

Spectral-based Features Ranking for Gamelan Instruments Identification using Filter Techniques

Aris Tjahyanto¹, Yoyon K Suprpto², Diah P Wulandari³

¹Information Systems Department, ^{2,3}Electrical Engineering Department

Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

e-mail: atjahyanto@gmail.com, yoyonsuprpto@gmail.com, diah_basuki@yahoo.com

Abstrak

Pada paper ini, kami menjelaskan upaya dalam menentukan ranking fitur berbasis spektral dengan memanfaatkan teknik filter yang digunakan untuk identifikasi instrumen gamelan Jawa. Model yang dipakai mengekstraksi sekelompok fitur berbasis spektral dari sinyal suara gamelan dengan menggunakan Short Time Fourier Transform (STFT). Ranking dari fitur ditentukan dengan memanfaatkan lima algoritma, yaitu ReliefF, Chi-Squared, Information Gain, Gain Ratio, dan Symmetric Uncertainty. Selanjutnya kami menguji ranking fitur secara validasi silang dengan menggunakan Support Vector Machine (SVM). Eksperimen menunjukkan bahwa algoritma Gain Ratio memberikan hasil terbaik, yaitu menghasilkan akurasi sebesar 98.93%.

Kata kunci: support vector machine, transkripsi otomatis, Gain Ratio, ekstraksi fitur

Abstract

In this paper, we describe an approach of spectral-based features ranking for Javanese gamelan instruments identification using filter techniques. The model extracted spectral-based features set of the signal using Short Time Fourier Transform (STFT). The rank of the features was determined using the five algorithms; namely ReliefF, Chi-Squared, Information Gain, Gain Ratio, and Symmetric Uncertainty. Then, we tested the ranked features by cross validation using Support Vector Machine (SVM). The experiment showed that Gain Ratio algorithm gave the best result, it yielded accuracy of 98.93%.

Keywords: support vector machine, automatic transcription, Gain Ratio, features extraction

1. Introduction

Feature selection is a process of finding an optimal feature subset, removes irrelevant or redundant feature. Feature selection is one of the important steps in machine learning especially for recognition tasks. The performance of recognition algorithms are usually dependent on the quality of the feature set. If the feature set contains redundant or irrelevant features, the algorithm may produce a less accurate or a less recognition rate. The feature selection problem has been studied by the statistics and machine learning communities for many years [1-4]. The feature selection algorithms can be categorized as filter, wrapper, and embedded methods based on the criterion functions. Filter methods uses statistical properties for evaluating feature subsets. The advantages of filters methods are fast and efficient to process high dimensional datasets, however filters approach do not consider the feature dependencies. Wrapper methods use a learning algorithm for evaluating the selected feature subsets. Embedded methods are similar to wrapper methods, but less computationally expensive and considering feature dependencies [5]. Feature extraction can be viewed as finding a subset of raw data while reducing the dimensionality.

Many algorithms have been developed to perform audio feature extraction; common methods such as temporal based and spectral based using Fast Fourier Transform (FFT), Short Time Fourier Transform (STFT), Discrete Wavelet Transform (DWT), and Continuous Wavelet Transform (CWT). There are various features have been proposed for audio signal, such as zero crossing rate, RMS energy, envelope, and spectrum representation [6]. We used a set of spectral-based features which has been previously developed for gamelan instruments identification [7].

There are several approaches has been developed to estimate the pitch and instruments in the automatic music transcription. We can use autocorrelation function [8] for identifying hidden periodicities in a time-domain signal. The autocorrelation function shows the peaks periodicity in a signal. Suprpto et al [9] [10] introduced a method to generate music transcription for gamelan using spectral density model to extract the waveforms of gamelan instruments sound using Adaptive Cross Correlation (ACC).

Another technique is pattern recognition approach that requires a set of features to identify the musical instruments [11] [12]. The common features that needed for recognition process such as pitch, frequency modulation, spectral envelope, spectral centroid, intensity, amplitude envelope, amplitude modulation, onset asynchrony and in harmonicity.

The goal of this paper is to get the minimal spectral-based features subset that extracted from gamelan recording using STFT. The selected features subset then validated by cross-validation techniques using support vector machine (SVM). There are two main reasons for addressing this tasks using SVM. First, accurate recognition of gamelan instrument is itself an important for automatic transcription. Second, because of the effectiveness of SVM [13] [14] and recently became one of the most popular recognition or classification methods. SVM have been used in a variety of applications such as text classification [15], facial expression recognition [16], gene analysis [17], [18] and many others.

In the proposed approach, Javanese gamelan instruments identification is accomplished through identification of individual blades or keys using an SVM classifier. Javanese gamelan is an ensemble of percussion instruments that mostly metallophone [19], xylophones, and gong type instruments which produce tones when struck with horn or wooden mallets. A complete set of gamelan consist of 72 instrument [20], for example: kendang, saron groups, bonang groups, kethuk-kenong and gongs. Group of saron consist of demung, saron, and peking. Those instruments play the core melody or balungan gendhing. Gamelan is one of percussion type musical instruments which do not produce harmonic sounds [21]. However, because of the handmade production, gamelan still produce the frequencies of non-integer overtone [22]. The frequency range of saron groups [7] can be seen at Table 1. Individual gamelan pitch are sometimes difficult to identify due to their overlapping in frequency, for example fundamental frequency of saron '1' equals to that of demung '1H'.

