TRADITIONAL MALAYSIAN MUSICAL GENRES CLASSIFICATION BASED ON THE ANALYSIS OF BEAT FEATURE IN AUDIO

N. M. Norowi[a], S. Doraisamy[b], R. Wirza[c]

[a,b,c]Faculty of Computer Science and Information Technology,Universiti Putra Malaysia
[a]noris@fsktm.upm.edu.my
[b]shyamala@fsktm.upm.edu.my
[c]rahmita@fsktm.upm.edu.my

Abstract - **Interest on automated genre classification systems is growing following the increase in the number of musical digital data collections. Many of these systems have been researched and developed to classify Western musical genres such as pop, rock or classical. However, adapting these systems for the classification of Traditional Malay Musical (TMM) genres which includes Gamelan, Inang and Zapin, is difficult due to the differences in musical structures and modes. This study investigates the effects of various factors and audio feature set combinations towards the classification of TMM genres. Results from experiments conducted in several phases show that factors such as dataset size, track length and location, together with various combinations of audio feature sets comprising Short Time Fourier Transform (STFT), Mel-Frequency Cepstral Coefficients (MFCCs) and Beat Features affect classification. Based on parameters optimized for TMM genres, classification performances were evaluated against three groups of human subjects: experts, trained and untrained. Performances of both machine and human were shown to be comparable.**

Keywords: **Genre Classification, Feature Extraction, Beat Feature, Music Information Retrieval, Traditional Malaysian Music**

1. INTRODUCTION

Interest on music information retrieval systems for the storage, retrieval and classification of large collections of digital musical files has grown in recent years. Metadata such as filename, author, file size, date and genres are commonly used to classify and retrieve these documents. Such manual classification is highly labour-intensive and costly both in terms of time and money (Dannenberg, Foote, Tzanetakis & Weare 2001).

An automatic classification system that is able to analyse and extract implicit knowledge of the musical files is therefore highly sought. One approach to musical classification that is currently being widely studied is classification by musical genres. Musical genres are labels created and used by humans for categorizing and describing music (Tzanetakis & Cook 2002). Examples of a few Western musical genres include Pop, Rock, Hip-hop, and Classical. Several systems for automated genre classification and retrieval of musical files have been researched and developed (Tzanetakis & Cook 2002; Wold, Blum, Keislar & Wheaton 1996; Aucoturier & Pachet 1993). However, most of these studies were conducted using only western dataset. This forms the main motivation in incorporating other non-Western musical genres, specifically Traditional Malaysian Music (TMM), in this study.

TMM encompasses all traditional music from Malaysia, both West Malaysia and Sabah and Sarawak (Mohd Ghouse 1992), e.g. *Dikir Barat, Joget, Wayang Kulit, Gamelan, Etnik Sabah* and *Inang*. In general, these musical genres have a strong sense of rhythm (Musical Malaysia 2005), partly due to the fact that TMM is traditionally played by ear as opposed to reading from written musical scores. Having the beat or rhythm clearly audible helps when the musical piece is being passed down orally through generations in the villages such as having clear gong hits. TMM is further discussed in section 3.

In general, the process of music genre recognition includes two main steps – feature extraction and classification. Feature extraction is a process where a segment of an audio is characterized into a compact numerical representation. Examples of audio features are STFT, MFCC, Low-Energy feature, etc. Once the features are extracted, standard machine learning techniques can be applied to initiate classification. Classification is the process whereby unlabeled instances are mapped unto different set of categories by making accurate predictions based on past observations. Further readings can be found discussed in Witten and Frank (1999).

This paper investigates the factors and audio feature set combinations affecting classification of TMM genre classification. The paper is organised as follows: Section 2 presents a discussion of related research. Section 3 gives a brief overview of the musical structures of TMM. The experimental framework is discussed in Section 4. Results and discussion follow in Section 5.

