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

Comparative Study of Musical Instrument Recognition

A. S. Patankar
International Institute of Information Technology,Hinjawadi,Pune
patankar.kargi@gmail.com

Abstract:

Identification of the musical instrument from a music piece is becoming area of interest for researchers in recent years. The system for identification of musical instrument basically performs three tasks: i) Pre-processing of input music signal; ii) Feature extraction from the music signal; iii) Classification.

The basic idea of this project is to identify musical instruments from audio recordings by extracting wavelet features. Further, two classifiers K-Nearest Neighbors (K-NN)and Support Vector Machine (SVM) are used to identify the musical instrument name. Also, results are calculated comparatively.

Keywords: Musical Instrument, Recognition, comparative study

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 exampleDue 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.

Related work

Eronen presented a musical instrument recognition system using 32 spectral and temporal features. Gaussian and k-NN classifiers are used for classification [1].

Arie A. livshin have used Gradual Descriptor Elimination (GDE) algorithm for feature selection and k-NN with Linear Discriminant Analysis(LDA) for classification[2].

An experimental study on feature analysis for recognition of musical instrument with the help of k-NN, Artificial Neural Network ,and Support Vector Machine(SVM) classifier have been performed by J. D. Deng, C.Simmermacher. They used MFCC features [3].

Brown distinguished between oboe and saxophone by calculating cepstral coefficients and applying a 'k' means algorithm to form clusters [4].

I. Kaminskyj and C. Pruysers 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. Using the proposed feature sets classification accuracy of 87.6 percent was claimed for 19 musical instruments [5].

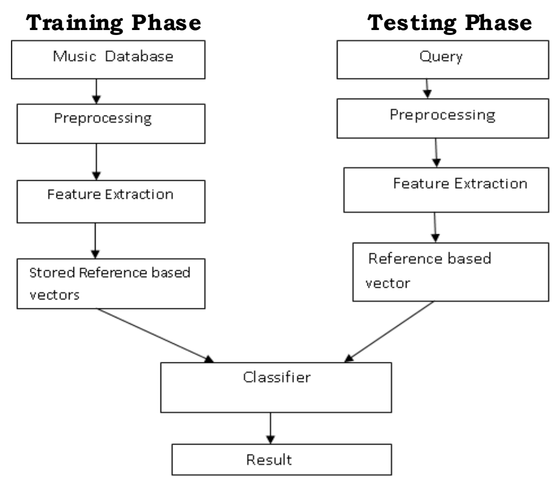

Figure 1: General block diagram of Musical Instrument Recognition

WORKFLOW

Dataset

The database for the experiments contains 150 samples which are taken from recorded database and recorded Sample tone at sampling frequency 11025 Hz.The recordings are categorized into general classes according to common characteristics of the scenes (10 flute,10Dilruba,10 Ghatam,10 Harmonium,10 MohanVeena,10 Mrudangam,10Santoor,10 Sarangi,10 Sitar,10 Tabala)and events the categorization of the scenes was somewhat ambiguous, some of the recordings are associated with more than one higher-level class. Every sound signal was stored with some properties that are also the initial conditions and criteria for the well-functioning of the algorithm. The sample database is split into training (10 samples of each instrument) and testing (5 samples of each instrument) sets(fig 1).

Properties of collected sound samples:

1.Audio type=Wavesound(.wav)
2.Frequency=11025 Hz
3.Bit rate=352 kbps
4.Duration=4 sec

Preprocessing

Passband Filter

A passband is the range of frequencies or wavelengths that can pass through a filter.For example, a radio receiver contains a bandpass filter to select the frequency of the desired radio signal out of all the radio waves picked up by its antenna. The passband of a receiver is the range of frequencies it can receive.A bandpass-filtered signal (that is, a signal with energy only in a passband), is known as a bandpass signal, in contrast to a baseband signal.The variations in the passband is the passband ripple,or the difference between the actual gain and the desired gain of unity.The stopband attenuation is required to stop particular frequencies which are not present in the specified passband frequency range.The stopband attenuation cannot be infinite, so we must specify a value with which we are satisfied.We measure passband ripple and the stopband attenuation in decibels(dB).

Feature Extraction Technique

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. The Wavelet Transform(WT) is a transform which provides a time frequency representation. There are two types of wavelet transform viz- Continuous Wavelet Transform and Discrete Wavelet Transform.

We are using the Discrete Wavelet Transform as this transform decomposes the signal into mutually orthogonal set of wavelet.

