Identical Instruments Detection using *Least Mean Square* Based Adaptive Filter

M Sa’adah, D P Wulandari and Y K Suprapto

Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

**Abstract.** Gamelan is one of Indonesian traditional music instrument. It have variations in terms of fundamental frequency, amplitude, and signal envelope, due to its handmade construction and playing style. When playing gamelan instruments, many identical music is played simultaneously. Detection of identical instruments are needed to indicate the changes of Gamelan features. We can find how many identical instruments are played at the same time or separated by very short time. We applied Least Mean Square (LMS) adaptive filter to detect identical instruments. The reference signal can have a fundamental frequency shift from the original signal to 10 Hz. The synthetic saron signal is used as reference signal. The performance can separate the identical instruments in 0.03 ms.

1. Introduction

Gamelan is one of the traditional musical instruments. Gamelan, which is played together like an orchestra. The difference between Gamelan and Orchestra is a conductor, where no one in the Gamelan plays acts a conductor to lead the game, while in the Orchestra there is a conductor.

Gamelan consists of 10 groups of instruments, namely saron, kenong, kempul, kendang, boning, gong, and so on. Gamelan has two standard notations namely Slendro and Pelog. A Gamelan expert named Sindusawarno, has conducted research on the similarity of the Pelog-Slendro scale [1].

One of Gamelan set is Balungan. It consist Demung, Saron, and Peking. The fastest tempo of gamelan ensemble is 300 ms. Separating two or more signals is a very important to find signal features. In addition, separating identical signals is very important to know the number of identical instruments played. This process can be used for music tagging. Kitahara identified the polyphonic music instrument like piano, flute and guitar, and he used a synthetic music instrument. Their sounds are separated by more than 10 ms and he did not play two or more identical instruments simultaneously within the same music notations. He identify each instrument played with different music notations but not an identical instruments [2]. Separating two or more identical music instruments is an extraordinary thing because the time delay among Sarons sound is very short, it is about 10 ms [3].

Gamelan is produced manually by a craftsman. The craftsmen is greatly rely on the hearing sensitivity of gamelan’s craftsmen themselves, while the hearing sensitivity of each gamelan craftsmen may different from one to another. So, the sound of gamelan from one craftsman to another craftsman can be different. It is because there is no standardization in manufacturing gamelan instrument [4]. The expression of each instrument was determined by its own particular characteristics. In other words, each component created its own specific melody, which could not be replaced or duplicated by others [5]. Fourier transform represents a signal in frequency domain, instead of in time domain. Wavelet does representing a signal in frequency domain too, but it still
represents the signal both in time domain and frequency domain. In this research we develop a method to identify two or more identical signals separated by very short time (msecond).

Gamelan is manually constructed and constructors tune the instrument with their own sense based on experience. As a result, fluctuation of frequency in each gamelan instrument is not set correctly. Because of that, we need to synthesize a gamelan signal for make a signal for reference signal. Ayers uses amplitude modulation to synthesize Gong Ageng and it can produces a tone that is very similar to the original. In this paper, we extract identical Saron signal using Least Mean Square (LMS), and use synthetic signal for the reference.

Reference signals are made automatically using several parameters, namely frequency, amplitude, phase and signal length. Furthermore, these parameters will be varied so that the reference signal will be similar to the input signal.

2. Fundamental Theory

2.1. Gamelan Signal

Gamelan is one of Indonesian traditional music instruments. It is made of iron, brass or bronze slabs. The gamelan spectrum varies greatly due to violence and hitting force although it still has the same basic frequency.

