A Study on Traditional Malay Musical Instruments Sounds Classification System

Norhalina Senan
Universiti Tun Hussein Onn Malaysia
86400 Parit Raja, Batu Pahat
Johor, Malaysia
+607-4538093
halina@uthm.edu.my

Rosziati Ibrahim
Universiti Tun Hussein Onn Malaysia
86400 Parit Raja, Batu Pahat
Johor, Malaysia
+607-4538027
rosziati@uthm.edu.my

Nazri Mohd Nawi
Universiti Tun Hussein Onn Malaysia
86400 Parit Raja, Batu Pahat
Johor, Malaysia
+607-4538002
nazri@uthm.edu.my

ABSTRACT
In recent year, most studies concerned with the classification of musical instruments sounds focus on western musical instruments. With the enormous amount of instruments data and features schemes, adapting the existing techniques for classifying the traditional Malay musical instruments sounds might not be as easy due to the differences in the sounds samples used. Thus, the existing framework and techniques that have been proposed for automatic musical instruments sounds classification system will be reviewed and evaluated especially on their performance in achieving the highest accuracy rate. As a result, a new framework for Traditional Malay Musical Instruments Sounds Classification System and the classification accuracy achieved in the preliminary experiment are presented.

Categories and Subject Descriptors
H.5.5 [**Information Interfaces and Presentation**]: Sound and Music Computing – *methodologies and techniques, signal analysis, synthesis, and processing, systems*

General Terms
Design and Experimentation

Keywords
Musical instruments sounds classification framework, Traditional Malay musical instruments

1. INTRODUCTION
With the advances of digital signal processing and machine learning techniques, automatic musical instrument sounds classification system has becomes an important aspect of music information retrieval (MIR). Mackay and Fujinaga [13] claimed that automatic musical instruments sounds classification using machine learning is better than human capability in producing a good result.

iiWAS2009, December 14–16, 2009, Kuala Lumpur, Malaysia.

This automatic classification system can have numerous potential applications. For instance, recognizing and analyzing the content of the instrument (sound signal) can lead to more knowledge about the different musical styles and can be further utilized for computer-assisted musical instrument tutoring [6]. Furthermore, it also can be enhanced as a validation or quality control tool in musical instrument manufacturing.

Several studies have been conducted regarding these issues [2][4][11]. However, almost all the studies are developed based on the Western musical instruments. Meanwhile, study on non-Western musical instrument especially on Traditional Malay musical instruments sounds is still lacking.

Thus, this paper discusses the overall issues and a review of the approaches and techniques involved in the development of the Traditional Malay musical instruments sounds classification system. Then the result of preliminary experiment is discussed. This result may lead to future improvement of Traditional Malay musical instruments sounds classification performance in subsequence experiment.

The remainder of this paper is organized as follows: Section 2 presents the problem statement. A review of the techniques and the processes are presented in the Session 3. The following Section 4 introduces the framework of Traditional Malay musical instruments sounds classification system. Section 5 presents the preliminary result and lastly in Section 6, the conclusion and future work of this study.

2. PROBLEM STATEMENT
With the various potential applications that can be developed from this automatic musical instrument sounds classification system, there were significance need to explore further in this area of research especially when it involved a new domain which is Traditional Malay musical instrument. Some of the researchers believe that different musical instruments sound have different characteristic or behavior [7]. Therefore, adapting the existing system for retrieval of Malaysian musical instruments contents might not be as that simple.

Moreover, the implementation of musical instruments sounds classification system still has restricted practical usability due to the certain problem especially to find the right feature extraction schemes for the musical instrumentals sounds. It is also fascinating to see that the feature schemes adopted in current research are all highly redundant. Due to a large number of sound features available, how to select or combine them to achieve higher classification accuracy is important [10].

