents a way of timbre manipulation. The algorithm uses frequency tracking analysis to decompose existing sounds' spectra. the proposed method provide tool used for modification of sound sample before they are resynthesized

Juan José Burred *et al.* [1], discussed a framework for representing timbre model which was important for classification of instrument sample. The proposed work is composed of spectro-temporal envelope with unknown sample PCA and used GMM as classifier.The analysis tasks includes classification of instrument samples and detection of instruments in monaural polyphonic mixtures**.**This algorithm provide a classification accuracy of 94.9% with a database of five classes. The algorithm enables accurate modelling of timbre with multiple descriptors.. The features extracted are spectral ,temporal,specro-temporal and statistical parameter.

Sang Hyun Park et al.[2] describes an approach to identify Musical Instrument Extraction through Timbre Classification.. The heart of any automatic musical classification or analysis system is the computing of feature vectors. Though different classifiers have been compared [4], the choice of features has a large effect to the recognition accuracy than the selected classifiers. Even if artificial neural networks classifiers give satisfactory scores Many different sets of parameters have been proposed so far. A large number of them are mainly originating from speech recognition or analysis area. There are various features that can be used to characterize audio signals. They can be divided generally into spectral and statistical domain.The algorithm uses a machine learning approach is presented to extract a musical instrument from a complex music using timbre classification such as KNN that provide 75.12% accuracy.

Marques et al. [8], introduced a method for robust and early detection of musical instrument. the proposed method uses different feature and SVM & GMM classifiers used. The algorithm uses large database of isolated notes . The paper by Slim Essid et al.[17] presents prototype strategy for detection of Instrument with much higher databases for both training and testing and used PCA to de-noise the feature space. The accuracy for recognition rates obtained range from 52.0 % for Clarinet up to 96.0 % for Oboe.

Based on the above mentioned work and by the results obtained from the previous work, we recommend a new approach for recognition and classification of Musical instrument that can overcome the disadvantages of the previous work. Our proposed work can be applied to India musical Instrument (Flute, Harmonium, Sitar).

.

3. PROPOSED TECHNIQUE

In this section, we consider the general flow of the various steps that are being performed in order to get the preferred result..The proposed approach consists of four main steps: Envelope analysis of sound sample, extraction of Timbre parameter(spectral), statistical analysis and classification. Figure 1 shows the general block diagram of the Musical note Identification system.

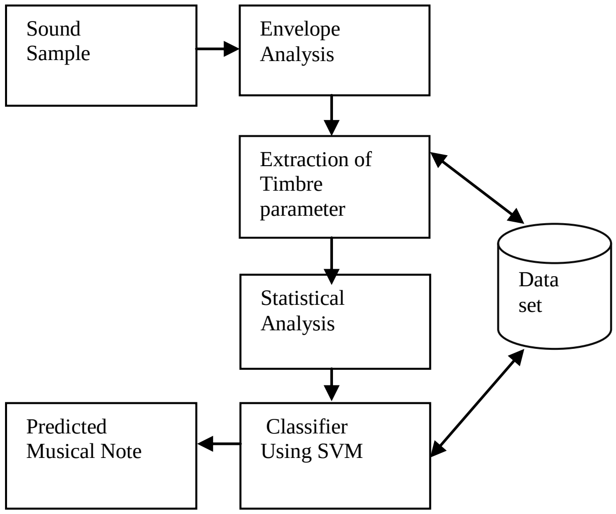


Figure 1 General block diagram of Musical note Identification

In the proposed approach, the first step is collection of audio database of Indian musical instrument (Flute,Harmonium,Sitar) with sampling frequency 44.1KHz with 3 second duration. A total of 150 sound sample are included in the database which comprises of Flute, Harmonium and Sitar & normalized at (16-bit) frequency.

In the second step, a Timbre recognition is done using two ways: First is Amplitude envelope (envelope analysis) is amplitude structure of sound. In which special case of envelope is used known as ADSR. Second is Spectrum envelope (Timbre features) such as ZCR ,periodicity, Amplitude, Spectral centroid, Spectral roll-off, Spectral flux. This spectrum parameter alone is often enough to classify or recognize sound.

In the last step, statistical analysis of the features is done and support vector machine (SVM) is used for final classification which classifies the Indian musical instrument in one of the two classes as woodwind(flute & harmonium) & string based(sitar).