The rest of this paper is organized as follows. Section 2 describes the research method how to get the optimal spectral-based feature subsets. Section 3 presents our experiments and discuss the results. Finally, Section 4 gives conclusions of our experiments.

Table 1. Saron group frequency range

Keys	Fundamental Frequency (Hz)		
	Demung	Saron	Peking
6L	231	463	925
1	267	533	1062
2	307	613	1225
3	349	698	1400
5	402	805	1599
6	463	925	1858
1H	533	1062	2158
2H	613	1225	2477

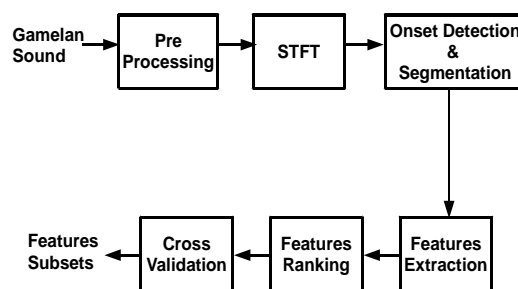


Figure 1. Flowchart of the research method

2. Research Method

A general view of the flowchart of the proposed system is depicted in Figure 1. The output of the proposed system is the selected feature subset for identifying the gamelan instruments. The first stage in our proposed system is preprocessing. Before a gamelan recording is subjected to the proposed methods, it is preprocessed in some way in order to make the following task easier.

The preprocessing consists of noise reduction, low-pass filtering, and sampling rate conversion. The second step is to create time-frequency representation or spectrogram from a gamelan recording. The 2D matrices spectrogram of the given gamelan recording is calculated

by the short-time Fourier transform (STFT) using Hamming window with window size approximately 2048 samplings and hopsize 6%.

Before extracting the features set, segmentation in the time-frequency domain was performed. The process of segmentation for the time-frequency representation requires note onset information. Note onset can be detected using sudden changes of acoustic energy approaches [24]. In the case of strong gamelan note, this abrupt energy changing will be very sharp. We can find the onset location using the peak detection function [25]. The features set then calculated based on the segmented spectrograms. The features set should contain useful information for identifying and differentiating gamelan instruments. In this paper, we used 34 features for gamelan instruments identification tasks. The features [26] have been calculated and additional features have been extracted including the statistical properties like mean, variance of the spectral envelope.

We compared the five feature ranking algorithms of the filters approach. They are Information Gain, Gain Ratio, Chi-squared, Symmetrical Uncertainty and Relief. Ranking algorithms produce a ranked list, according to the evaluation of criterion function. For the sake of performance comparison, we also consider the cross validation accuracy. We calculated the cross validation accuracy in terms of SVM classifier.

2.1. Time-frequencies Analysis

The goal of automatic gamelan transcription is to extract the sequence of gamelan notes from gamelan recording. Gamelan notes are any system that represents the pitch of a gamelan sound. This paper is part of the project aims to develop a system that extracts note events from gamelan sounds spectrogram.

Spectrogram is a spectro-temporal representation of the sound. Spectrogram provides a time-frequency portrait of gamelan sounds. The STFT has been the commonly used method for generating time-frequency representations or spectrograms of musical signal. The result of STFT can be plotted on a 2D or 3D spectrogram (as shown in Figure 2) as a function of time and frequency, and magnitude is represented as the height of a 3D surface spectrogram or intensity in 2D spectrogram. However, STFT suffers from the common shortcoming that the length of the window determines the time and frequency resolution of the spectrograms [27] [21].

The size of the window used for STFT is related to the time resolution and frequency resolution. If we apply a short window, we will have good time resolution. However, if we implement a long window, we will get high frequency resolution but low time resolution. For pitch analysis such as automatic gamelan note transcription, the frequency resolution of the spectrograms is more important than the time resolution [21]. Then STFT with long window is good enough for automatic gamelan note transcription.

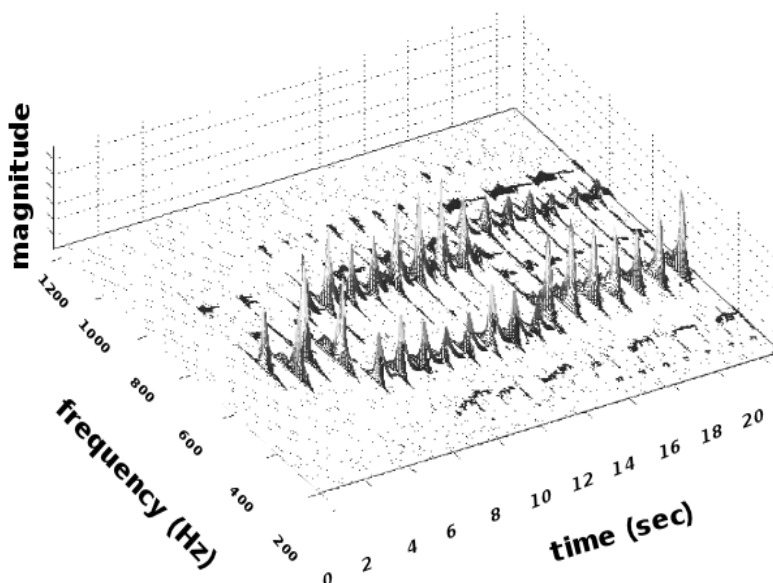


Figure 2. 3D surface spectrogram of gamelan recording

2.2. Segmentation

Segmentation is an important process in automatic gamelan notes transcription that give significant effect for instruments recognition performance. In this paper we used simple segmentation method based on onset information. The simple approach is to find points where the magnitude of energy exceeds a local or global threshold (as shown in Figure 3). The segmentation task can be viewed as a process for finding the notes boundaries. Boundaries are determined at time-frequency domain by looking for the onset time of the notes in the amplitude envelope. An onset can be defined as the instant when the player strike the gamelan blade or the moment when a new note begins [27].