Hence, by simply selecting all the features available, it might give poor performance. This is because some features give poor separability among different classes and some are highly correlated [10]. Furthermore, Deng et al. [2] also found that some of the features that they used in their research were not reliable for giving robust classification result. All these show that one of the most crucial issues in musical instruments sounds classification is to select the right features extraction schemes from sounds database. Consequently, this study has a significant importance to find better mechanisms for this problem.

Therefore, the objectives of this research are (a) to adapt the existing feature schemes into traditional Malay musical instrument sound; (b) to design and formulate a new feature selection method for identifying a good feature combination schemes; and (c) to validate the generated feature combination schemes using classifier.

3. RELATED WORKS

This section discuss briefly the overview of traditional Malay musical instruments and several topics related on musical instruments sounds classification such as feature extraction schemes, feature selection techniques and classification algorithms used to validate the performance of selected features.

3.1 Traditional Malay Musical Instruments

The traditional Malay musical instruments are believed to have originated from different countries and cultures. For instance, the *kompang* was brought to the Malay Region by the traders from the Middle East in thirteenth century [1]. Besides Arab countries, some of the instruments were also invented from other country such as angklung which believed was brought by migration from Indonesia [14].

The traditional musical instruments play an important role in traditional Malay culture. The instruments were mainly used to accompany traditional dance such as *kuda kepang* and *mak yong*, wedding ceremony, traditional theater such as *wayang kulit* (shadow puppet) and religious function such as M*aulud Nabi* and *berzanji* [14].

In general, the traditional Malay musical instruments can be classified into four (4) categories which are membranophones, idiophones, chordophones and aerophones [19]. Membranophones and idiophones are also known as percussion instruments. These instruments are the largest and most important instruments in Malay traditional music. Table 1 shows the category of the Malay musical instruments.

3.2 Automatic Musical Instruments Sounds Classification System

Automatic musical instruments sounds classification system is a systematic approach that able to identify the complex features of the musical signals from the musical instruments database automatically. This is concern as the first step in developing a wide variety of potential applications [2][13].

Generally, automatic musical instruments sounds classification process involved three (3) main stages which are feature extraction, feature selection and classification.

There are various algorithms that have been explored in solving problem for each stage in automatic musical instruments sounds classification system. However, there are still several remains problem that need to be tackled in producing a good classification system.

One of the most crucial issues of automated musical instruments sounds classification system is to find the best feature schemes or properties [2][4][10]. This is important because features are feed to pattern recognition system as the input and are the basis in the lead of the classification process.

3.2.1 Feature Extraction

Feature extraction can be considered as a transformation process of input data into a reduced digital representation set of features schemes. The purpose of feature extraction is to obtain the relevant information from the input data to execute certain task using small set of features instead of the large original size data.

In musical instruments sounds classification, various features schemes have been extracted and adopted by past research either by individual sets or combination of them. Normally, the features used consists both spectral and temporal features. It has been highlighted by previous work that the combination of both spectral and temporal features is essential in order to provide an accurate description of sounds timbre [12].

This study use two (2) different extraction feature categories proposed by [2], which are mel-frequency cepstral coeffients (MFCC) features and perception-based. The mean and standard deviation are calculated for each of the features for the classification purposes.

3.2.1.1 Mel-Frequency Cepstral Coefficients

Mel-frequency cepstral coefficients features have been used not only in musical instrument classification but also in other audio processing area such as music genre and speech processing [2]. The MFCC value is computed directly from the power spectrum [15]. Typically, the first thirteen (13) coefficients have been found to be most useful in musical instrument sounds features. The effectiveness of MFCC in identifying different type of audio features have been discovered in [4][16]. The following formula is used to obtain the value of MFCC features [2]:

$$mel(f) = 2595\log_{10}\left(1+\frac{f}{700}\right) \tag{1}$$

3.2.1.2 Perception-based

Perception-based features were extracted from multiple segments either in time-domain or frequency-domain of a sample signal. There are six (6) features in this category which are zero crossing, zero-crossing rate, root-mean-square, spectral centroid, bandwidth and flux.