2. PREVIOUS RESEARCH

Digital audio in general is categorized as speech, music and noise. Wold, Blum, Keislar and Wheaton (1996) analyse and compare audio features such as rhythms, pitch, duration, loudness and instrument identification to classify various groups of audio such as speech, gender, animal sounds and sound effects. However, music classification was not emphasized in their study.

Classification of audio based on musical genres was studied by Tzanetakis and Cook (2002) and Aucoturier and Pachet (1993). Both studies categorized audio features into three categories; timbral related features, rhythmic related features, and pitch related features. Audio data that are to be classified cannot be represented as raw audio data, such as samples of amplitude value of digital audio signals. Hence, some form of parameterisation is required. Parameterisation of audio data is based on audio analysis, which can be done using several methods such as Fourier transform, wavelet transform, statistical methods and so on (Wieczorkowska 2000).

Timbral related features are based on the Short Time Fourier Transform (STFT), which is used to determine the phase content of short local sections in a signal as it changes over time. The features are used in music-speech discrimination and speech recognition. Examples of timbral related features are such as spectral centroid, spectral roll off and time domain zero crossing, which measure the spectral shape, the changes in spectral shape and the noisiness of a signal respectively. Another feature, Mel-Frequency Cepstral Coefficients (MFCC), is also based on the STFT, but is typically used to provide compact representation of the spectral envelope, especially in speech recognition. Tzanetakis and Cook (2002) provide a detailed discussion on timbral related features and the account of their experiment.

Beat features analyses the signals that calculate the rhythmic structure of music based on their wavelet transform (Tzanetakis, Essl & Cook 2001). It involves time-frequency analysis, which is useful to music classification as its algorithm is similar to human hearing. The main beat can be defined as the regular periodic sequence of pulses corresponding to where a human would tap his foot whilst listening to music. Although achievable, extraction of beat features is very difficult. Whilst it is trivial for human to do so to a music, it is not so with machines. Kosina (2002) and Dixon (1999) give good overviews on beat tracking methods.

Pitch content feature extraction is based on multiple pitch detection techniques. Pitch features convert frequency into musical pitches so that it corresponds to a musical note with a specific pitch, where the musical notes are then labelled using the MIDI note-numbering scheme (Tzanetakis & Cook 2002). Although they can be useful, their pitch content cannot characterize musical genres fully. The conversion to MIDI is also unnecessary sometimes in certain genres where its tonal scale is different (Suyoto & Uitdenbogerd 2004). This is certainly applicable to TMM, where certain genres are based on microtonality. Therefore, pitch features are not investigated further in this study.

Li and Tzanetakis (2003) investigate the effects of different feature sets combinations for optimum classification performance. Features incorporated in the study were: FFT, MFCC, Pitch and Beat. Although optimum feature sets combinations from this study was obtained, it was also suggested that they might not be generic to all genres but applicable only to western genres that was used in their study.

Several studies investigate the use of existing classifiers for musical genre classification (Tzanetakis & Cook 2002; Li & Tzanetakis 2003). Classifiers vary in terms of robustness, speed, memory usage and complexity. For instance, OneR is a primitive form of classifier as it produces simple rule based on one attribute only, but it is useful in determining a baseline performance as a benchmark for other learning schemes (Witten & Frank 1999). Emphasis on the importance of understanding different classifiers is also discussed at length by Kaminskyj (2000).

Lippens, Martin, Mulder and Tzanetakis (2004) observe the viability of automated classification by comparing its result against classification by humans. Using the same dataset, he found that automated classification by far outperforms random classification, but human classification performed better by 20%. In a separate research, which evaluated human performance in classifying musical genres, the result obtained was 70% correct (Perrot & Gjerdigern 1999), whilst automated classification result by Tzanetakis and Cook (2002) was modest at 60%.