It is possible to distinguish different note onsets in a gamelan recording. For notes from gamelan instruments (such as demung, saron, peking, and bonang) the sudden change of energy produces hard onsets that are shown as an abrupt energy increments. The gamelan recording has a set of features that sometimes make the note boundaries diffuse. So it is hard to identify the boundaries between the notes, especially if the notes with locations that are in close proximity such as notes candidate for saron '3' and bonang '3' (as shown in Figure 3).

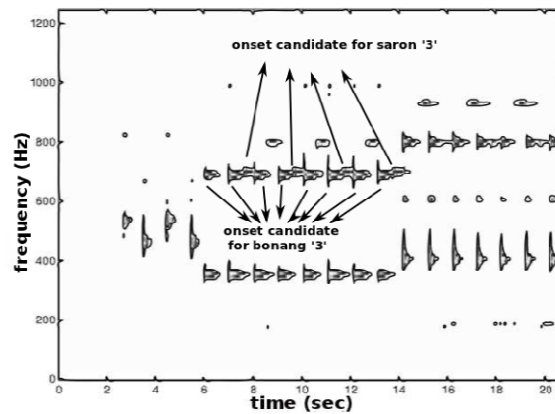


Figure 3. Thresholded spectrogram

After getting the spectrograms from gamelan recording, onset candidates or events corresponding to energy changes are detected. The algorithm looked for energy changes [28] from different frequency channel according to gamelan instruments. The most salient or prominent ones are considered as note onsets. The weak onset candidates are considered as a valid onset if their amplitude above a global threshold value (see Figure 4). The onsets detected by an onset detection algorithm will be used to segment the spectrogram.

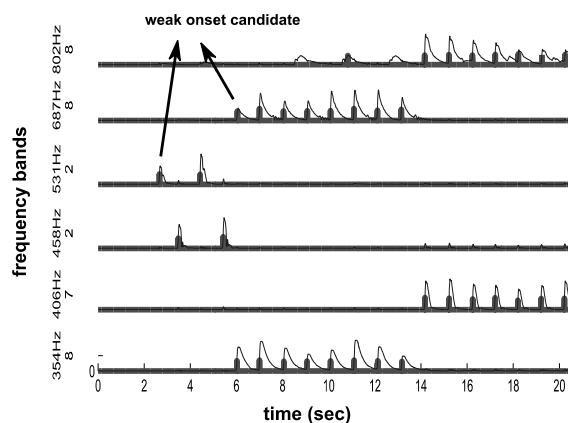


Figure 4. Onsets candidate

2.3. Feature Extraction

The main idea of feature extraction is to perform recognition process more effective and efficient, so the process requires a small and simple data space. To obtain the features set, we first segmented the spectrogram based on the onsets information (see Figure 5). Then the features such as spectral centroid, spectral flux, mean and variance of the segment will be extracted and calculated. Before performing feature extraction, it is important to decide what features should be included for recognition process.

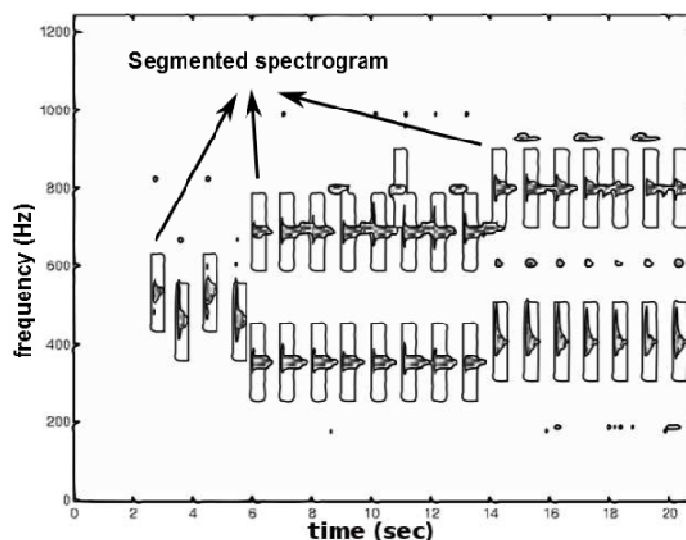


Figure 5. Onset based segmentation

Stan Z. Li et al [29] explained about perceptual features, mel-cepstral features and their combinations for audio classification. They used short time energy (STE), zero crossing rates (ZCR), brightness, and bandwidth for capturing the perceptual characteristics of the sounds. They also used cepstral coefficients (CC) for capturing the shape of the frequency spectrum of a sound. Based on the CC, most of the original sound signal can be reconstructed again, these features give a complement to the perceptual characteristics.