3.2.2 Feature Selection

Feature selection, also known as feature reduction, can be defined as a technique of choosing the most relevant features for building robust classifier. By removing the irrelevant features, the performance of the classifier can be improved by reducing the "curse of dimensionality", enhancing generalization capability, and reducing learning and computational cost. In addition, many researcher also claimed that even an optimal classifier difficult to classify accurately if the poor features are presented as the input [2][10][13].

The feature selection process comprises four (4) basic steps which are subset generation, subset evaluation, stopping criterion, and result validation [9]. Subset generation is a search process, which produces candidate features subsets for evaluation based on certain *search strategy*. Each candidate subset is evaluated and compared with previous best feature based on certain *evaluation criterion*. If the current feature is better, it will replace the previous best feature. The process is repeated until the stopping criterion is satisfied.

The feature selection algorithms can be classified into two (2) main categories which are filter and wrapper algorithms [8]. The filter algorithm use the initial set of the features, and then applies the selected feature subset to the clustering algorithm, whereas, the wrapper algorithms incorporates the clustering algorithm in the feature search and selection. Essid, Richard and David [5] claimed that the wrapper algorithm more efficient than the filter algorithm, but more complex.

Liu and Wan [10] studied the feature selection for automatic classification of musical instrument sounds using filter algorithm which is sequential forward selection technique. This technique is convenient to provide a sub-optimized set of features. The results shows that the modified k-NN classifier using 19 selected features (6 temporal, 8 spectral, and 5 coefficients) achieves highest accuracy of 93%.

Whereas, study on wrapper algorithms also have been presented by several researchers. For instance, Essid, Richard and David [5] and Mackay and Fujinaga [13] applied Genetic Algorithm (GA) in their work. The GA perform better in [13] but less efficient when compared to others wrapper algorithm which is Inertia Ratio Maximization using Feature Space Projection (IRMFSP) in [5]. Essid, Richard and David [5] claim that the selection of fitness function for GA structure also can affect the overall performance.

3.2.3 Feature Validation via Classification

Classifier is used to verify the performance of the selected features. The accuracy rate achieved by the classifier is analysed to identify the effectiveness of the selected features. Achieving a high accuracy rate is important to ensure that the selected features are the best relevance features that perfectly serve to the classification architecture which able to produce a good result.

However, the performance of the overall classification system is not only depends on the features used. There is also significance to ensure that the classifier is able to analyze and extract the implicit information of these features into an intelligible form [17]. There are various classification algorithms that have been used in musical instruments sounds classification system such as Support Vector Machine (SVM) [2][5][9], *k*-Nearest Neighbours (*k*-NN) [2][4][10] and Artificial Neural Network [2][11].

The classification of the instrument into individual and four (4) groups of instrument's family which are brass, woodwind, piano and string has been discussed by [2]. They used five (5) classifiers which are SVM, k-NN, Naïve Bayes, multi-layer perceptron (MLP) and Radial Basic Function (RBF). In individual classification, 3-NN achieved average accuracy of 98.4% over four instruments. However, even full feature set would not help much in classified woodwind instrument. Meanwhile, 1-NN produced highest accuracy of 96.5% for "Selected 17" features in family classification.

Liu and Wan [10] analyzed 351 instruments sounds from five (5) different families. The main objective of their research is to identify the effectiveness of selected features on classification performance. Three (3) classifiers are used which are Nearest Neighbour (NN), k-NN and Gaussian Mixture Model (GMM). The result shows that the performance increases when more features are used. The best feature sets for different classifiers are different. The k-NN classifier using the best 19 features achieves highest accuracy of 93%.

The new feature selection and classification strategy were introduced by [5] using pairwise classification technique with Hastie-Tibshirani approach. Ten (10) individual instruments were used in this study. SVM with RBF kernel is the most successful classifier with average accuracy rate of 87%.