4.FEATURE EXTRACTION

Feature extraction provides a compact numerical representation of the musical pieces. We have chosen to represent each signal by a unique feature vector over the texture window. This vector will only contain timbral descriptors. The Timber is closely related to the instrumentation of the music performance. It is a perceptual attribute of sound that describes difference between two tones played in same pitch & volume. Most objective signal descriptors associated to timbre describe the spectral distribution of sound. We will use the following features to characterize timbral texture. These latter's can be divided into time-domain, spectral-domain,

A.Statistical features:

They consist of Mean, Variance, Standard deviation, Kurtosis, Average power and Average Magnitude. These features give general statistics of the given signal. Statistical analysis is carried out in time domain.

B.Timbre Time-domain:

These coefficients are called time-domain because they are computed using the time envelop of the signal.

4.1. Zero Crossings rate

It gives a measure of the noisiness of a signal. Zero crossings rate for musical signals is higher for musical signals than speech.

$$ZCR = \frac{1}{2}\sum_{n=1}^{N} |sign(x(n) - sign(x(n-1))| \qquad (1)$$

4.2. Periodicity

Periodicity is continuous function that representing the strength of periodicity(or tempo).Many method such as autocorrelation ,Fourier transform are used but most commonly used method is Autocorrelation. It is used to detect fundamental frequency ,phase or frequency. it is used for pitch tracking.

$$r(\Gamma) = \sum_{k=0}^{n} x(n)x(n+\Gamma) \qquad (2)$$

4.3 .Amplitude

Amplitude is root mean square energy of signal. It is computed as Intensity feature. It measure of loudness of signal

$$Erms=\sqrt{\frac{\sum_{n=0}^{N-1} E(n)^2}{N}} \qquad (3)$$

C. Timbre Spectral-domain:

4.4 Spectral Centroid

Spectral centroid is the centre of gravity of the magnitude spectrum of the Short Time Fourier Transform (STFT)." It measures the spectral brightness of a sound

$$Spectral\ Centroid = \frac{\sum_{n=0}^{N-1} f(n)\, s(n)}{\sum_{n=0}^{N-1} s(n)} \qquad (4)$$

Where f(n) is the frequency and s(n) is the power

4.5. Spectral roll-off

Spectral roll-off is another measure of spectral shape. It is defined as frequency bin F below which 85% of magnitude distribution is concentrated.

$$Spectral\ roll\text{-}off = 0.85 * \sum_{n=0}^{N} f(n) \qquad (5)$$

4.6. Spectral Flux

The spectral flux is defined as the squared difference between the normalized magnitudes of consecutive spectral distributions. It is a measure of local spectral changes in the signal.

$$Spectral\ flux = = \sum_{n=1}^{N} (N_t[n] - N_{t-1}[n])^2 \qquad (6)$$

5.CLASSIFICATION

The next step after extracting unique features from the sound sample classifying the Indian musical instrument as flute , harmonium and Sitar.. The feature used in this algorithm for classification are statistical & spectral parameters. The classifier used for this purpose is Support Vector Machine (SVM). SVM belong to a group of supervised learning methods which are normally used for classification and pattern recognition. Supervised learning is a machine learning algorithm that uses a known dataset i.e. the training dataset to make predictions for a new dataset i.e. the testing dataset. SVM performs classification tasks by constructing hyper planes in a high dimensional space which is used for classifying data points belonging to different classes. The accuracy of SVM classifier gets better with the increase in the number of samples in the training dataset.

The classification using SVM is done by finding the best hyper plane which distinguishes between the data points of two classes. The hyper plane having largest gap between two classes is best hyper plane for SVM. [21] .The proposed approach uses a simple binary SVM classifier.

Consider a supervised SVM two class classification. Consider that the training data is represented by $\{m_i, n_i\}$, where i = 1,2,….,N, and $n_i \in \{-1, +1\}$, where N is the number of training samples, n= +1 for class s_1 and n_i = -1 for class s_2.It is possible to find minimum one hyper plane which is defined by vector v and bias b, which can separate the classes accurately [20]. Thus, find v and b such that

$$n_i(v.m_i + b) > 0, i = 1,2, \ldots\ldots, N \quad (7)$$

If there is a hyper plane satisfying eq. (1), the set is known as linearly separable. Thus, it is always possible to rescale v and b such that

$$\min n_i(v.m_i + b) \geq 1, i = 1,2, \ldots\ldots, N$$

i.e. such that the closest point to the hyper plane has a distance of 1/||v||. Then, eq. (7) become

$$n_i(v.m_i + b) \geq 1 \quad (8)$$

Among the distinguishing hyper planes, the one for which the distance to the closest point is maximum is called optimal separating hyper plane (OSH). Since the distance to the closest point is 1/||v||, finding the OSH requires solving the following problem:

Minimize

$$\frac{1}{2} v.v \quad (9)$$

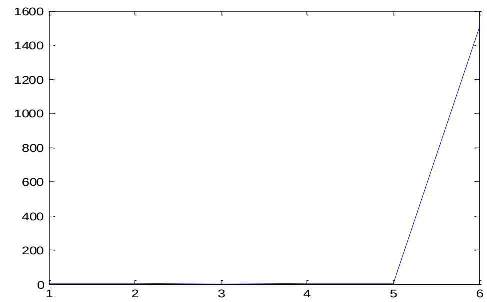
following constraints(8)

Thus, the aim of the proposed algorithm is to classify the Indian Musical Instrument using SVM binary classifier into two classes as sitar (string based) and Flute & Harmonium (woodwind based).One of the most important advantages for the SVM is that it guarantees generalization to some extent. The decision rules reflect the regularities of the training data rather than the in capabilities of the learning machine. Because of the many nice properties of SVM, it has been widely applied to virtually every research field.

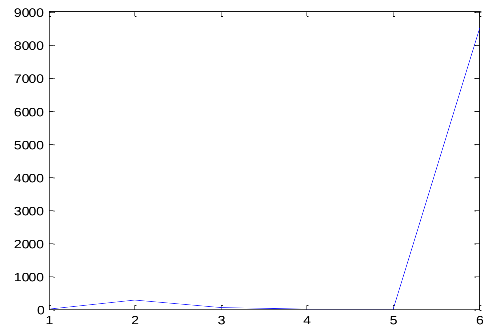
6.EXPEIMENTATION AND RESULTS

The proposed algorithm is tested over the database of 150 sound sample of Indian Musical Instrument . The database consists of sound sample of 50 for each Musical instrument such as Flute, Harmonium, Sitar. For testing 45 samples are taken separately. The database is generated from standard CD recordings. We have considered 3 second (.wav) wave files. Waves are sampled at 44.1 KHz (16-bit) frequency. Features are extracted from each sample and form feature vectors.

Figure 2 shows the Feature vector plot of Spectral parameter which represent each musical sample in classification stage



Feature vector plot of Flute



Feature vector plot of Harmonium

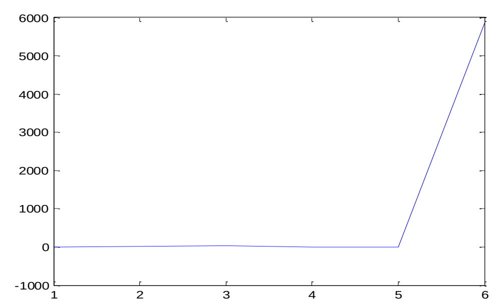


Figure. 2 Feature vector Plot of Sitar

Tabular representation illustrating the accuracy obtained by SVM classifier. The classification results show that SVM classifier yields an average accuracy of as high as 88.88 %. Thus, SVM yields accurate classification of Indian Musical Instrument into two classes as woodwind & string based gives 93.33 and the instrument can be further classified in woodwind (flute & Harmonium)gives 89.65 for statistical analysis.

TABLE 1. SVM Classification Accuracy

Type	Statistical Analysis	Type	Spectral Analysis
Sitar(Flute Harmonium)	93.33	Flute& (Harmonium, sitar)	68.88
Flute & Harmonium	89.65	Harmonium & Sitar	81.8

7.CONCLUSION

In the proposed work, an audio signal processing based approach is used for automated musical instrument classification based on timbre time domain ,spectral domain & statistical features. Our research work is concerned with the classification between Indian musical instrument (Sitar,Harmonium,Flute) using SVM classifier. We have tested our proposed system over the database of 150 sound sample taken directly from Indian Musical instrument. The SVM classifier can significantly support in recognizing Indian Musical Instrument with an average accuracy as high as 88.88%

The main aim of the proposed research is to provide Distinguishes musical instruments . This system will provide help to an application of detection of singers and an MIR based on timbre correspondence . The future work includes timbre synthesis which is versatile tools that could be useful for many composers and signal processing researchers. The proposed work can be extended to various Musical instrument .

8.REFERENCES

[1] Juan José Burred, Axel Röbel, and Thomas Sikora, "Dynamic spectral envelope modeling for timber analysis" *speech & language processing* ,vol 18 ,no 3 March 2010.