3. TRADITIONAL MALAYSIAN MUSIC

Traditional Malay music are mainly derivative, influenced by the initial overall Indian and Middle Eastern music during the trade era and later from colonial powers such as Thailand, Indonesia, Portuguese and British who introduced their own culture including dance and music. A thorough overview on the origin and history of TMM can be found in (Matusky 1993). The taxonomy of TMM depends on the nature of the theatre forms they serve and their instrumentations. Categorization of TMM genres has been studied extensively by Ang (1998). Music of these genres is usually disseminated non-commercially, usually performed by persons who are not highly trained musical specialists, undergoes change arising from creative impulses and exists in many forms (Nettle 1993). The musical ensembles usually include *gendangs* or drums that are used to provide constant rhythmic beat of the songs and gongs to mark the end of a temporal cycle at specific part of the song (Becker 1968).

One common attribute that is shared by most TMM genres is that they are generally repetitive in nature and exist in 'gongan'-like cycle. 'Gongan' is defined as a temporal cycle marked internally at specific points by specific gongs and at the end by the lowest-pitched gong of an ensemble (Matusky 1993). It is an important structural function as it divides the musical pieces into temporal sections. Once every measure has been played, musicians continue playing in a looping motion by repeating the cycle from the beginning again until one of the lead percussionists signals the end of the song by varying their rhythms noticeably. Traditional Malaysian music does not have a chorus that plays differently than other parts of the songs, which is the usual occurrence in western music. Its repetitiveness and constant rhythms are two aspects that are taken into account to facilitate classification by genre later.

4. EXPERIMENTS

The aim of this study is to investigate the factors and audio feature set combinations affecting classification of TMM. This study is carried out through supervised machine learning, where a percentage of data is divided into training data and the remaining used for testing.

4.1 Dataset

The dataset consists of TMM genres with a few western musical genres included in the preliminary study. The dataset collection method is described below.

4.1.1 Dataset Collection

Fifteen genres were involved in this study; ten of which were of TMM and five of western, totaling up to 330 musical files. The breakdown for each genre and its number of musical files are listed in Table 1.

A relatively small dataset was used in this experiment due to the difficulty in obtaining digital files of TMM, as traditional Malay musical culture is fast corroding with little preservation in digital format (Shriver 2003). Whilst it was much easier to obtain dataset for western music, the number was also kept small to match the size of TMM dataset.

For most of the experiments where track length and location were not the controlled factors, the length of each file was kept at 30 seconds each and started from the beginning of the song, replicating the experimental set up by Tzanetakis and Cook (2002). Approximately 2 hours 45 minutes of music were used. Musical files for this experiment were obtained from the Malaysia National Arts Academy, Sultan Salahuddin Abdul Aziz Shah's Cultural and Arts Centre at Universiti Putra Malaysia, Student's Cultural Centre at Universiti Malaya and also personal collections of audio CDs from many individuals.

4.1.2 Dataset Treatment

The dataset became available in both digital and analog format. Quite a number of musical data for TMM genres were in analog format and were digitized manually. All of the digital music files were then converted into wav files; the only audio format supported by the existing feature extraction tool used at the time of study. The whole dataset was later trimmed to specific length and location in the file by executing certain audio commands through batch processing before extraction began.

Table 1. Overall number of musical files for each genre

	Division	Genre	Number
1	Traditional Malay	*Dikir Barat*	31
2	Traditional Malay	*Etnik Sabah*	12
3	Traditional Malay	*Gamelan*	23
4	Traditional Malay	*Ghazal*	17
5	Traditional Malay	*Inang*	10
6	Traditional Malay	*Joget*	15
7	Traditional Malay	*Keroncong*	43
8	Traditional Malay	*Tumbuk Kalang*	13
9	Traditional Malay	*Wayang Kulit*	17
10	Traditional Malay	*Zapin*	10
11	Western	Blues	30
12	Western	Classical	30
13	Western	Jazz	19
14	Western	Pop	30
15	Western	Rock	30
		Total	**330**

4.2 Genre Classification Components

This section discusses feature extraction and classification using two software tools: 1) Musical Research System for Analysis ad Synthesis (MARSYAS) and 2) Waikato Environment for Knowledge Analysis (WEKA), respectively.

