INTERNATIONAL JOURNAL OF RESEARCHES IN BIOSCIENCES, AGRICULTURE & TECHNOLOGY
© VISHWASHANTI MULTIPURPOSE SOCIETY (Global Peace Multipurpose Society) R. No. MH-659/13(N)
www.vmsindia.org

MUSICAL INSTRUMENT RECOGNITION USING DISCRETE WAVELET TRANSFORM AND k-NN CLASSIFIER

Patankar A.S.

International Institute Of Information Technology, Pune (M.S) India
Email: *patankar.kargi@gmail.com*

Abstract: The musical instrument sound can be described by four factors: pitch, loudness, duration, and timbre. The timbre can be referred as a feature of an audio that allows us to differentiate two sounds that are of same pitch, loudness and duration. The musical instruments can be generally classified into three categories as wind instruments (e.g.flute), string instruments (e.g. violin) and rhythm instruments (e.g. drum).Variation in musical pitch recognition is primarily due to highly heritable differences in auditory functions not tested by conventional audiologic methods. In pattern recognition applications, finding compact and efficient feature set is important in overall problem solving.The aim of project is to identify and classify the musical instruments. The new prosposed wavelet coeffcient features found compact and efficient with existing traditional features.

Keywords: Musical, Instrument, Recognition, Transform, Classifier

Introduction:

Automatic sound source recognition plays an important role in developing automatic indexing and database retrieval applications. These applications have potential in saving the humans from time taking searches through huge amounts of digital audio material available today. For instance, it would be most useful if we could find sound samples that sound similar as a given sound example. Music content analysis in general has many practical applications, including e.g. structured coding, automatic musical signal annotation, and musician's tools. Automatic musical instrument recognition is a crucial subtask in solving these difficult problems, and may also provide useful information in other sound source recognition areas, such as speaker recognition. However, musical signal analysis has not been able to attain as much commercial interest as, for instance, speaker and speech recognition. This is because the topics around speech processing are more readily commercially applicable, although both areas are considered as being highly complicated. Through constructing computer systems that listen, we may also gain some new insights into human perception. Recently, music data analysis and retrieval has become an emerging research field in signal processing.

Due to tremendous applications of musical instruments, musical instrument recognition (MIR) has attracted the attention of various researchers. In addition to this, the advancement in digital signal processing and data mining techniques has led to intensive study on music signal analysis like , content-based music retrieval, music genre classification, duet analysis, Musical transcription , Musical Information retrieval and musical instrument detection and classification. Musical Instrument detection techniques have many potential applications such as detecting and analyzing solo passages, audio and video retrieval, music transcription, playlist generation, acoustic environment classification, video scene analysis and annotation, etc.

Fig.-1: Musical Instruments

Wavelet transformation is one of the most popular of the time-frequency transformations. The fundamental idea of wavelet transforms is that the transformation should allow only changes in time extension, but not shape. This is effected by choosing suitable basis functions that allow for this. Changes in the time extension are expected to confirm to the corresponding analysis frequency of the basis function. The wavelet analysis provides spectro-temporal information of the music signal. The wavelet packet analysis decomposes a signal into packets by simultaneously passing the signal through a low decomposition filter (LDF) and a high decomposition fillter (HDF) in a sequential tree like structure. There are a large number of mother wavelet filters that can be used for this purpose. In this experiment we considered upto fifth level decomposition of various mother wavelet .Mother wavelets are nothing but transformation functions. The name wavelet because they are baby waves(oscillatory functions).The name mother is because they form the basis for various transformation process. The LDF and HDF are intimately related as the HDF is calculated by passing the LDF through a quadratic mirror filter (QMF) .The QMF firstly reverses the coefficients of the LDF and then reverses the sign of every second coefficient. This produces a filter that passes high frequency components. Passing a signal through two of these fillters produces two packets: (1) the Approximation (from the LDF) and (2) the Detail(from the HDF). This is referred to as level 1 decomposition. The level 1 packets can then be passed through another pair of filters to produce a total of 4 packets (level 2 decomposition). This operation can be continued indefinitely, although after a certain point, which has to do with the signal sample length, the packets from one instrument become less distinguishable from that of other instruments, affecting classification accuracy.