The frequency of saron tones is not always the same because the production of saron is different. Making saron depends on the gamelan master which affects the dissolution of saron, without any standards. This results in varied tones. In addition, the basic factors can also be different from the techniques used by the player and the beating power. Besides that, a gamelan can also change because of the gamelan instruments. Factors that are influenced by several factors, such as: very old and frequent gamelan instruments, occur in weathering in the original wood.

| Saron | Fundamental Frequency of Saron Slendro (Hz) | 1 | 2 | 3 | 5 | 6 | 1’ |
|-------|------------------------------------------|---|---|---|---|---|----|
| 1     | 521.5                                    | 616.8 | 715.4 | 814.7 | 939.8 | 1080.8 |
| 2     | 544.1                                    | 619.0 | 716.1 | 814.6 | 942.4 | 1082.3 |
| 3     | 542.4                                    | 620.7 | 712.8 | 819.8 | 934.6 | 1086.1 |
| 4     | 542.8                                    | 620.4 | 713.0 | 820.2 | 934.3 | 1086.3 |
| 5     | 543.0                                    | 620.4 | 712.9 | 820.1 | 934.2 | 1086.2 |
| 6     | 542.8                                    | 620.3 | 713.1 | 820.0 | 934.2 | 1086.2 |
| 7     | 537.0                                    | 617.9 | 710.6 | 816.7 | 929.7 | 1044.3 |
| 8     | 544.3                                    | 618.9 | 716.1 | 814.5 | 942.5 | 1082.1 |
| 9     | 539.6                                    | 617.9 | 711.2 | 816.0 | 934.0 | 1082.7 |
| 10    | 541.1                                    | 617.4 | 710.5 | 815.6 | 932.8 | 1083.7 |
| 11    | 514.9                                    | 589.3 | 681.9 | 776.0 | 896.7 | 1039.3 |
| 12    | 536.9                                    | 611.4 | 709.9 | 815.2 | 928.0 | 1073.2 |
| 13    | 544.2                                    | 618.9 | 716.2 | 814.5 | 942.6 | 1082.1 |
| 14    | 541.8                                    | 620.0 | 714.9 | 814.4 | 940.5 | 1083.7 |
| 15    | 541.5                                    | 616.8 | 715.4 | 814.8 | 939.8 | 1080.8 |
| Min   | 514.9                                    | 589.3 | 681.9 | 776.0 | 896.7 | 1039.3 |
| Max   | 544.3                                    | 620.7 | 716.2 | 820.2 | 942.6 | 1086.3 |
| Average | 539.3                                    | 614.8 | 711.2 | 814.8 | 932.4 | 1077.0 |

In the Table 1, it can be seen that the frequency of saron tones is different. To find out about the elements contained in the signal, it is necessary to do digital signal processing. For example, by using FFT or STFT, it can be seen elements of balungan signals such as attack, decay, sustain, release, pitch, harmonic frequency, music color, sampling frequency and quantization level. During attack area the amplitude of the signal grows very fast. During decay period, the amplitude and frequency density is uncertain. The signal can not control. The signal swings irregularly and the amplitude become smaller.
within the time. After decay period, the signal goes to sustain area. During that period, the amplitude goes smaller exponentially and the frequency becomes stable. Based on its character, the beginning of signal can detect. If amplitude grows rapidly in very short time, we detect it as the beginning of signal or if we find an uncontrolled behavior, amplitude and frequency density, we detect it as decay area.

2.2. Fast Fourier Transform
Fourier transform is a process that is widely used to convert the domain of a function into the frequency domain. In digital signal processing, such as to analyze the signal of a gamelan, the basic frequency for each tone must be known. Equation (1) is the fourier transformation.

$$FT(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t} dt$$ (1)

Where: FT (ω) is a function in the frequency domain, ω is the radial frequency, x (t) is the input signal from fourier transformation with t is the unit of time (seconds).

The result of Fourier transform is the frequency spectrum. The frequency spectrum states the composition of the frequency arrangement of a signal. This aims to form a frequency table that can be used as a reference in the analysis of the next stage. One method used to determine the frequency of a signal is to use Fourier transform.