4.2.1 Feature Extraction

The features were extracted from the music files through MARSYAS-0.2.2; a free framework that enables the evaluation of computer audition applications. MARSYAS is a semi-automatic music classification system that is developed as an alternative solution for the existing audio tools that are incapable of handling the increasing amount of computer data (Tzanetakis & Cook 2000). It enables the three feature sets for representing the timbral texture, rhythmic content and pitch content of the music signals and uses trained statistical pattern recognition classifiers for evaluation. The feature extractor will produce numerical outputs in the form of Attribute Related File Format (ARFF) files. These ARFF files are later plugged into WEKA for classification as described below.

4.2.2 Classification

The ARFF files were used with selected classifiers using WEKA (Waikato Environment for Knowledge Analysis). WEKA is a workbench offering various dataset processing tools and collections of machine learning schemes. It enables pre-processing, classifying, clustering, attributes selection and data visualizing (Witten & Frank 1999). In this experiment, the J48 classifier was utilized and applied with stratified ten-fold cross validation, as it provides fast classification and was relatively easy to use compared to other classifiers as suggested by Witten and Frank (1999). Results from using WEKA are presented in the form of confusion matrices where the predicted value is plotted against the actual value to get the percentage of correctly classified instances. Classifier results can be evaluated in terms of accuracy and reliability (Pandya & Macy 1996). However, in this study, only the classification accuracy is focused.

4.3 Experimental framework

The study was conducted in three phases. The first phase was done to identify the factors that affect automatic genre classification of TMM. The second phase deals with analysing the feature sets that is significant in improving classification accuracy. The last phase of the study involves a comparison study between automated performances versus human.

4.3.1 Preliminary

The factors listed in Table 2 below were investigated in this preliminary study. Detailed descriptions of the experimental framework can be found in (Norowi, Shyamala & Wirza 2005).

Table 2. Factors and values investigated in the study

	Factors	Values
1	Dataset size	10 songs, 30 songs
2	Dataset Length	10 seconds, 30 seconds, 60 seconds
3	Dataset Location	00:00:00, 00:01:00
4	Classifier	ZeroR, OneR, J48
5	No. of cross-validation folds	3, 5, 10 folds

In general, larger dataset yields better classification result. However, dataset size should not be too large as it brings in possibilities of noise and outliers (Witten & Frank 1999). Hence, the effect of two dataset sizes was investigated, at ten and thirty songs per genre. For this preliminary study, both western and traditional Malaysian music genres were included together.

The second set focused on finding whether the track length has any significant role on classification. The lengths were kept at 10 seconds, 30 seconds and 60 seconds each. The track lengths were measured in seconds rather in 'gongan' cycles in order to follow the framework by Tzanetakis and Cook (2002), which had kept their measurement in seconds. The third set studied the effect of track location, where one batch had songs excerpts that started from zero second and the other at sixty seconds. The remaining of the sets dealt with classification parameters such as the classifier utilized and the number of cross-validation folds applied.

4.3.2 Feature set analysis

This phase of the study analyses the effect of several audio feature sets on classification results. Five different feature set combinations were investigated as shown in Table 3. The feature set combinations used were similar to those studied by Li and Tzanetakis (2003). However, their study was performed on western dataset. In our study, we investigated the effect of feature set combinations on TMM. Although they had found that beat feature had not significantly improved classification, we wish to see if this behaviour is similarly applicable when beat feature is utilized with TMM dataset. In addition, pitch feature was not investigated due to microtonality that can be found of some TMM genres as previously explained in section 2.