4. FRAMEWORK

Figure 1 shows the framework of this study which consist six (6) main activities which are data acquisition, sounds editing, data representation, feature extraction, feature selection and classification.

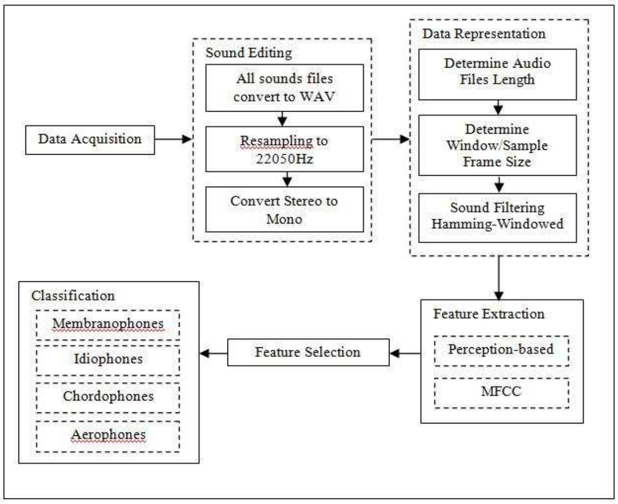

Figure 1. A Framework of Traditional Malay Musical Instruments Sounds Classification System

The brief description for each phases of this framework are as follows:

4.1 Data Acquisition

The 150 sounds of traditional Malay musical instruments were downloaded freely from personal web page at www. rickshriver.net/hires.htm and *Warisan Budaya Malaysia* web page at http://malaysiana.pnm.my/kesenian/Index.htm. The distribution of the sounds into categories is shown in Table 1.

4.2 Sound Editing

The original collection came in MP3 and WAV files format with assortment of sample rate which are 22.1 kHz and 44.1 kHz. In order to utilize Matlab function in this experiment, all the sound files is converted into WAV format. Then the sounds is down-sampled to 22.1 kHz, and convert to mono. The reason is to reduce the computational time compared using stereo file with high sample rate. Schmidt and Stone [18] also discovered that mono files provide better overall models. All these processes is done using sound editing software.

Table 1. Data Sets

Family	Instrument	Number of Sounds
Membranophone	Kompang, Geduk, Gedombak, Gendang, Rebana, Beduk, Jidur, Marwas, Nakara	41
Idiophone	Gong, Canang, Kesi, Saron, Angklung, Caklempong, Kecerik, Kempul, Kenong, Mong, Mouth Harp	75
Aerophone	Serunai, Bamboo Flute, Nafiri, Seruling Buluh	23
Chordophone	Rebab, Biola, Gambus	11
Total		150

4.3 Data Representation

There are two (2) different experimental sets are tested in this phase. In the first experiment, the original data collections is trimmed into three (3) different data sets with different interval time. The first data set (A) comprises sound files range from 0.1 to 10 seconds, the second data set (B) with range from 0.1 to 20 seconds, and the third data set (C) with range from 0.1 to 30 seconds. This is done in order to assortment the number of the sounds samples and to examine whether the length of audio files plays important role in determining the classification result. Then, the second experiment is focused on identifying whether the size of segmented frame has significant consequence to the output. For that, every audio file is segmented into frames of two (2) different sample sizes which are 256 and 1024 with overlap about 50%. The overlap procedure is to ensure there are no missing signals during the segmentation process. In order to improve the quality of the sounds, each frame is then be hamming-windowed.

4.4 Feature Extraction

In this phase, two (2) categories of features schemes which are perception-based and MFCC features are extracted. Both of the feature schemes represent the temporal and spectral features. The spectral features are computed from the Fast Fourier Transform (FFT) of the segmented signals. The mean and standard deviation are then calculated for each of the features. All 37 extracted features from two (2) categories are shown in Table 2. The first 1-11 features represent the perception-based features and 12-37 are MFCC's features.