There are four feature sets [12] suitable for audio signal processing:

(i) Temporal features

- Autocorrelation coefficients (AC): signal spectral distribution in the time domain
- Zero crossing rates: using short windows (ZCR) and long windows ($IZCR$)
- Local temporal waveform moments, including the first four statistical moments (temporal centroid Tc , temporal width Tw , temporal asymmetry Ta , and temporal skewness Tk)
- Amplitude Modulation features (AM), meant to describe the "tremolo", "graininess" or "roughness"

(ii) Cepstral features

- Mel-frequency cepstral coefficients ($MFFC$), tend to represent the spectral envelope over the first few coefficients.

(iii) Spectral features

- A subset of features obtained from the first four statistical moments, namely the spectral centroid (Sc), spectral width (Sw), spectral skewness or spectral asymmetry (Sa), and spectral skewness (Sk)
- Audio Spectrum Flatness (ASF) and Spectral Crest Factors (SCF)

- Spectral slope (Ss), spectral decrease (Sd), spectral variation (Sv), spectral rolloff or frequency cutoff (Fc), and spectral flatness (So).
 - Frequency derivative of the constant-Q coefficients (Si)
 - Octave Band Signal Intensities ($OBSI$)
- (iv) Perceptual features
- Relative specific loudness (Ld), sharpness (Sh), and spread (Sp).

In this paper, we provide 34 spectral features, such as: fundamental frequency, spectral centroid (Sc), two spectral rolloff (Fc), spectral flux (SF), spectral skewness (Sa), spectral kurtosis (Sk), spectral slope (Ss), and spectral bandwidth (Sw). These features are then combined as a feature set of a gamelan sound. The feature set is normalized by dividing each feature component by a real number so the result is between -1 and 1. The normalized feature set is considered as the final representation of the gamelan sound.

Spectral skewness (Sa) is a measure of the asymmetry of the spectrum around the mean value. If $Sa < 0$ indicates more energy on the right side. If $Sa > 0$ indicates more energy on the left side. Spectral kurtosis K is a measure of the peakedness or flatness of the shape of power spectrum distribution. Positive kurtosis $K > 3$ indicates a peaked distribution, the standard normal distribution has a kurtosis $K = 0$, and negative kurtosis $K < 3$ indicates a flatter distribution [30]. Those features (spectral skewness, spectral moment, spectral kurtosis, and spectral entropy) were implemented using statistical function.

Spectral centroid (Sc) is a measure of the center of gravity of the spectrum. The spectral centroid is computed by multiplying the value of each frequency by its magnitude, then the sum of all these divided by the sum of all the magnitudes. The spectral centroid (Sc) [31] [29] [32] can be defined as Eq. (1),

$$Sc = \frac{\left(\sum_{i=1}^N f(i) \times M[f(i)] \right)}{\sum_{i=1}^N M[f(i)]} \quad (1)$$

where $M[f(i)]$ is the magnitude for the frequency f at bin i , N is the number of frequency bins.

Scheirer and Slaney defined the spectral rolloff point (Fc) as the 95th percentile of the power spectrum distribution [33]. Spectral rolloff is the frequency when 95% of the signal energy is contained. Spectral rolloff (Fc) is defined as Eq. (2),

$$\sum_{i=1}^{Fc} M[f(i)] \geq 0.95 \sum_{i=1}^N M[f(i)] \quad (2)$$

2.4. Feature Ranking

The goals of feature selection are improving computational efficiency but preserving or even increasing recognition rate. It becomes important to the success of the tasks that apply machine learning approach especially when the data have many irrelevant or redundant features. In general, the features selection algorithms can be categorized as wrapper approach and filter approach [34] [1].

The five filter-based feature ranking techniques being compared are described below. Those techniques are Information Gain (IG), Gain Ratio (GR), ReliefF (RF), Chi-Squared (CS) and Symmetric Uncertainty (SU), and available in the Weka data mining tool [44]:

- (i) **ReliefF (RF)** is an extension of the Relief algorithm developed by Kira and Rendell [36]. The main idea of Relief algorithm is to evaluate the worth of a feature or attributes based on how well their values can be used to distinguish among the instances. Relief algorithm cannot handle incomplete data and only limited to two-class problems. The ReliefF is the extended version of Relief. ReliefF can handle incomplete data and not limited to two class problems. However, if we apply the algorithm for a highly noisy data that have many irrelevant features and/or mislabeling, the performance of ReliefF can get worse [37].
- (ii) **Chi-Squared (CS)** can be used to evaluate the worth of a feature by calculating the value of the Chi-Squared with respect to the class. The null hypothesis is the assumption that the two features are unrelated, and it is tested by Chi-Squared formula from Plackett [38]. If we got a large value of CS, then we can determine that the feature is an important feature.
- (iii) **Information gain (IG)** can also be used for determining the feature rank. The main idea of *IG* is to select features based on entropy. Entropy is a measure of how mixed up or uncertainty or the disorder degree of a system. [39] [40]. *IG* measures the number of bits of information gained about the class prediction when using a given feature to support the prediction. Information gain [40] of the feature or attribute *A* is defined as Eq. (3),

$$IG(A) = E(C) - E(C | A) \quad (3)$$

where $E(C)$ is the entropy of classes C and $E(C | A)$ is the conditional entropy of C given A when the value of the attribute A is known.