Table 3. Features tested

	Feature Set
1	STFT
2	MFCC
3	STFT + Beat Feature
4	MFCC + Beat Feature
5	STFT + MFCC + Beat Feature

4.3.3 Performance Evaluation

In this final phase of study, human subjects were included to compare their performances with automated classification. Only TMM genres were involved. However, unlike western music, where most listeners can generally classify songs into several different genres, traditional Malaysian musical genres are not as familiar to Malaysians. Shriver (2003) reasons that the fast shifting of the cultural milieu towards westernisation in Malaysia is responsible towards the de-emphasis of traditional Malaysian music. A performance comparison between machine and human will not be valid when tested against random Malaysian. Hence, three groups of human subjects were set up to validate this study.

Subjects from the first group all had formal musical background in TMM, where twelve co-eds students of the Malaysian National Arts Academy had participated. Subjects from the second and third groups were non-musicians, all of which were staff and postgraduate students at the faculty, with age ranging from early twenties to late fifties. To imitate a training-testing process that took place with automated classification, the second group received a basic introduction to TMM genres before the testing began. The last group received none.

5. RESULTS

The results of all three phases of the experimentation are presented as follows.

5.1 Phase 1 – Identifying the Factors Affecting Classification

The result of these factors are discussed below:

1) The larger dataset size, the better

At 10 songs per genre, 84% of correct classification had been achieved. In comparison, 90% was achieved at 30 songs per genre (Figure 1). Based on this, the dataset collection has been extended more aggressively, where the total number of dataset obtained is presented in Table 1 and is used in subsequent experiments.

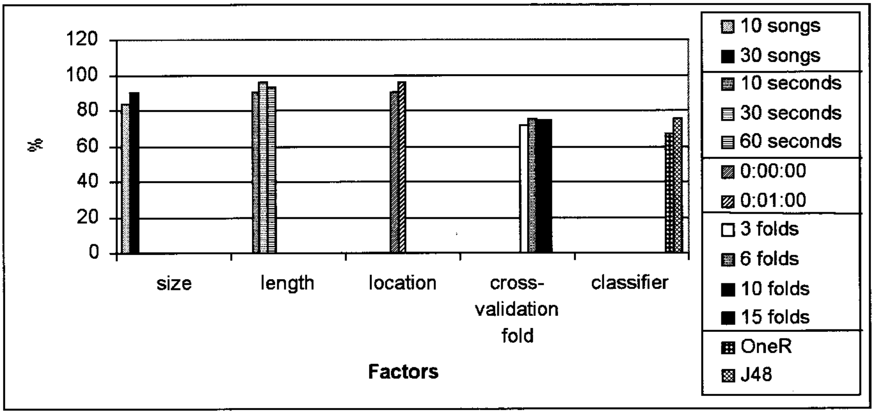

Figure 1. Classification performances between two different dataset sizes

2) Longer dataset length are not necessarily better

Initially, it was thought that longer musical excerpt yields higher result, but it was not always the case. 90% of correct classification was achieved with dataset of 30 seconds long compared to 96% with excerpts of 10 seconds long (Figure 1). A slight increase did occur when using excerpts of 60 seconds long (93%), but at a higher storage and computational cost. 30 seconds excerpts was accepted as a justified length as it returned reliable results and commonly used by many researches in their experiments (Tzanetakis & Cook 2002).

3) The beginning of a track is better than mid track for dataset location

Classification result was better on tracks that started at zero second (96%) compared to tracks that started at sixty seconds (90%) (Figure1). One possible explanation is that during the chorus, all the instruments are played at once, causing the overall energy of the audio signal to be higher, which leads to misclassification e.g. classical being misclassified as rock.

4) Increasing cross validation fold does not always yield better results
At three folds, 72% of the instances were correctly classified. When increased to six folds, the performance had increased to 75%. However, at ten folds, the performance dropped slightly to 74%, and remained constant when tested with fifteen folds and higher (Figure 1). This suggests that after a certain optimum point, increasing the number of cross validation folds would not aid classification performance. With this small dataset, six-folds is seen to be the optimal number. However, ten folds is still reliable and most widely used for basic classification.

