Information and Knowledge Management www.iiste.org
ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)
Vol.3, No.10, 2013

A Novel Techniques for Classification of Musical Instruments

Dr. D.S. Bormane, Ms. Meenakshi Dusane
Rajarshi Shahu College of Engineering,
Department of Electronics and Telecommunication Engineering, Pune, India.
Email: bdattatraya@yahoo.com[1] , mpwagh.2009@gmail.com[2] ,

Abstract

Musical instrument classification provides a framework for developing and evaluating features for any type of content-based analysis of musical signals. Signal is subjected to wavelet decomposition. A suitable wavelet is selected for decomposition. In our work for decomposition we used Wavelet Packet transform. After the wavelet decomposition, some sub band signals can be analyzed, particular band can be representing the particular characteristics of musical signal. Finally these wavelet features set were formed and then musical instrument will be classified by using suitable machine learning algorithm (classifier).

In this paper, the problem of classifying of musical instruments is addressed. We propose a new musical instrument classification method based on wavelet represents both local and global information by computing wavelet coefficients at different frequency sub bands with different resolutions. Using wavelet packet transform (WPT) along with advanced machine learning techniques, accuracy of music instrument classification has been significantly improved.

Keywords: Musical instrument classification, WPT, Feature Extraction Techniques, Machine learning techniques.

I. INTRODUCTION

The classification of musical instrument, where the idea is to build a computer system that listen to musical note and recognize which instrument is playing. Musical instrument classification plays an important role in developing automatic indexing and database retrieval applications. Musical content analysis in general has many applications including structure coding, music transcription, music indexing, and automatic musical signal annotation, creation of spatial sound effects. Musical instrument is a crucial subtask in solving these difficult problems and also provides useful information in other sound source.

Music is not only for entertainment and for pleasure but also for a wide range of purposes due to its social and physiological effects. Efficient and accurate automatic information music processing will be an extremely important issue, and it has been enjoying a growing amount of attention. Content based music genre classification is a fundamental component of music information retrieval system and has been gaining importance and enjoying a growing amount of attention with emergence of digital music on the internet. Up-to-date the reported very little work has been done on automatic musical instrument classification having accuracy very low [9]. This suggests feasibility of automatic musical instrument classification.

A musical instrument sound is said to have four perceptual attribute: pitch, loudness, duration and timbre. These four attributes make it possible to distinguish musical sound from each other. For musical instrument pitch is almost defined and is almost equal to the fundamental frequency. The physical counterpart of loudness is intensity which is proportional to the square of amplitude of acoustic pressure.

Automatically extracting music information is gaining importance as a way to structure and organize the increasingly
large numbers of music files available digitally on the Web. It is very likely that in the near future all recorded music in human history will be available on the Web. Automatic music analysis will be one of the services that music content distribution vendors will use to attract customers [8].

Genre hierarchies, typically created manually by human experts, are currently one of the ways used to structure music content on the Web. Automatic musical genre classification can potentially automate this process and provide an important component for a complete music information retrieval system for audio signals. In addition it provides a framework for developing and evaluating features for describing musical content. Such features can be used for similarity retrieval, classification, segmentation, and audio thumb nailing and form the foundation of most proposed audio analysis techniques for music [1]. Classification of musical instruments into a single unique class based computational analysis of music feature representations [9].

The paper is structured as follows. Introduction is given section I. A review of related work is in Section II. Feature extraction and the specific feature sets for describing timbral texture, rhythmic structure, and pitch content of musical signals are described in Section III. Section IV motivation V deals with the proposed evaluation and Section VI results.

II. Related Work

The basis of any type of classification system is the extraction of features. A large number of different feature sets, mainly originating from the area of music recognition have been proposed to represent musical signals. Typically they are based on some form of time frequency representations. Although a complete overview of feature extraction is beyond the scope of this paper, but some relevant music feature extraction references are provided.