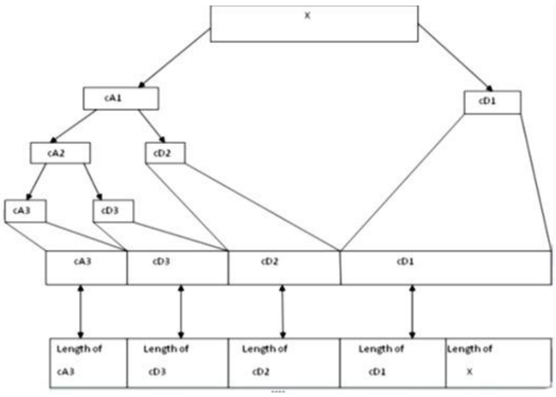

Fig.-2: Decomposition Tree Project idea

Block Diagram

The processes performed in the project are given in block diagram shown in Fig.3 are as follows

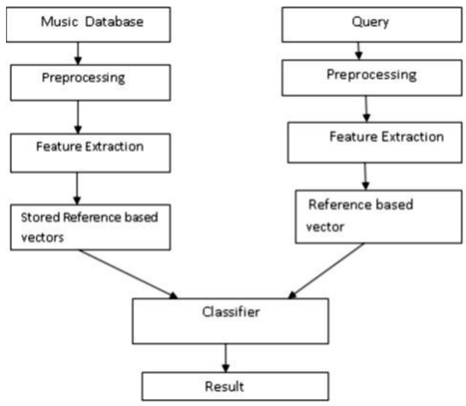

Fig.-3: Block Diagram

1. Database: Apply a data set of small ensemble recordings as an input.
2. Preprocessing: Remove noise present in the input.
3. Feature Extraction: Extraction of features takes place by using wavelet coefficient algorithm.
4. Reference feature vector: In this stage, vector matrix is generated for each applied input.

4.1.1 Testing

Input Query: Apply a testing input query.

The remaining processes like Preprocessing, Feature Extraction and Reference feature vector are carried out similar to the above mentioned training process.

4.1.2 Classification

The classification will be done by comparing the stored reference vectors with the input signal reference vector.

4.1.3 Result

It is the final stage, in which comparison is performed between the stored set of samples and the applied processed inputs. Finally, the playing instrument will be recognized.

Material and Method:

Project implementation

Wavelet Transform:

Recently wavelet transforms have found widespread use in various fields of music signal processing. By using Wavelet transform, it is possible to extract the desired time frequency components of a signal corresponding to music signals. After the wavelet decomposition some sub band signals corresponding music signals can be analysed. Each chosen sub band can be analysed in detail. In analysis we get different characteristics information regarding the particular instrument concentrated in a particular band. Hence we proposed Wavelet Packet Transform. As we said after the wavelet decomposition, some sub band signals can be analyzed, particular band can be representing to particular music instrument. Wavelet means a short wave. Wavelet transform are mostly used in many areas of signal processing and image processing like filtering of noisy data, compression, fingerprint compression, edge detection, medical electronics etc. The wavelet transform(WT) is a transform which provides a time frequency representation. It is capable of providing the time and frequency information simultaneously. Hence, it gives a time frequency representation of the signal. This transform is a new mathematical tool for local representation of non stationary signals. It involves mapping of signals to a time frequency joint representation. The temporal aspects of the signals are preserved. WT provides multiresolution analysis (MRA) with dilated windows. The high frequency analysis is done using narrow windows and low frequency analysis is done using wide windows. When we use frequencies are a narrow window, time resolution is better and the high frequencies are resolved in time domain. When the window is wide, the time resolution is poor, but frequency resolution is good and low frequencies are resolved in frequency domain.

There are two types of wavelet transform viz-
Continuous Wavelet Transform and
Discrete Wavelet Transform.

Discrete Wavelet Transform

Discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage it has over Fourier transforms is temporal resolution: it captures both frequency and location information (location in time).

DWT theory requires two sets of related functions called scaling function and wavelet function given by

$$\phi(t) = \sum_{n=0}^{N-1} h[n]\sqrt{2}\phi(2t-n)$$

and

$$\psi(t) = \sum_{n=0}^{N-1} g[n]\sqrt{2}\phi(2t-n),$$

h[n] =an impulse response of a low-pass filter

g[n] = an impulse response of a high-pass filter

h[n] and g[n]= quadrature mirror filters

Different 104 features which are shown in the Table- 1 are extracted and reference feature set is formed separately using temporal, spectral , cepstral , perceptual and wavelet sub band histogram based features . The results are tested using KNN classifier. The wavelet sub band histogaram based features shows good accuracy than all other features.