5) Specific classifiers yield better results than others
Three different classifiers were tested, ZeroR, OneR and J48 to demonstrate this. With ZeroR and OneR, results obtained were just above satisfactory at 52% and 67% respectively whilst 75% was obtained with J48 classifier (Figure 1). J48 was a better classifier to use than the other two primitive classifiers.

5.2 Phase 2 – Analysis of the Effect of Beat Feature on Classification

Figure 2 shows that there is no improvement when beat feature is utilized in classifying western genres, except with Rock. Since Rock is rhythmic in nature, this result suggests that rhythmic music, as is the case with TMM, would benefit from the addition of beat features.

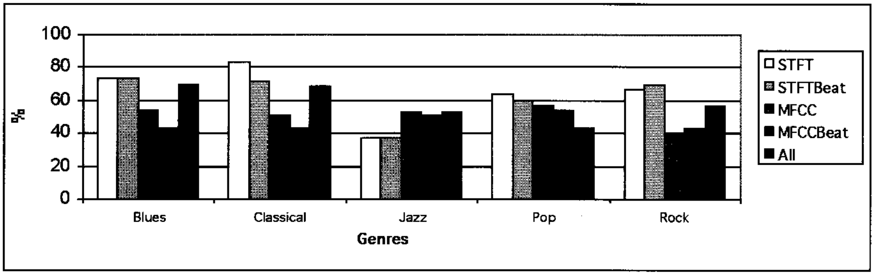

Figure 2. Effect of Feature Sets on Western Musical Genres

Figure 3 shows that the combination of STFT + MFCC + Beat features, in comparison to other feature combinations, had returned the highest classification result for most TMM genres. Furthermore, when other features were combined with Beat feature, better results were obtained (i.e. STFT with Beat and MFCC with Beat produced better results than STFT and MFCC alone). Seven TMM genres had shown positive results with the addition of Beat feature, indicating that it is important in classification of TMM. This is expected; due to the rhythmic, repetitive and clearly audible beat pulses of TMM, which is different than western music.

Figure 4 illustrates the importance of certain features in classifying TMM. The average percentage of correctly classified instances across ten genres was measured.

It is clear that the feature combinations STFT + MFCC + Beat achieved the highest classification result at 69.1%, followed closely by MFCC + Beat and MFCC (both 67.5%), STFT (64.9%) and STFT + Beat (62.8%). With 69.1% at its highest, it shows that the STFT + MFCC + Beat feature combination is slightly better than the 61% obtained by Tzanetakis and Cook (2002) and indirectly comparable to the result by Perrot & Gjerdigen (1999) which deduced that human classification of musical genre is approximately 70% correct. Genres such as *Dikir Barat, Ghazal,* and *Keroncong* are amongst the top three genres that classified well whilst *Inang* and *Zapin* are the worst. Genres such as *Inang* and *Zapin* are quite broad in nature and lack constant rhythm, which is a possible contribution to the low score. The same pattern is seen with Jazz in western music, where it is prone to misclassification in comparison to a more steadily rhythmic hip-hop (Tzanetakis & Cook 2002).

5.3 Phase 3 – Performance Evaluation

The result for classification performance on both machine and human in general is comparable. Tested on the very same dataset, the experts scored 81.8% followed by the training group at 70.6%. The untrained group scored 49.6%. Interestingly, automated classification outperformed the untrained group and random classification at 66.3% and 10% respectively (Figure 5). Its performance is also very close to the trained group, implying that the training-testing process in human is just as important as in automated classification.

However, one other issue became apparent from the results – the average Malaysian did not seem to know their traditional musical genres. The results seem to re-emphasize the issue that traditional Malaysian music is a fast disappearing art. Seeing that not many Malaysians are familiar in classifying their traditional musical genres, an automatic traditional Malaysian musical genre classification system is certainly needed!