4.5 Feature Selection

Initially, the existing algorithm which is Fuzzy-Rough Sets will be employed to select the best features. To our knowledge, this algorithm is not yet applied in musical instruments domain. For that, the effectiveness of the algorithm will be analyzed. After that, the algorithm will be enhanced in order to improve the performance of the algorithm and the classification accuracy respectively.

Table 2. Features Description

Number	Description
1	Zero Crossings
2-3	Mean and Standard Deviation of Zero Crossings Ratios
4-5	Mean and Standard Deviation of Root-Mean-Square
6-7	Mean and Standard Deviation of Spectral Centroid
8-9	Mean and Standard Deviation of Bandwidth
10-11	Mean and Standard Deviation of Flux
12-37	Mean and Standard Deviation of the First MFCC's

4.6 Classification

The data set with the selected features is further assessed using the classifier. This data set is classified into four (4) different families as stated in Table 1. The existing classifiers that will be used are Neural Network and Support Vector Machine. Then, the result will be compared. This is done in order to identify the effectiveness of the selected features. The performance of the selected features will be determined from the accuracy rate of sound classification produced by the classifier. Finally, a report of data analysis from the testing will be tabulated.

5. PRELIMINARY EXPERIMENT AND RESULT

First experiment is done based on the proposed framework. However, the selection process is not being completed. The purpose of this experiment is to examine the performance of the extracted features schemes (perception-based and MFCC). Two (2) different interval audio files length was examined which are 0.1 to 2 seconds and 0.1 to 10 seconds. Multi-Layered Perceptron (MLP) Neural Network is used as the classifier. The database is splited into two parts: training and testing with 70:30 ratio.

From the Table 3, it can be seen that, the combination of features schemes (perception-based and MFCC) achieved highest accuracy rates up to 99.57% for the data set with the interval time from 0.1 to 10 seconds and the size of segmented frame is 256. This finding is associated with the result produced [2][3][12]. It shows that the combination of various features able to represent the actual properties of the sounds and produce highest accuracy accordingly.

Table 3. Preliminary Result

Features Schemes	% Training and Testing	Data A (0.1 to 2 sec) (%)	Data B (0.1 to 10 sec) (%)
Perception + MFCC	**70:30**	**96.91**	**99.57**
Perception	70:30	57.87	81.23
MFCC	70:30	85.85	75.64

6. CONCLUSION AND FUTURE WORK

In conclusion, automatic musical instruments sounds classification is still an open problem. From the literature, we found that one of the most crucial issues in musical instruments sounds classification system is to find the best feature schemes or sounds properties. This means that more attention should be given in the data representation, feature extraction and features selection process.

In addition, a number of techniques have been applied in the past that differ in the features used to describe the important of classification strategy. However, there are potential ways to improve the algorithms especially on features selection issues that have major influence to the classification performance. Result from the preliminary experiment shows that highest accuracy can be obtained from the combination of several features schemes. However, it involved a large number of features schemes which involved a high computational time. This factor might affect the overall classification performance.

At the end of this study, the expected result to be derived is an enhancement of feature selection algorithm that can effectively select a best feature combination schemes which can improve the performance of musical instruments sounds classification system. Thus, for future work, feature selection process using the existing technique will be done in order to examine the significant of this process towards Traditional Malay musical instruments sounds classification performance.

7. REFERENCES

[1] Asmawati, I., Salina, A.S., Aini, H., Che, H.A., and Mohd, R.M.Z. 2006. "Analysis of the Sound of the Kompang for Computer Music Synthesis." 4th Student Conference on Research and Development (SCOReD 2006), IEEE, 95-98.

[2] Deng, D., Simmermacher, C., and Cranefield, S., 2008. "A Study on Feature Analysis for Musical Instrument Classification", IEEE Transactions on System, Man, and Cybernetics-Part B: Cybernetics, Vol. 38, No. 2.