Automatic classifications of musical instruments have a long history. Mel frequency cepstral coefficients (MFCC) [7] are perceptually used for music instrument recognition. They provide a compact representation of the spectral envelope. Spectral envelope is the key information of instrument characteristic and timbre. In this most of the signal energy is concentrated in the first coefficients.

More recently music genre classification includes specifically, three feature sets for representing timbral texture, rhythmic content and pitch content can be used [1]. The performance and relative importance of the proposed features is investigated by training statistical pattern recognition classifiers using real world audio collections.

Gavat [3] proposes a technique of multi-resolution analysis and spectral analysis. By using no of features like Mel frequency Cepstral coefficients (MFCC), zero crossing rate and FFT coefficients realize the classification of musical pieces in music genre class. The wavelet transform was used to obtain the signal representation at different levels. Time and frequency features are extracted for music classification [4]. A new design of a music note recognition system based quantization on time delay networks, self organizing maps, and linear vector quantization was proposed [5] which can be used for music transcription, song retrieval. In this method they used 76 NN to cover the range of instrument from A1 to C8 which gives the correct recognition rate. The application of such type of musical note recognition can be done for the music writing.

The pitch dependency method has been fully exploited for musical instrument classification [6]. This Fundamental frequency represents the pitch dependency of each feature. Musical instruments are analyzed by discrimination function based on the Bays decision rule. One of the novel method of music instrument recognition is done by HMM (hidden Markov Model). Spectral envelop is the key information of instrument characteristic and timbre [7]. By decomposing an instrument sound into harmonics and noise components and then by estimating the amplitudes of harmonic component the classification of musical instrument can be achieved. Feature extraction is the main step in the classification of musical instruments. A new music feature extraction method addressed by Tao Li and Qi Li [13]. They used wavelet coefficients at various frequency sub bands for music classification techniques and also used different types of classification methods like KNN, GMM and STFT. They used the features like timbral texture, rhythmic content, pitch content, spectral centroid, roll off, zero crossing, LPC, MFCC, spectral flux. The automatic classification of music instrument of such type of method can be extended for identifying emotional content of music.

There is in fact a growing interest for music information retrieval (MIR) applications amongst which is the most popular are related to music similarity retrieval, artist identification, musical instrument recognition. Currently in MIR related classification system usually do not into account the midterm temporal properties of the signal (over several frames) and lie on the assumption that the observations of the features in different frames are statically independent. The aim of Cyril Joder to demonstrate the usefulness of the information carried by the evolution of these characteristics over time. In this described temporal integration for audio classification with application to musical instrument classification [8]. In this system they used the methods like HMM, GMM and DTW.

Tao Li and Qi Li [9] addressed the comparative study on content based music genre classification. By computing the histograms on the Daubechies wavelet (DWCHs) coefficient the local and global information of music signals can be captured. They compared the effectiveness of new feature and previously studied features by using the machine learning algorithms. A new statistical approach to musical instrument classification using non-negative matrix factorization introduced [10]. The new feature set generates a vector space to describe the spectrogram representation of a music signal. The space is modelled statistically by a mixture of Gaussians (GMM). They presented the new feature set based on NMF of the spectrogram of a music signal for the description of the vertical structure of music.

A novel technique suggested by Qian Ding [11] for the classification of recorded musical instruments sounds based on Neural Networks. The classification of musical instruments was on the basis of a limited number of parameters and timbral features. Regarding the task of classification, they designed a two layer Feed-Forward Neural Network (FFNN) using back propagation training algorithm. George Tanetakis [12] had done the classification of musical instrument by using individual partials. The strategy presented that explores the spectral disjointness among instruments by identifying isolated partials, from which a number of features are extracted. The information contained in those features, in turn is used to infer which instrument is more likely to have generated that partial. If several isolated partials are available, the summarized into a single, more accurate classification. A possible shortcoming of this algorithm would be its dependency on other tools to perform tasks like onset detection, estimation of number of simultaneous instruments and estimation of the fundamental frequencies. A fair compararision is only possible if the same signals and the same number of instruments are considered. A direct comparision with other methods for instrument recognition was not presented their due to several practical constraints.