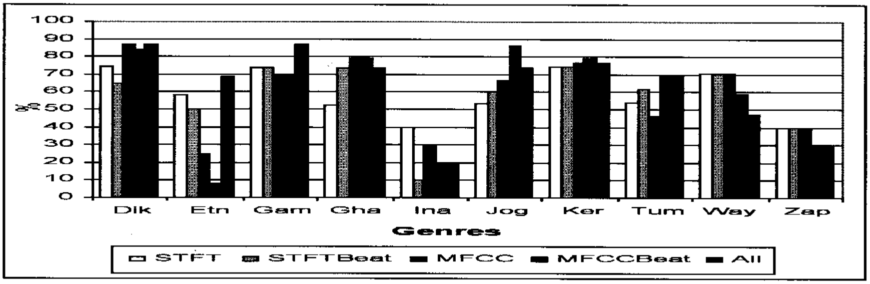

Figure 3. Effect of Feature Sets on Traditional Malaysian Musical Genres

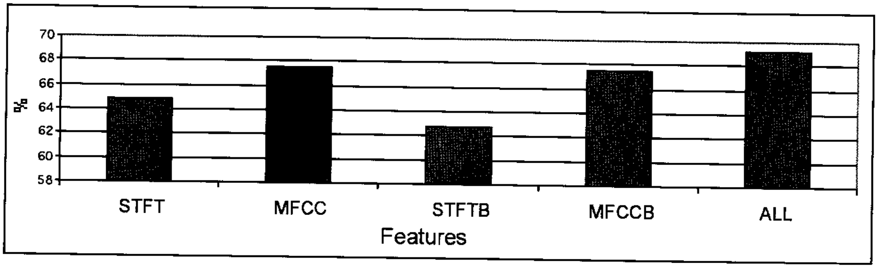

Figure 4. Feature Sets in Order of Importance

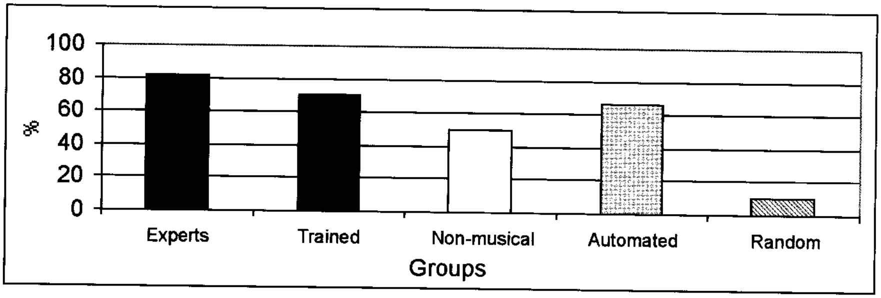

Figure 5. Feature Sets in Order of Importance

6. CONCLUSION

This study has shown that a beat feature, in addition to existing features, is important in improving genre classification of TMM. This had been anticipated, as the musical structure of TMM is different than western music. It was also found that classification results can be improved by taking into consideration the factors such as dataset size, track lengths, etc., as previously discussed. The effect of different feature sets on classification results of TMM genres were also presented. It was found that automated classification is comparable with human classification. Although difficult, the dataset size needs to be increased. However it is hoped that this small-scale study will trigger a larger scale research in this direction in the near future.

10. REFERENCES

Ang, M. 1998, 'A Layered Architectural Model for Music: Malaysian Music on the World Wide Web', Ph.D. dissertation: UPM, pp. 371.

Aucoturier, J., and Pachet, F. 1993, 'Representing Musical Genre: A State of the Art', *Journal of New Music Research.* vol. 32, no. 1, pp. 83-93.

Becker, J. 1968, 'The Percussive Patterns in the Music of Mainland Southeast Asia', *Ethnomusicology,* vol. 2, no. 2, pp. 173-191.