[2] Sang Hyun Park," Musical Instrument Extraction through Timbre Classification", NVIDIA Corporation Santa Clara, CA 95050

[3] Bartsch, M.A., and Wakefield, G.H. "Singing Voice Identification Using Spectral Envelope Estimation,."*IEEE Transactions, Speech and Audio Processing, VOL. 12, pp.100-109, March* 2004 .

[4] James W. Beauchamp. "The sound of music : analysis, synthesis, and perception of musical Sounds, pages 1{89. Springer, New York, 2007.

[5] Daniel Mintz "Toward Timbral Synthesis: a new method for synthesizing sound based on timbre description schemes" June 2007

[6] J. F. Schouten, "The perception of timbre," in Proc. 6th Int. Congr.Acoust., Tokyo, Japan, 1968, vol. GP-6-2, pp. 35–44.

[7] T. H. Park, "Towards Automatic Musical Instrument Timbre Recognition",Ph.D. Thesis, Princeton University, 2004.

[8] J. B. Allen, " Short term spectral analysis, synthesis and modification by discrete fourier transform." *IEEE Trans. On Acoustics, Speech and Signal Processing,* Vol. ASSP-25,NO-3 , June 1977.

[9] Marques and P. J. Moreno, "A study of musical instrument classification using gaussian mixtures models and support vector machines," *Compaq Cambridge Research Laboratory, vol.* Tech. Report 99-4, 1999.

[10] Swe Zin Kalayar Khine, Tin Lay New, Haizhou Li ,"On Timbre Based Perceptual Feature For Singer Identification'' *Institute for Infocomm Research 21 Heng Mui Keng Terrace, Singapore.*

[11] Poli, G.D., and Prandoni, P. .Sonological Models for Timber Characterization,. *Journal of New Music Research*, vol. 26, pp. 170-197

[12] Christianini, N., and J. Shawe-Taylor, "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods", Cambridge University Press, Cambridge, UK, 2000.

[13] Y. Stylianou, O. Cappe, E. Moulines, "Continuous Probabilistic Transform for Voice Conversion". *IEEE Transactions on Speech and Audio Processing* Vol. 6, No. 2, pp. 131-142 March 1998.

[14] B. Kostek, "Musical instrument classification and duet analysis emplozmusic information retrieval techniques," *Proc. IEEE*, vol. 92, no. 4, pp. 712–729, Apr. 2004.

[15] C.W. Hsu and C. J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 415–425,Mar. 2002.

[16] Hiromasa Fujihara, Masataka Goto, Tetsuro Kitahara, and Hiroshi G. Okuno,*"* A Modeling of Singing Voice Robust to Accompaniment Sounds and Its Application to Singer Identification and Vocal-Timbre-Similarity-Based Music Information Retrieval", *IEEE Transactions, Speech and Audio Processing,*Vol. 18, No. 3, March 2010

[17] J. Herre, O. Hellmuth, T. Kastner, and C. Ertel, "Amultiple feature model for musical similarity retrieval," in *Proc. 4th* Int. Conf. Music Inf. Retrieval (ISMIR 2003), 2003, pp. 217–218.

[18] Zhang, T., and Kuo, C.C.J." Content-Based Audio Classification and Retrieval for Data Parsing", Kluwer Academic, USA, 2001.

[19] E. Pampalk, "Computational models of music similarity and their application in music information retrieval," *Ph.D. dissertation, Universitat Wien*, Vienna, Austria, 2006

[20] *Slim Essid, Ga¨el Richard, Bertrand David."* Musical Instrument Recognition On Solo Performance", *GET - ENST (T´el´ecom Paris) 46*, rue Barrault - 75634 Paris Cedex 13 - France

[21] Panagiotis Giotis," Instrument Timbre Transformation using Gaussian Mixture Models", *Master's Thesis, Master in Sound and Music Computing Universitat Pompeu Fabra* P.O. Box 138 Roc Boronat Str., 08018, Barcelona, SPAIN

[22] X. Zhang, Z. W. Ras, "Analysis of Sound Features for Music TimbreRecognition", IEEE CS International Conference on Multimedia and *Ubiquitous Engineering* (MUE 2007). April 26-28, Seoul, Korea, pp. 3-8.

[23] D. G. Bhalke, C. B. Rama Rao and D. S. Bormane, " Hybridization of Fractional Fourier Transform and Acoustic Features for Musical Instrument Recognition," *International Journal of SignalProcessing, Image Processing and Pattern Recognition*, Vol.7, No. 1, 2014 .