III. Feature Extraction

Many different features can be used for music classification e.g. content based features including tonality, pitch and beat, symbolic features extracted from the scores and text based features extracted from the song lyrics. But for the accurate classification of musical instruments we are interested in only content based features. The content based features are classified into timbral texture features, rhythmic content features and pitch content features [9]. Typical timbral features include spectral centroid, spectral Roll off, Spectral Flux, energy, Zero Crossings, Linear Prediction coefficients, and Mel Frequency cepstral Coefficients (MFCCs). Among all these features MFCCS are dominantly used in music recognition.

III.A. Timbral Textural Features:

Timbral texture features(spectral features) are used to differentiate mixtures of sounds that are possibly with the same or similar rhythmic and pitch contents. To extract the timbral features the music signal is divided into frames that are statically stationary usually by applying a windowing function, at fixed intervals. The window, remove the edge effects. Then the timbral texture features are then computed for each frame by calculating mean and variance (stastical features).

- Spectral centroid: is the centroid of the magnitude spectrum of short term Fourier transform and it is a measure of spectral brightness. Formally spectral centroid C_t is defined as

$$C_t = \frac{\sum_{n}^{N-1} f(n)x(n)}{\sum_{n}^{N-1} x(n)} \quad . \qquad \text{........ (1)}$$

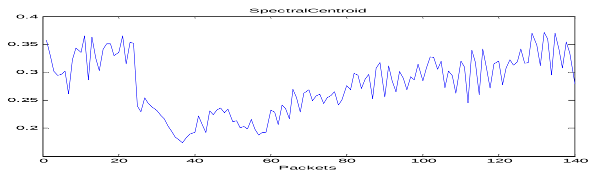


Fig 1. Spectral centroid for guitar signal

- Spectral Roll off: is the frequency below which 85% of the amplitude distribution is concentrated. It measures the spectral shape.

$$\sum_{n=0}^{M} f(n) = 0.85 \times \sum_{n=0}^{M} f(n) \qquad \text{.......... (2)}$$

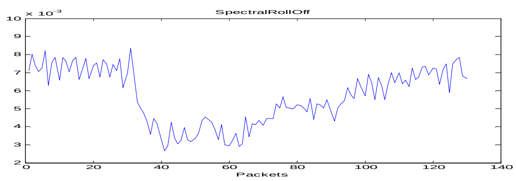


Fig.2. Spectral roll off for guitar signal

- Spectral Flux*:* The spectral flux is defined as the squared difference between the normalized magnitudes of successive spectral distributions

Spectral flux= $$\sum_{K=2}^{K} M(f_k) - M((f_{k-1})$$ .... (3)

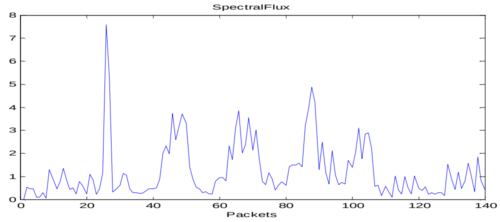


Fig 1. Spectral flux for guitar signal

- Zero crossings is the number of time domain zero crossings of the signal. It measures the noisiness of the signal. Normally, the time domain zero crossings Zt is defined as

$$Z_t = \frac{1}{N} \sum sign(x[n]) - sign(x[n-1])$$ .... (4)