- (iv) **The Gain Ratio (GR)** is an extended version of Information Gain. *GR* is more efficient and effective than Information Gain and can be used to evaluate the correlation of attributes with respect to the class concept of an incomplete data set in [41] [42] [35]. The gain ratio of A is defined as the information gain [40] of A divided by its intrinsic value $IV(A)$ using Eq. (4),

$$GR(A) = \frac{IG(A)}{IV(A)} \quad (4)$$

where $IV(A) = -\sum_{i=1}^k \left(\frac{|A_i|}{N} \times \log_2 \frac{|A_i|}{N} \right)$ which $|A_i|$ is the number of instances where

attribute A takes the value of A_i , k is the number of distinct values of attribute A , and N is the total number of instances in the dataset.

- (v) **Symmetric Uncertainty (SU)** is a correlation measure between the features and the class, and it is obtained by [44] [1] Eq. (5),

$$SU = 2 \frac{E(C) - E(C | A)}{E(A) + E(C)} \quad (5)$$

where $E(A)$ and $E(C)$ are the entropies based on the probability associated with feature A and class value C .

2.5. Cross Validation

As discussed in the previous section, we need to make a comparison of performance between different ranking approaches using cross validation method. Cross validation is a statistical method of evaluating and comparing learning algorithms by dividing data into two portion of data for training and validating or testing the model. The goal is to compare the performance of different ranking approaches and find out the best approach for the gamelan instruments recognition. Cross validation can also be used to understand the generalization power of a classifier.

We used 10-fold cross-validation procedure. The procedure randomly partitioned the N instances in the original dataset into 10 approximately equal sets of size $N/10$. Then 9 partitions are used for training and the last is used for testing. This whole process then repeated 10 times. In this research, cross validation were implemented using LibSVM [44]. LibSVM provides several kernel functions: linear, polynomial, radial basis function (RBF), and sigmoid. SVM is a popular tool for data classification or recognition. The main idea of SVM is to project the data into a high dimensional space and find a hyperplane that separate between the two class with the maximal margin [45].

To train an SVM, we must select the proper value of the kernel parameters. For RBF kernel, we need determine the value of γ and C that can be obtained using grid search based on the procedure Grid-Search [44]. Grid search can be conducted to choose the best parameters. The grid-search can be done by selecting the value of γ and C from the following sets: $\log_2(C) \in \{-3, -2, \dots, 12\}$ and $\log_2(\gamma) \in \{-6, -5, \dots, 10\}$. In this research, the gamelan instruments are classified into 31 classes, they are 7 demung tones, 7 saron tones, 7 peking tones, and 10 bonang tones. Classification of multiclass can be achieved by SVMs. There are two common approach for multiclass SVM: one-against-all (OAA) and the one-against-one (OAO) [13]. In OAA approach, an SVM is created for each class by differentiating the class against all classes ($M-1$). Then the number of SVMs created in OAA is M . All the N training data are used in constructing an SVM for a class. The SVM for class k is created using the set of training data and their outputs (x_i, y_i) . For OAO approach, an SVM is constructed for every pair of classes by training it to differentiating the two classes. The number of SVMs created in this approach is $M(M-1)/2$.

3. Experiments and Discussion

The database used in our experiments is composed of approximately 2790 segmented audio, collected from Elektro Budoyo ITS gamelan set. All audio are 16-bit, mono-channel, and frequency sampling 44100 Hz. The training data set consists of $N = 2790$ audio samples for $M = 31$ classes. We produced the sounds data samples by randomly hitting the keys or bars of metal with their own hammer at center, upper, and lower areas [46] as shown at Figure 6.

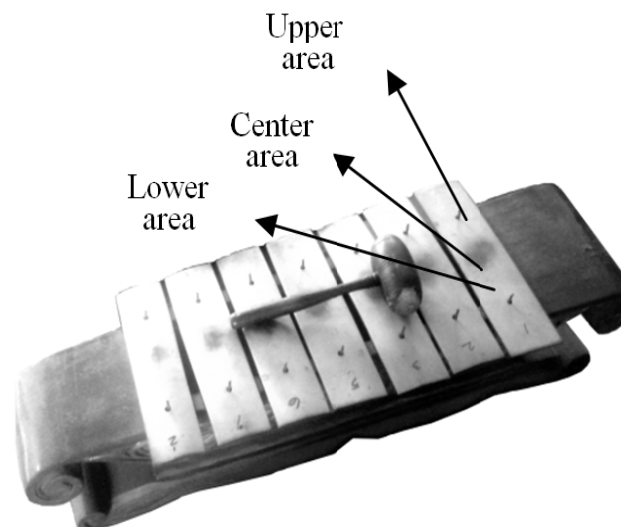


Figure 6. Different struck area for data collecting

After the feature data set was calculated and extracted, then we randomly partitioned the data into training data sets and testing data sets. The training features data were ranked in descending order using the five techniques. The ranking of features obtained for the training

data is presented in Table 3. The first 7 features are consistently ranked as the top. The first 4 features predicted by the five techniques gives the same results, although features 5, 6 and 8 are reversed in some rankings.