2.3. Reference Signal
Building a reference signal is an important part of the filter process using Least Mean Square. A reference signal is a signal that will be filtered to be similar to the original signal. The difference between the input signal and the output signal is called an error. The smaller error, will indicate that the input signal is close to the original signal, and vice versa.

2.4. Adaptive Filter
The Adaptive Filter is a computing device that model the relationship between two signals by way of real time. Adaptive filters are usually associated with the broader topic of statistical signal processing. The operation of signal filtering by definition implies extracting something desired from a signal containing both desired and undesired components. One of them, can be used as noise cancellation in the Gamelan [8].

3. Design System

![Figure 1. Design system of adaptive filter](image)

3.1. Blind Signal
In the block diagram in Fig. 1, there is a blind signal which is the signal to be analyzed. In this study, the blind signal is an input signal in the form of a saron music signal instrument. This blind signal can be in the form of one or more saron signals, meaning that this blind signal is an identic signal that will be detected.
Input signal is also called blind signal. Blind signal is the signal to be analyzed. Blind signal consists of several identical instruments. The first step is taking data from gamelan instruments. The gamelan instruments used in this study were saron. Because the tunes of the gamelan set are created with different basic frequencies so that later the gamelan will have a varied base tone, then the next step is the Fast Fourier Transform (FFT) analysis which is used to find out the basic frequency of the saron used as the research data, which then this basic frequency can be used as a reference to the reference signal which is a synthetic signal.

3.2. Reference Signal

The reference signal is generated following the rules regarding the gamelan tuner. The reference signal is formed by generating a sinusoidal signal with the desired frequency. Next, modulation of the standard envelope will be carried out [4].

\[
X_p(n) = A \sin(2 \pi \phi \left( f \frac{f_p}{f_s} x + \theta \right) \times 0.7693e^{-0.03^2 n^2})
\]  

Where \(X_p(n)\) is standart envelope. \(A\) is a amplitude of signal, \(f\) is fundamental frequency of reference signal, \(\theta\) is phase. Equation 2 only applies to saron. The difference in generating reference signals for saron with other gamelan devices is a multiplier. This multiplier factor is the value of the Gamelan signal reduction.

3.3. Least Mean Square Adaptive Filter

Adaptive digital filters are currently widely used for signal processing applications. The adaptive algorithm uses the value of the criterion of performace, the measurements of the input and desired signals so as to modify the parameters of the filter to improve the performance. This digital filter is widely used to obtain the desired signal spectral characteristics, eliminating unwanted signals (such as noise or interference signals). The adaptive digital filter structure is shown in Fig. 2.

![Figure 2. Principle of adaptive filter](image)

The block diagram in Fig. 2 indicates that, if the value of \(N(n)\) is known, then after subtracting this from the mixed signal \(d(n)\), the original signal \(X(n)\) is obtained. \(S(n)\) is synthetic signal as reference.

4. Experiments

This experiment uses the gamelan signal from the Balungan group, saron. Saron has 5 tones. The input signal is in the form of WAVE files (*.wav). Signal analysis needs to be done to find out the elements in a signal. The gamelan signal has several elements, namely attack, decay, sustain, release, pitch, harmonic frequency, sound color, sampling frequency and quantization level. Signal analysis can also
be done with Fourier analysis to find out the signal frequencies that make up the signal wave. There are two Fourier transforms, used, namely Fast Fourier Transform (FFT) to convert signals from the time domain to the frequency domain, and the Short Time Fourier Transform (STFT) which is a combination of windowing and FFT.

![Figure 3. Saron signal in time domain](image)

Fig. 3 shows the original Saron signal in time domain. The characteristics of the saron signal, such as attack, decay, sustain and release. Attack is when the instrument first sounds so that the beginning of the signal appears, in this phase the wave moves up from zero to peak. Decay is a continuation of the attack where the wave will drop at the medium level. While sustain is when the waves will survive at the medium level. And the last is release where the volume drops again and then disappears.