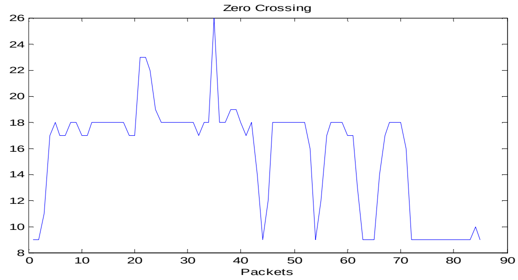


Fig.2.Zero crossing for guitar signal

- Mel-Frequency Cepstral Coefficients: Mel-frequency cepstral coefficients (MFCC) are perceptually motivated features that are also based on the STFT. After taking the log-amplitude of the magnitude spectrum, the FFT bins are grouped and smoothed according to the perceptually motivated Mel-frequency scaling. Finally, in order to decor relate the resulting feature vectors a discrete cosine transform is performed. Although typically 13 coefficients are used for speech representation, we have found that the first five coefficients provide the best genre classification performance.
- Energy is simply the sum of the amplitudes present in a frame and defined as,

Energy= $$\sum_{n}^{N-1} (x[n])^2$$ .... (5)

III.B Time domain features:

ADSR Envelope:

Fig 1. ADSR Envelope

Time domain features (Temporal features) related to the shape of the envelope of the musical note. Attack time is the time difference between onset and end-of-attack. Onset time is the time at which the note begins to sound. Decay time is the time difference between end-of-attack and the forward position where the amplitude is 25% of the amplitude at end-of-attack. Roll-off Rate is the rate of decrease of amplitude from end-of-attack to the forward position where the amplitude is decreased by 25% of the amplitude at end-of-attack.

III.C. Rhythmic Content Features:
Rhythmic content features are those characterize the regularity of the rhythm, the beat, tempo and the time signature. The feature is calculated by periodic changes from the beat histogram. In beat histogram we select the two highest peaks and compute their relative amplitude and the ratio between relative amplitudes, and the period length of each. By adding the overall average of the amplitude a total of six features are calculated [1]-[2].

III.D. Pitch Content Features:
The pitch content features describe the melody and harmony of information about music signals and are extracted based on various pitch detections algorithms. The pitch content features typically include the amplitudes and periods of maximum peaks in the histograms. Basically the dominant peaks of the autocorrelation function, calculated via the summation of envelopes for each frequency band obtained by decomposing of the signals and the pitch content are extracted from the pitch histograms. Also pitch can be calculated by different pitch detection techniques like DFT, FFT, Cepstral and many more.

IV. Motivation:

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.

The motivation in studying wavelet transform was provided by the fact that signals can be modelled suitably by combining translation and dilation of a simple, function called a wavelet. 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.

The WPT system was proposed by Ronald Coifman to allow finer and adjustable resolution of frequencies at high frequencies. Wavelet packets are required to find the parameters of the signal hidden in specific bands.

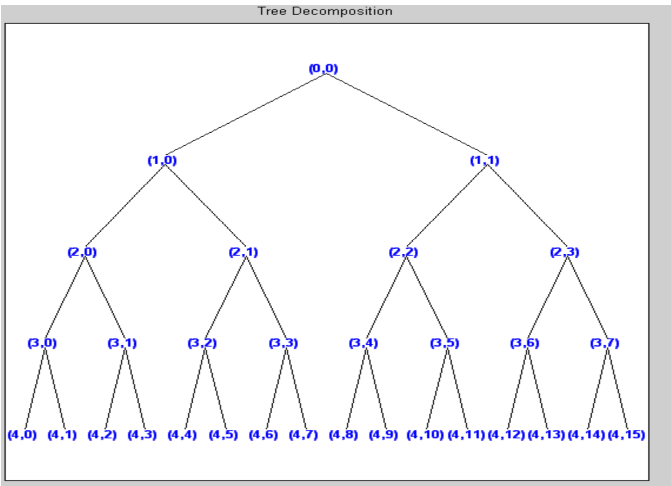


Fig.2 wavelet Packet Tree

V. Evaluation:

The musical instruments are commonly sorted into 4 families according to their vibration nature, which are brass, keyboard, string and woodwinds. Currently, our small instrument database has 16 wav files and we are considering the ten C, D & E notes.