[3] Ding, Q. and Zhang, N. 2007. "Classification of Recorded Musical Instruments Sounds Based on Neural Networks," IEEE Symposium on Computational Intelligence in Image and Signal Processing. April 2007. Pp. 157-162.

[4] Eronen, A. 2001. "Comparison of Features for Musical Instrument Recognition". IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics. 19-22.

[5] Essid, S., Richard, G., and David, B. 2005. "Musical Instrument Recognition by Pairwise Classification Strategies". Audio, Speech, and Language Processing, IEEE Transactions on Speech and Audio Processing,14(2), 1401–1412.

[6] Ferguson, S. 2006. "Learning Musical Instrument Skills Through Interactive Sonification", NIME '06: Proceedings of the 2006 Conference on New Interfaces for Musical Expression, 284-389.

[7] Golzari, S., Doraisamy, S., Sulaiman, M.N., Udzir, N. I. and Norowi, N. M. 2008. "Artificial Immune Recognition System with Nonlinear Resource Allocation Method and Application to Traditional Malay Music Genre Classification," Lecture Notes in Computer Science. Springer-Verlag, Heidelbeg. Vol. 5132. Pp. 132-141. ISBN: 978-3-540-85071-7.

[8] Jennifer, G. D. and Carla, E.B. 2004. "Feature Selection for Unsupervised Learning". Journal of Machine Learning Research. 845-889.

[9] Liu, H. and Yu, L. 2005. "Towards Integrating Feature Selection Algorithms for Classification and Clustering". IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 4.

[10] Liu, M. And Wan, C. 2001. "Feature Selection for Automatic Classification of Musical Instrument Sounds". JCDL '01: Proceedings of the 1st ACM/IEEE-CS Joint Conference on Digital Libraries. 247-248.

[11] Livingston, J. & Shepard, N. 2005. Musical Instrument Identification Using Wavelets and Neural Networks.University of Texas. Retrieved May 23, 2008, from www.nathanshepard.net/documents/Musical_Instrument_Identification_Using_Wavelets_and_Neural_Networks.pdf

[12] Loughran, R. , Walker, , O'neill, J. M. and O' Farrell, M. 2008. "Musical Instrument Identification Using Principal Component Analysis and Multi-Layered Perceptrons," IEEE International Conference on Audio Language and Image Processing. July 2008. Pp. 643-648.

[13] Mackay. C. and Fujinaga, I. 2005. "Automatic Music Classification and the Importance of Instrument Identification". Proceedings of the Conference on Interdisplinary Musicology (CIM05).

[14] Mohd, H.A. 2004. "Idiosyncratic Aspects of Malaysia Music: The roles of the Kompang in Malay Society. Retrieved June 2, 2008, from UNESCO at http://portal.unesco.org/culture/en/files/21753/10891249663abdullah.pdf/abdullah.pdf

[15] Molau, S., Pitz, M., Schulter, R., and Ney, H. 2001. "Computing Mel-Frequency Cepstral Coefficients on the Power Spectrum". IEEE International Conference on Acoustics, Speech, and Signal Processing. 73-76. Vol. 1.

[16] Namunu, C.M., Changsheng, X., and Ye, W. 2003. "A SVM-based Classification Approach to Musical Audio". The International Society for Music Information Retrieval Conferences.

[17] Norowi, N. M., Doraisamy, S. and Wirza. R. 2005. "Factors Affecting Automatic Genre Classification: An Investigation Incorporating Non-Western Musical Forms, " ISMIR 2005. 6th International Conference on Music Information Retrieval Proceeding.

[18] Schmidt, A.P. and Stone, T. K. M. "Music Classification and Identification System,". Retrieved August, 1, 2009, from www.trevorstone.org/school/MusicRecognitionDatabase.pdf

[19] Shriver, R. 2003. "Digital Stereo Recording of Traditional Malaysian Musical Instruments". AES 114th Convention. Journal of the Audio Engineering Society. 22-25.