Dannenberg, R., Foote, J., Tzanetakis, G. and Weare, C. 2001, 'Panel: New Directions in Music Information Retrieval.' *In Proceedings of the 2001 International Computer Music Conference,* International Computer Music Association, pp. 52-59.

Dixon, S. 1999. 'A Beat Tracking System for Audio Analysis', *In Proceedings of the Diderot Forum on Mathematics and Music,* Austrian Computer Society, pp. 101-110.

Kaminskyj, I. 2001, 'Multi-feature Musical Instrument Sound Classifier', *Mikropolyphonie, WWW Journal,* Melbourne, Australia, Vol.6.

Kosina, K. 2002, 'Music Genre Recognition', Diplome Thesis, Hagenberg University.

Li, T. and Tzanetakis, G. 2003, 'Factors in Automatic Musical Genre Classification of Audio Signals', *IEEE Workshop on Applications of Signal Processing and Audio Acoustics,* WASPAA 2003, pp. 143-146.

Lippens, S., Martens, J., Mulder, T. and Tzanetakis, G. 2004, 'A Comparison of Human and Automatic Musical Genre Classification', *IEEE International Conference on Acoustic Speech and Signal Processing* (ICASSP), Montreal, Canada.

Matusky, P. 1993, 'Malaysian Shadow Play and Music: Continuity of an Oral Tradition', Kuala Lumpur: Oxford University Press.

Mohd Ghouse Nasuruddin. 1992, 'The Malay Traditional Music', Kuala Lumpur: Dewan Bahasa dan Pustaka.

Musical Malaysia – An Educational Resource (web site) 2006, Available at: http://www.unitedcoopyorkshirebrassband.com

Nettl, B. 1993, 'Folk Music', Microsoft Encarta'94. Redmond, WA: Microsoft Corporation.

Norowi, N.M., Doraisamy, S. and Wirza, R. 2005, 'Factors Affecting Automatic Genre Classification: An Investigation Incorporating Non-Western Musical Forms', *In Proceedings of the 6th International Conference on Music Information Retrieval (ISMIR 2005),* pp. 13-21.

Pandya A.S. and Macy, R.B. 1996, 'Pattern Recognition with Neural Networks in C++', CRC Press.

Perrot, D. and Gjerdigen, R. 1999, 'Scanning the Dial: An Exploration of Factors in Identification of Musical Styles', *In Proceedings of Society for Music Perception and Cognition.* pp. 88 (abstract).

Shriver, R. 2000, 'Digital Stereo Recording of Traditional Malaysian Musical Instruments', *Audio Engineering Society Convention Paper.*

Suyoto, I. and Uitdenbogerd, A. 2004, 'Exploring Microtonal Matching', *In Proceeding of 5th International Conference on Music Information Retrieval (ISMIR 2004).*

Tzanetakis, G. and Cook, P. 2000, 'MARSYAS: A Framework for Audio Analysis', *Organized Sound,* Cambridge University Press, vol. 4, no.3.

Tzanetakis, G. and Cook, P. 2002, 'Musical Genre Classification of Audio Signals', *IEEE Transactions on Speech and Audio Processing,* vol. 10, no. 5.

Tzanetakis, G., Essl, G. and Cook, P. 2001, 'Automatic Musical Genre Classification of Audio Signals', *In Proceeding of International Symposium of Music Information Retrieval (ISMIR).*

Wieczorkowska, A. 2000, 'Towards Musical Data Classification via Wavelet Analysis', *12th International Symposium on Methodologies for Intelligent Systems,* pp. 292-300.

Witten, I. and Frank, E. 1999, 'Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations', Morgan Kauffman Publisher.

Wold, E., Blum, T., Keislar, D. and Wheaton, J. 1996, 'Content-based Classification, Search, and Retrieval of Audio', *IEEE Multimedia,* vol.3, no.3, pp. 27-36.