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

Haar Wavelet Transform

Haar is one of the example of DWT.We are using it for obtaining coefficients of matrix. Haar wavelet is a sequence of rescaled "square-shaped" functions which together form a wavelet family or basis. Wavelet analysis is similar to Fourier analysis in that it allows a target function over an interval to be represented in terms of an orthonormal basis. The property can of non continuous be an advantage for the analysis of signals with sudden transitions, such as monitoring of tool failure in machines.

We have done the 5 level decomposition on the applied input samples and obtained the statistical parameters i.e Mean,Standard deviation and Variance.

Formulae for obtaining the statistical parameters are as follows-

Mean:Mean is the average of all values in the sample.

Standard Deviation:Standard deviation is the average distance from mean of the dataset to each point.

Variance:Varience is the square of standard deviation(fig 1).

Classification

k-NN classifier

k-Nearest Neighbor classifiers are based on the idea that an object should be predicted to belong to the same class as the objects in the training set with the biggest similarity.To use kNN classification, it is required to have a suitable similarity search system in which the training data is stored. Classification is done by analyzing the results of a kNN query. kNN classifiers use the class borders of the objects within the training set and thus do not need any training or model building apriori. As a result, kNN classifiers cannot be used to gain explicit class knowledge to analyze the structure of the classes. kNN classification is also known as a lazy learning algorithm. The parameter that determines the neighborhood size, k, is very important to the classification accuracy achieved by a kNN classifier.

Euclidean distance is used in k-NN in this paper.

Support Vector Machine

To determine the exact position of the maximum margin hyperplane and to find out the support vectors, a dual optimization problem is formulated which can be solved by algorithms like Sequential Minimal optimization.The problem of linear separation is that there is not always a hyperplane that is able to separate all training instances. For many real world applications, it is not possible to find a hyperplane that separates the objects of two classes with sufficient accuracy. To overcome this problem the feature vectors are mapped into a higher dimensional space by introducing additional features that are constructed out of the original ones. Since this mapping is not linear, hyperplanes in the so-called kernel spaces provide much more complicated separators in the original space.

However, the training of SVMs tends to take large periods of time, especially for multi-class variants calculating many binary separators. The models built by SVMs do not provide any explicit knowledge that might helto understand the nature of the given classes(fig 1).

Result:

Formula for calculation of average efficiency is

Average Efficiency(%) = (No. of samples correctly identified/Total no. of samples) *100

The Average Efficiency of classification using the k-NN classifier and SVM classifier is described in Table -1(fig 2).

Table 1: Average Efficiency of Musical Instrument Recognition System

Name of musical instument	% efficiency (k-NN)	% efficiency (SVM)
Dilruba	80	100
Flute	80	100
Ghatam	80	100
Harmonium	100	100
Mohanveena	60	100
Mrudangam	100	100
Santoor	100	100
Sarangi	80	80
Sitar	100	80
Tabala	80	80
Average efficiency	**86**	**94**

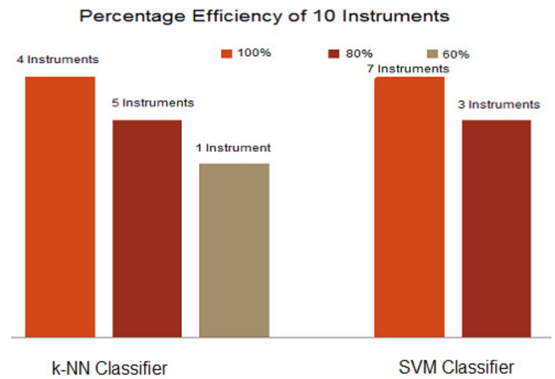

Figure 2: Bar graph average efficiency of k-NN and SVM

Conclusion:

We have used Discrete Wavelet Transform with Haar for decomposing the applied input music sample and generated the feature vector which contains features like mean,variance and standard deviation.We have used k-NN classifier as the first method for classification with accuracy of 86 percent and the second method is SVM with 96 percent.After comparing the classification results obtained from both k-NN and SVM ,we can conclude that the SVM classification method is suitable over k-NN method for performing and giving better classification of the applied inputs.

References:

[1] A. Eronen, A. Klapuri, 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, 2000.

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

[3] J. D. Deng, C. Simmermacher, and S. Crane field, A study on feature analysis for musical instrument classification, IEEE Trans. On System,Man and Cybernetics-part B: cybernetics, Vol. 38, no. 2,April 2008.

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

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

[6] M. Popescu, A.; Gavat, I.; Datcu, Wavelet analysis for audio signals with music classification applications, in proc. of speech technology and human computer dialogue, pp. 16.,2009.