Table-1: Types of features

Feature No	Feature
1-39	MFCC, ΔMFCC, ΔΔMFCC
40-42	Mean SC, Std SC, variance SC
43-45	Mean SS, Std SS, variance SS
46-48	Mean SF ,Std SF,variance SF
49	LAT
50-52	Mean SK , STD SK , variance SK
53-100	Mean , STD , Variance of sub band of each histogram (Five level decomposition)
101	LAT
102-104	Mean ZCR, Std ZCR, variance ZCR

Result and Discussion:

Fig.4 shows the decomposition Flute E2 note in separate mode. Fig 4,5 and 6 shows the accuracy of musical instrument recognition using different mother wavelet. Db4, Db5 wavelets shows better accuracy. The choice of decomposition level is five because for five level decomposition we achieved better accuracy.

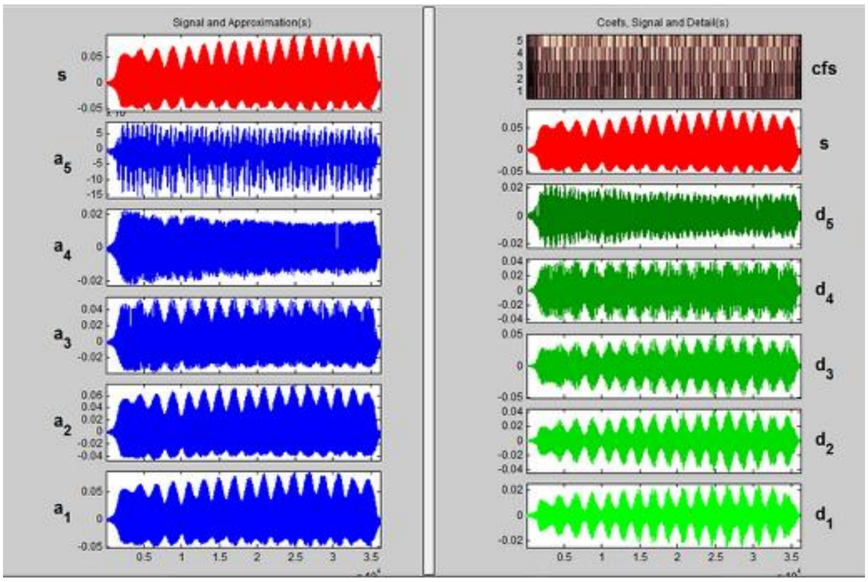

Fig.- 4: Separate Mode

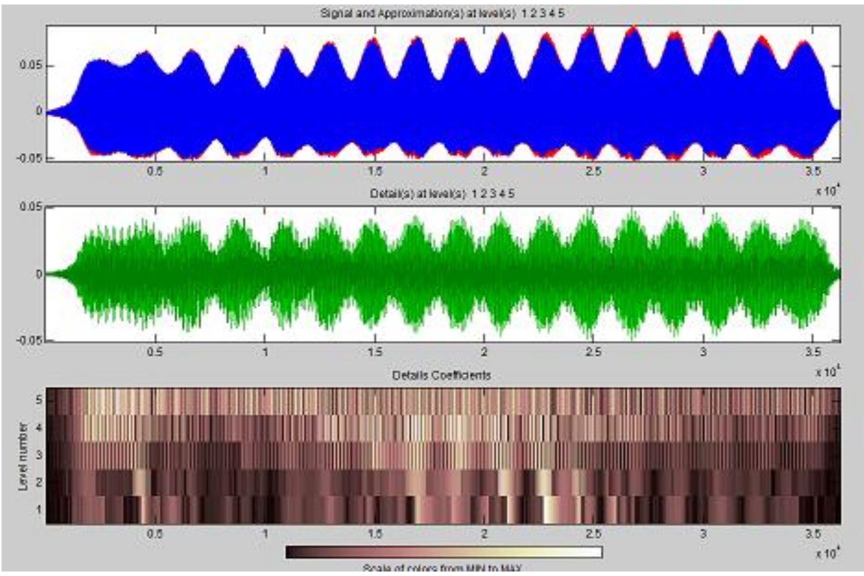

Fig.- 5:Superimpose Mode

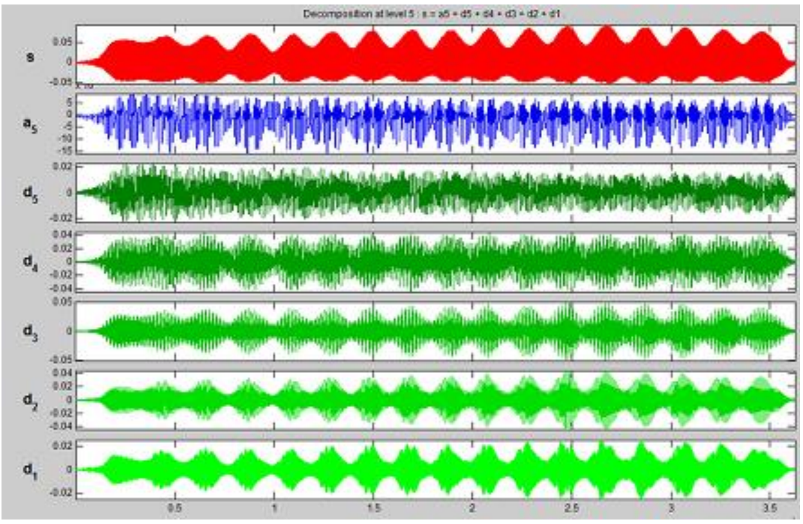

Fig. -6: Full Decomposition Mode

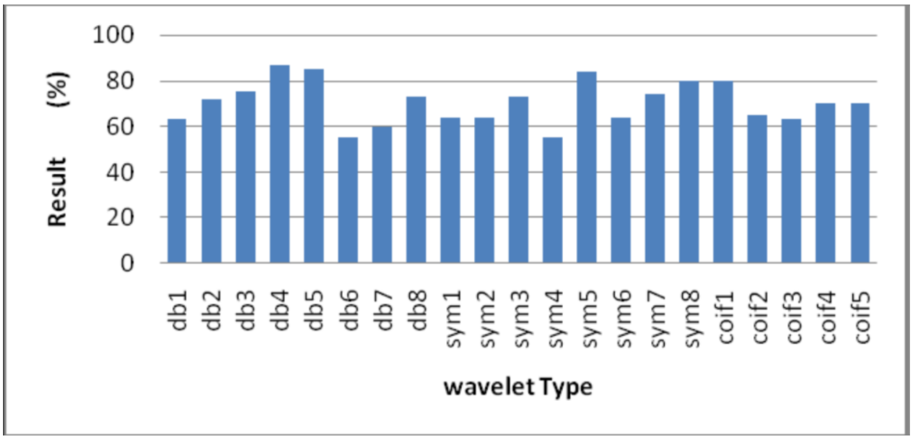

Fig.- 7: Result using different wavelet types

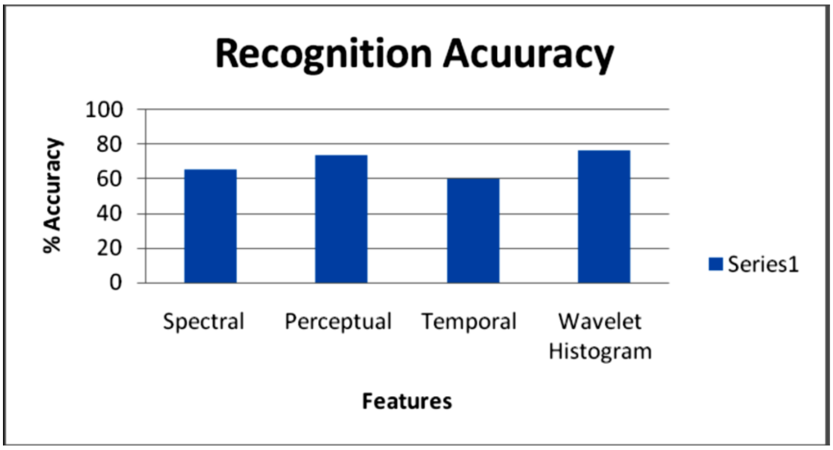

Fig. -8 :Recognition accuracy using different features

Fig.7and 8 show the overall accuracy using spectral features (40-48), Perceptual features (1-39), Temporal features(101-104), Wavelet histogram features (53-100). Eronen and Klapuri(2000) presented a musical instrument recognition system using 32 spectral and temporal features. Martin and Kim(1998) used features calculated from the log-lag correlogram rather than features based on the Short-Time Fourier Transform (STFT) to classify instruments hierarchically. Kaminsky and Materka(1995) examined the RMS of a group of instruments and reduced this data using PCA . Tzanetaki proposed a new technique in music genre classification. Three feature sets for representing timbre texture, rhythmic and pitch content are proposed by them. Kaminskyj and C. Pruysers(2005) described the technique to improve the Musical Instrument Recognition by adding Wavelet packet based features. The author claimed that the recognition accuracy is improved by five percentages by adding Morlet and Daubechies