Table 2. Spectral-based features

No	Features	Number of features
1	Fundamental Frequency	1
2	Spectral Centroid	1
3-4	Spectral Rolloff	2
5	Spectral Flux	1
6	Spectral Skewness	1
7	Spectral Moment	1
8	Spectral Kurtosis	1
9	Spectral Entropy	1
10	Spectral Slope	1
11	Spectral Bandwidth	1
12	Mean	1
13	Standard Deviation	1
14	Mode	1
15	Median	1
16	Variance	1
17-25	Percentile	9
26-34	Quantile	9

For each ranking method, investigation of recognition accuracy on the testing data as a function of the features has been done in ascending order and descending order. Recognition rate or accuracy was taken from prediction accuracy performed by SVMs. Accuracy results as a function n number of features in ascending order are presented in Figure 7, for descending order are presented in Figure 8. We measured the performance for subsets consisting of the n ranked features. Where n varies between 1 and 34, started from the least important features for ascending order and from the most important features for descending order.

The SVM perform very well when all features or subsets of the original features are used. The peak accuracy was reached on the 19 until 22 best features in ranking by all techniques at accuracy of more or equal to 98.87%, and increasing the subsets did not improve the accuracy. Then the rest of the features can be deleted due to non-significant influence for the performance. Interestingly, the GR technique show the peak at accuracy of 98.93% (as shown at Table 5), the highest accuracy achievable using the five techniques.

Table 3. The first twenty of feature ranks on the spectral-based gamelan features; see description in the text; all entries denote feature numbers shown in Table 2

Methods	Feature Rank																			
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
CS	1	3	2	4	8	5	6	11	10	28	34	29	30	15	31	32	12	33	7	26
IG	1	3	2	4	5	8	6	11	34	28	32	29	33	31	15	30	7	10	27	12
SU	1	3	2	4	5	8	6	11	34	32	31	29	28	30	15	33	10	7	12	27
GR	1	3	2	4	8	5	6	11	31	32	18	14	17	30	15	29	10	34	33	28
RF	1	3	2	4	6	8	5	13	27	32	26	33	28	30	15	29	31	12	10	11

Table 4. Accuracy for gamelan dataset as a function of the worst n ranked features (ascending order); for $n=24..34$

Methods	Accuracy (%) for the worst n ranked features used for classification										
	34	33	32	31	30	29	28	27	26	25	24
CS	98.87	98.53	97.97	97.74	95.99	95.88	94.69	92.66	81.36	68.47	68.36
IG	98.87	98.53	97.97	97.74	95.99	95.37	94.69	92.66	81.36	79.66	79.60
SU	98.87	98.53	97.97	97.74	95.99	95.37	94.69	92.66	81.36	79.66	79.49
GR	98.87	98.53	97.97	97.74	95.99	95.88	94.69	92.66	81.36	81.41	81.41
RF	98.87	98.53	97.97	97.74	95.99	95.08	95.20	92.66	92.43	92.43	92.43

Figure 7 shows the degradation in the recognition rate or accuracy when the number of features subsets is reduced. A comparison of the five methods shows that the accuracy over 90% achieved with RF subsets are better than another results (see Table 4). All the techniques show the same behavior without any significant differences. The accuracy is almost same until the subsets are reduced to 26 or less features, then the accuracy tends to decrease with reducing the subsets (see Table 4 and Figure 7).

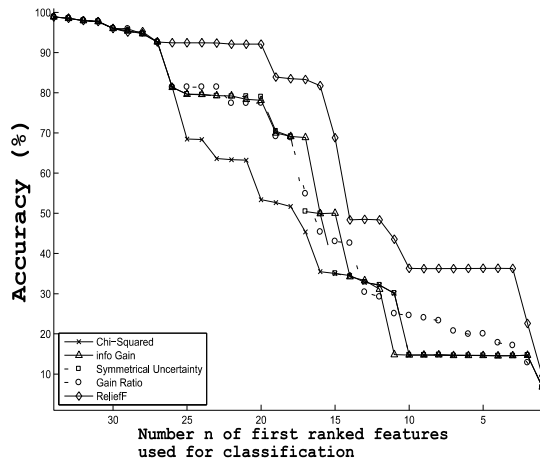


Figure 7. Accuracy for gamelan dataset as a function of the worst n ranked features (ascending order)

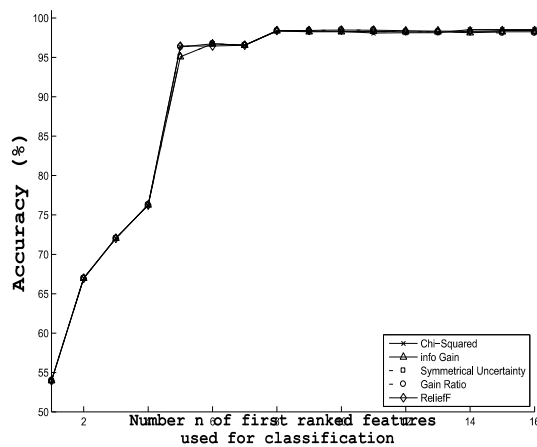


Figure 8. Accuracy for gamelan dataset as a function of the best n ranked features (descending order)

Table 5. Accuracy for gamelan dataset as a function of the best n ranked features (descending order); for $n=19...23, 30...34$

Methods	Accuracy (%) for the best n ranked features used for classification										
	34	33	32	31	30	...	23	22	21	20	19
CS	98.87	98.87	98.87	98.87	98.87		98.87	98.87	98.81	98.81	98.81
IG	98.87	98.87	98.87	98.87	98.87		98.87	98.87	98.87	98.81	98.53
SU	98.87	98.87	98.87	98.87	98.87		98.87	98.87	98.81	98.81	98.81
GR	98.87	98.87	98.93	98.93	98.93		98.93	98.93	98.93	98.59	98.59
RF	98.87	98.87	98.87	98.81	98.87		98.87	98.87	98.87	98.87	98.87

For descending order, the accuracy is quite stable until the subsets reduced to 7 or less features. The seven features are fundamental frequency, spectral roll off 40%, spectral centroid, spectral roll off 90%, spectral flux, spectral kurtosis and spectral skewness. The first best feature give accuracy of 53.96%, the second best features give 66.95%, the third best features give 72.03%, and the seven best features give accuracy of 96.55% (as shown at Figure 8).