The signal is subjected to wavelet decomposition. In our work we used the four levels of decomposition. Daubechies wavelet filters are commonly used for music retrieval. The decomposition of music signal using wavelets produces a set of sub band signals at different frequencies corresponding to different characteristics. This motivates the use of wavelet technique for feature extraction method. The wavelet coefficients are distributed in various frequency bands at different resolutions.

In this method, the silence part of the musical signal is removed first i.e. from starting and from end. Then WPT of music signal is taken. In WPT, the music signal is decomposed into 16 bands using 4 level decomposition. The energies of all 16 bands are calculated. Then the ratio of the energy in each band to energy in first band is calculated i.e. E_2/E_1, E_3/E_1, E_4/E_1 and so on. This normalises the parameters. And then statistical parameters are calculated for these energy ratios to classify the musical instrument. Here we are calculating the statistical parameters like average, mean, standard deviation and variance.

The algorithm contains following steps:

1. Read wav file
2. The 4 level wavelet decomposition of the music signal is obtained by using WPT.
3. Calculate the energy of each sub band.
4. Also compute the ratio of the energy in each sub band to the energy of first sub band.
5. Calculate the statistical parameters like variance, standard deviation and mean.

The output of 4 level decomposition is given:

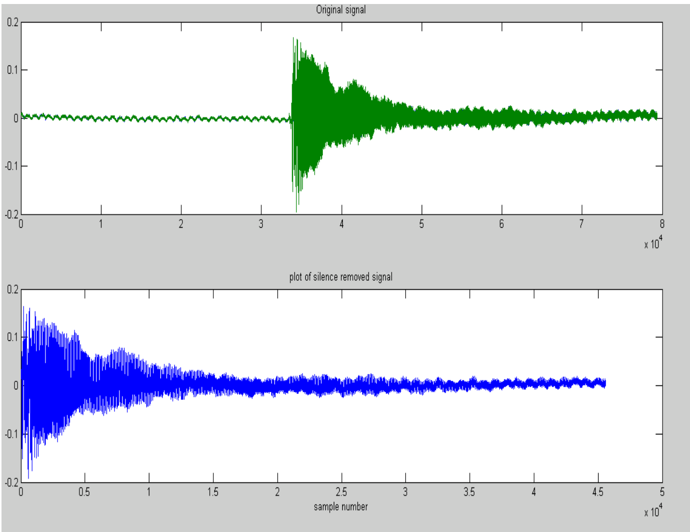
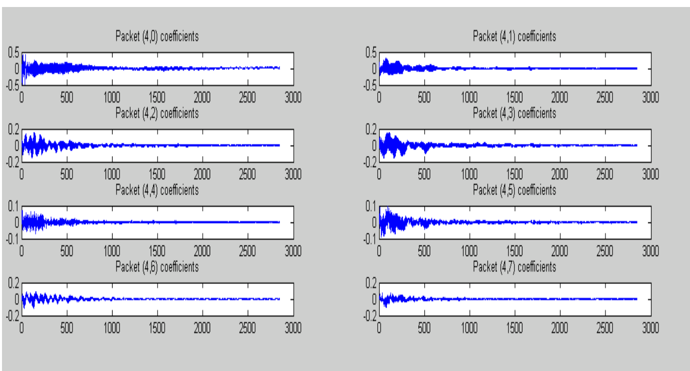