features. Popescu, Gavat and Mithai Datcu(2009) suggested a technique of Wavelet analysis for audio signals with music classification application. They used multi-resolution (wavelet) analysis and spectral analysis based features. The Proposed approach uses a no. of features like Mel Frequency Cepstral Coefficient (MFCC), Zero Crossing rate and FFT Coefficients combined with wavelet based features. Using the proposed features the accuracy of 75 percent was claimed . Tao Li and Qi Li(2006) proposed features based on wavelet coefficients at various frequency sub bands of daubechies wavelet for music genre classification and emotional content of the music . The author used features like timbral , texture, rhythmic content, pitch content, spectral centroid, and roll off ,zero crossing, LPC, MFCC, spectral flux. The author claimed the result up to 80 percent.

From literature survey it is seen that, finding the unique characteristic of musical instruments is a crucial step and poses challenges for the researchers. People are using common features like MFCC, spectral features, MPEG features and applying to different classifiers. The main problem in Musical Instrument Recognition and classification is to find optimal and compact feature set which allow the classifier to build it quickly and easily.

Conclusion:

We have proposed an innovative feature set for musical instrument recognition based on wavelet coefficient histograms, which is compact and optimized. The result reflects that musical instruments information is available in specific frequency band information. The energy in different sub band can easily retrieved using wavelet transform. The overall accuracy of the system using Wavelet Transform recorded is 76.83% compared to 73.82 % using MFCC and other features. The Db4 and Db5 give good accuracy. An improvement of the results may be possible by adding other timberal features, features selection algorithms and different machine learning algorithms.

References:

Brown, J.(1998). Computer Identification of Musical Instruments Using Pattern Recognition with Cepstral Coefficients as Features. J. Acoust. Soc. Am. 105, 1933-1941.

Deng, J.D., Simmermacher, C. and S. Cranefield (2008). A study on feature analysis for musical instrument classification, IEEE Trans. On System, Man, and Cybernetics part B: cybernetics, Vol. 38, no. 2.

Eronen, A. and Klapuri, A.(2000). Musical Instrument Recognition Using Cepstral Coefficients and Temporal Features. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 753-756 (2000).

Eronen, A. and Klapuri, A. (2000). Musical instrument recognition using Cepstral Coefficient and Temporal Feature, IEEE proceedings on International Conference on Acoustic, Speech and Signal Processing vol. 2, pp. 753-756.

Farooq, O. and Datta,S.(2004). Wavelet based robust sub-band features for phoneme recognition. IEEE Vis Image Signal Process. 151(4), 187193.

Kaminsky, I. and Materka, A.(1995). Automatic Source Identification of Monophonic Musical Instrument Sounds. In: IEEE Int. Conf. On Neural Networks, 1, 189-194(1995).

Livshin, A.A. and X. Rodet, X.(2004). Musical instrument identification in continuous recordings, in Proc. Int. Conf. Digital Audio Effects, Italy.

Popescu, M., Gavat, A. and Datcu, I. (2009). Wavelet analysis for audio signals with music classification applications, in proc. of speech technology and human computer dialogue, pp. 16.
Speech and Audio Processing, vol. 17, no. 1, pp. 174186.

Pruisers, C., Kaminskyj, I. and Schnapp (2005). Wavelet analysis in musical instrument sound classification, in international symposium on signal processing and its applications, pp. 14.

Martin, K. D. and Kim, Y. E. (1998). Musical Instrument Identification: A Pattern Recognition Approach. In: 136th meeting of the Acoustical Society of America, Cambridge, MA 02139.

T. Li, Q. Li, and Ogihara, M. (2006) Music feature extraction using wavelet coefficient histograms, US Patent 7,091,409 B2.

Tzanetakis, G. and Cook, P. (2002). Musical Genre Classification of Audio Signals, IEEE Transactions on Audio, Speech and Language Processing, vol. 10, no. 5, pp. 293302.

****