4. Conclusion

In this paper, we have presented in details our approach to perform feature ranking using five filter-based ranking methods. Although they all perform in a similar way, accuracy of the SVM classifier has been significantly influenced by the feature ranking. It shows that Gain Ratio (GR) technique gave better result than the other four techniques. The highest accuracy 98.93% for GR was reached using the 21 best features.

Five filter-based ranking methods have been evaluated. The first seven features predicted by the five techniques gives the same results. The first seven features are: fundamental frequency, spectral roll off 40%, spectral centroid, spectral roll off 90%, spectral flux, spectral kurtosis and spectral skewness. Those features give accuracy of 96.55% for gamelan instrument identification.

References

- [1] P Langley. *Selection of Relevant Features in Machine Learning*. Proc. of the AAAI Fall Symposium on Relevance. 1994.
- [2] M Dash, H Liu. *Feature Selection for Classification*. Intelligent Data Analysis. 1997; 1: 131–156.
- [3] I Guyon, A. Elisseeff. An introduction to Variable and Feature Selection. *J. Mach. Learn. Res.* 2003; 3:1157–1182.
- [4] Y Sun. Iterative RELIEF for Feature Weighting: Algorithms, Theories, and Applications. *Pattern Analysis and Machine Int., IEEE Trans. On.* 2007; 29:1035–1051.
- [5] Y Says, I N Inza, P Larranaga. A Review of Feature Selection Techniques in Bioinformatics. *Bioinformatics.* 2007; 23(19): 2507–2517.
- [6] O Lartillot, P Toivainen. *A MATLAB Toolbox for Musical Feature Extraction from Audio*. Proc. Int. Conference on Digital Audio Effects the 10th. Bordeaux, France: DAFX. September 2007; 1–8.
- [7] A Tjahyanto, Y K Suprpto, M H Purnomo, D P Wulandari. *FFT-based Features Selection for Javanese Music Note and Instrument Identification using Support Vector Machines*. Proc. IEEE Int. Conf. on Comp. Science and Automation Eng. CSAE. Zhangjiajie, China. May 2012; 439–443.
- [8] J C Brown, B Zhang. Musical Frequency Tracking using the Methods of Conventional and "Narrowed" Autocorrelation. *Journal Acoustical Society of America.* 1991; 89(5): 2346–2354.
- [9] Y K Suprpto, M Hariadi, M H Purnomo. Traditional Music Sound Extraction Based on Spectral Density Model using Adaptive Cross-correlation for Automatic Transcription. *IAENG International Journal of Computer Science.* 2011; 38(2):1–8.
- [10] Y K Suprpto, D P Wulandari, A Tjahyanto. Saron Music Transcription using LPF-cross Correlation. *Journal of Theoretical and Applied Information Technology.* 2011; 32(1): 71–79.
- [11] K D Martin, Y E Kim. *Musical Instrument Identification: A Pattern-recognition Approach*. 136th Meeting of the Acoustical Society of America. ASA. Oct 1998; 1–11.
- [12] S Essid, G Richard, B David. *Instrument Recognition in Polyphonic Music Based on Automatic Taxonomies*. IEEE Trans. on Audio, Speech, and Lang. Proc. 2006; 14(1): 68–80.
- [13] V Vapnik. *Statistical Learning Theory*. New York: Wiley, 1998.
- [14] C J C Burges. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery.* 1998; 2(2): 121–167.
- [15] T Joachims. *Text Categorization with Support Vector Machines: Learning with Many Relevant Features*. Proceedings of ECML-98, 10th European Conference on Machine Learning, C. Nedellec and C. Rouveirol, Eds., Chemnitz, DE: Springer Verlag, Heidelberg, DE. 1998; 137–142.
- [16] I Kotsia, I Pitas. *Real Time Facial Expression Recognition from Image Sequences using Support Vector Machines*. Proc. of the IEEE Int. Conference on Image Processing. 2005; 2: II–966–9.
- [17] I Guyon, J Weston, S Barnhill, V Vapnik. Gene Selection for Cancer Classification using Support Vector Machines. *Mach. Learn.* 2002; 46(1): 389–422.
- [18] L Wang, J Zhu, H Zou. Hybrid Huberized Support Vector Machines for Microarray Classification and Gene Selection. *Bioinformatics.* 2008; 24(3): 412–419.
- [19] Sumarsam. *Cultural Interaction and Musical Development in Central Java*. Chicago: The University of Chicago Press. 1992, 1995.
- [20] W A Purwadi. *Seni Karawitan Jawa Ungkapan Keindahan dalam Musik Gamelan*. Yogyakarta: Hanan Pustaka. 2006.
- [21] A Klapuri. *Signal Processing Methods for the Automatic Transcription of Music*. Ph.D. Dissertation. Tampere University of Technology. March 2004.
- [22] A Tjahyanto, Y K Suprpto, D P Wulandari. *Model Analysis-By-Synthesis Aplikasi Pembangkit Suara Gamelan Sintetik*. Proc. Sem.Nas. Aplikasi Teknologi Informasi. SNATI. Yogyakarta: JTI Ull, June 2011.
- [23] A Krishnaswamy. *Application of Pitch Tracking to South Indian Classical Music*. Proc. International Conference on Multimedia and Expo. 2003; 3: 557–560.
- [24] G Hu, D Wang. *Auditory Segmentation based on Onset and Offset Analysis*. Trans. Audio, Speech and Lang. 2007; 25(2): 396–405.
- [25] J P Bello, C Duxbury, M Davies, M Sandler. On the Use of Phase and Energy for Musical Onset Detection in the Complex Domain. *Signal Processing Letters, IEEE.* 2004; 11(6): 553–556.
- [26] A Tjahyanto, Y K Suprpto, M H Purnomo, D P Wulandari. *Features Selection for Gamelan Instrument Identification Using RELIEF-F and Support Vector Machines*. Proc. The 13th Seminar on Int. Tech. and Its App., Electrical Eng., ITS SITIA, May 2012.
- [27] J P Bello. *Towards the Automated Analysis of Simple Polyphonic Music: A Knowledge-based Approach*. Ph.D. Dissertation. Queen Mary, University of London. January 2003.
- [28] A Klapuri. *Sound Onset Detection by Applying Psychoacoustic Knowledge*. Proceedings of the Acoustics, Speech, and Signal Processing on 1999 IEEE International Conference - ICASSP. Washington, DC, USA: IEEE Computer Society. 1999; 06: 3089–3092.
- [29] S Z Li. Content-Based Audio Classification and Retrieval using the Nearest Feature Line Method. *IEEE Transaction on Speech and Audio Processing.* 2000; 8(5): 619–625.
- [30] L T DeCarlo. On the Meaning and Use of Kurtosis. *Psychological Methods.* 1997; 2(3): 292–307.