Fig3. Original signal of guitar C2 notes and its silence removed signal



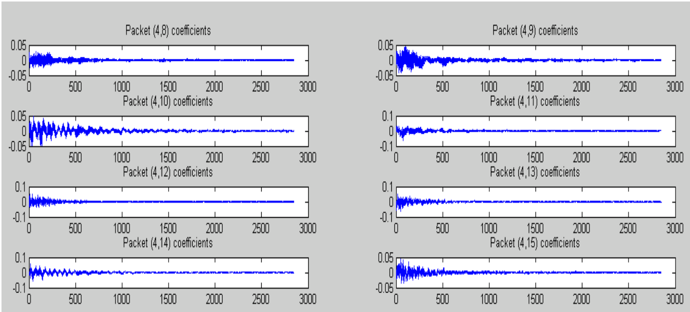


Fig 4 shows wavelet packet coefficients guitar C2 notes

VI. Result:

We have conducted experiments on sixteen musical instruments which are Cornet, Trombone, Trumpet, Tuba, Celesta, Harpsichord, Piano, Guitar, Sitar, Flute, Violin, Horn, Harmonica, Oboe and Saxophone. All the work is implemented in MATLAB to get results. Table 1 shows the individual instrument recognition accuracy. In this we are considering the ten instruments for checking the accuracy by using Wavelet transform. The results are getting for 14 instruments. It means we can better parametrically represent musical instruments using the wavelet transform.

Family	Instrument	Testing Notes	Recognized Notes	Accuracy
String Instruments	Bass	10	7	100%
	Guitar	10	8	80%
	Sitar	10	10	100%
	Violin	10	6	60%
Keyboard Instruments	Celesta	10	1	10%
	Harmonica	10	6	60%
	Harpsichord	10	10	100%
	Piano	10	5	50%
Woodwind Instruments	Flute	10	10	100%
	Horn	10	5	50%
	Oboe	10	10	100%
Brass Instruments	Coronet	10	7	70%
	Trombone	10	8	40%
	Tuba	10	10	100%

Table 1: Individual instrument recognition accuracy

VII. Conclusion:

In this paper we proposed a new feature extraction method for the classification of the musical instruments. We have conducted experiments on sixteen musical instruments. Only for the fourteen instruments we get the results hence to improve the result of this system we have to look specific musical instrument features which clearly characterize the musical instrument. There is scope to investigate new specific features which can clearly distinguish musical instruments. We are working progressively in this direction.

References :

[1] George Tzanetaki *"Music classification of audio signals"* IEEE Transactions on speech and audio processing, Vol. No. 5, July 2002.

[2] I. Kaminskyj and C. Pruysers *"Wavelet analysis in musical instrument and sound classification"* Proc. Of IEEE 2005.

[3] Inge Gavat and Mithai Datcu *"Wavelet analysis for audio signals wih music classification application"* Proc. Of IEEE 2009.

[4] Macro Gridmaidi and Padraig Cunningham *"A wavelet packet representation of audio signals for music genre classification using different ensemble and feature selection techniques"* Proc. Of IEEE 2003.

[5] Khalid Youssef *"Music note recognition system based on time delay neural network"* Fourth International Conference on Natural Computation IEEE 2008.

[6] Ketsuro Etahara *"Musical instrument identification based on F0 dependent multivariate normal distribution"* Proceedings of IEEE 2003.

[7] Antti Erronen *"Music instrument recognition using ICA based transforms of features and discriminatively trained HMMs"* Proc. Of 2003.

[8] Cyril Joder and Slim Essid *"Temporal Integration for audio classification with application to musical classification"* IEEE Transactions on Speech and audio processing, Vol. No. 1, January 2009.

[9] Tao Li and Qi Li in *"Comparative study on content based music genre classification"* Proc. Of 2003.

[10] Andre`s Holzapfel and Yannis Stylianou *"A stastical approach to musical genre classification using non-negative matrix factorization"* Proc. Of IEEE 2007.

[11] Qian Ding ng *"Classification of Recorded Musical instruments Sounds based on Neural Networks"* Proceedings of the 2007 IEEE symposium on Computational Intelligence in Image and Signal Processing.

[12] George Tzanetikis *"Musical Instrument classification using individual partials"* IEEE Transactions on Speech and audio processing, Vol. No. 19, January 2011.

[13] Tao Li & Qi Li *"Music feature extraction using wavelet coefficient histograms"* United States Patent August 2006.

[14] Dr. Shaila D. Apte, *"Digital Signal Processing"*, Second Edition, Wiley India.