- [31] E Wold, T Blum, D Keislar, J Wheaton. Content-based Classification, Search, and Retrieval of Audio. *IEEE Multimedia*. 1996; 3(3): 27–36.
- [32] W Chu, B Champagne. A Noise-Robust FFT-Based Auditory Spectrum With Application in Audio Classification. *IEEE Transactions on Audio, Speech, and Lang. Proc.* 2008; 16(1): 137–150.
- [33] E Scheirer, M Slaney. *Construction and Evaluation of a Robust Multifeature Speech/Music Discriminator*. Proc. IEEE Int. Conf. Acoustic, Speech, Signal, Process. 1997; 2: 1331–1334.
- [34] P Bermejo, L dela Ossa, J A Gamez, J M Puerta. Fast Wrapper Feature Subset Selection in High-dimensional Datasets by means of Filter re-ranking. *Knowledge-Based Systems*. 2012; 25: 35–44.
- [35] I H Witten, E Frank. *Data Mining - Practical Machine Learning Tools and Techniques*, 2nd ed. San Francisco, CA: Morgan Kaufman. 2005.
- [36] K Kira, L A Rendell. *A Practical Approach to Feature Selection*. Proceedings of the 9th. Int. Conf. on Machine Learning. Morgan Kaufmann. 1992; 249–256.
- [37] M Robnik-Sikonja, I Kononenko. Theoretical and Empirical Analysis of Relief-F and R-Relief-F. *Machine Learning Journal*. 2003; 53: 23–69.
- [38] R L Plackett, K Pearson. The Chi-Squared, *Test International Statistical Rev.* 1983; 51(1): 59–72.
- [39] Y Yang, J O Pedersen. *A Comparative Study on Feature Selection in Text Categorization*. Proceedings of the Fourteenth International Conference on Machine Learning, ICML '97. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. 1997; 412–420.
- [40] M M Valdivia, M D Galiano, A M Ruez, L U Lopez. Using Information Gain to Improve Multi-modal Information Retrieval Systems. *Inf. Processing and Management*. 2008; 44: 1146–1158.
- [41] J Quinlan. *C4.5: Programs for Machine Learning*. San Francisco, CA: Morgan Kaufman, 1993.
- [42] J Chen, H Huang, F Tian, S Tian. A Selective Bayes Classifier for Classifying Incomplete Data based on Gain Ratio. *Knowledge-Based Systems*. 2008; 21: 530–534.
- [43] W Altidor, T M Khoshgoftaar, J V Hulse. *Robustness of Filter-based Feature Ranking: A Case Study*. Proc. of the Twenty-Fourth International Florida Artificial Intelligence Research Society Conf., 2011.
- [44] C C Chang, C J Lin. LIBSVM: A Library for Support Vector Machines, Tech. Rep., 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/>.
- [45] O Chapelle, V Vapnik, O Bousquet. Choosing Multiple Parameters for Support Vector Machines. *Machine Learning*. 2002; 46: 131–159.
- [46] Y K Suprpto, I K E Purnama, M Hariadi, M H Purnomo. *Sound Modeling of Javanese Traditional Music Instrument*. Proc. Int. Conf. on Instrumentation, Communications, Information Technology, and Biomedical Engineering, ICICI-BME. Bandung: ITB. 2009; 1–6.