![Figure 4. FFT of saron signal](image)

Fig. 4 shows the original and synthetic Saron signal in frequency domain. From Fig. 4 we can know the fundamental frequency. The fundamental frequency includes harmonics and non-harmonics frequency of Saron.

The synthesis signal is used as a reference signal. Furthermore, the twin signal extraction process will be carried out using Least Mean Square (LMS). Every music notation signal is separated from the ensemble and produces an output signal so that it represents a music notation. When music notation is played it can be detected using amplitude detection.

The saron signal used is in the form of a file.wav record which has a fs of 48000. In addition to the blind signal there is also a reference signal. The reference signal used is Saron's synthetic signal. The LMS adaptive filter has several parameters that can affect its performance. These parameters are step size ($\mu$), signal length and iteration. Different parameters will produce different error values. In this research, these parameters become control variables. The step size used is 0.0003. The step size value is obtained after varying the step size value from 0.00005 to 0.0004. For each parameter changing, the
minimum MSE will be calculated. Mean Squared Error (MSE) is used for evaluating forecasting methods. Each error or remainder is squared. Then added up and added to the number of observations. The square of the difference between the results and values is actually called the quadratic loss function. MSE will produce a value that shows the average error and can easily distinguish between an error with another error. Step size 0.0003 has the lowest MSE value. The graph of step size effects on MSE can be seen in Fig. 5.

![Figure 5. Step size vs MSE](image)

The next experiment is done to find out until the density of how many identical signals can be detected. This test is carried out with two blind signal saron1 using a reference signal synthetic saron1. After testing by varying the value of time delay starting from 500 to 8500, then the MSE value is obtained as shown in Fig. 6. Fig. 6 show that the time delay affects the error value. The minimum MSE with time delay = 0.03 ms with MSE value is 0.015768. The maximum MSE is on time delay = 0.01 with MSE value is 0.040251. It concludes that the best detection for identical instrument when the distance is 1500. The performance gained with 0.03 ms accuracy.

![Figure 6. Graph of Distance Tolerance](image)
From the graph in Fig. 6, it can be seen that the value of MSE fluctuates but tends to fall. So that the greater the time delay, the MSE value will be lower, this happens because the signal density is very low, thereby reducing the mixing between the two signals and causing the detection process to be more accurate. Fig. 7 shows the coefficients of the adaptive filter with 60 times iteration. Then the identical instruments can be detect, it can be seen on Fig. 8.

![Figure 7. The Adaptive Filter Coefficients Signal with Time Delay = 0.03 ms](image1)

![Figure 8. Identical Instruments with Time Delay = 0.03 ms](image2)

This experiment is conducted to determine the effect of shifting the frequency of synthetic signals as a reference signal. This test is carried out by shifting the frequency of synthetic signals into frequencies higher or lower than the original frequency. From the Fig. 9 can be seen that the frequency of synthetic signal as reference signal is affects the error value. The minimum MSE is on frequency = 581 Hz with MSE value is 0.023539. The maximum MSE is on frequency = 600 Hz with MSE value is 0.026062. From the signal coefficient, after frequency=591 Hz we cannot detect the peak of identical instruments. So, from that experiments we know that the tolerance frequency is until 10 Hz.

![Figure 9. Graph of Frequency Tolerance](image3)
5. Conclusion
By implementing LMS adaptive filter, the identical instrument can be detected and separated in 0.03 ms. The reference signal can have a fundamental frequency shift from the original signal to 10 Hz. Moreover, this method can also detect how many instrument played.

6. References
[1] Sumarsam, _Cultural Interaction and Musical Development in Central Java_. Chicago: The University of Chicago Press, 1995.
[2] T. Kitahara, M. Goto, K. Komatani, T. Ogata, and H. G. Okuno, “Instrument Identification in Polyphonic Music: Feature Weighting to Minimize Influence of Sound Overlaps,” vol. 2007, 2007.
[3] Y. K. Suprapto, M. H. Purnomo, and M. Hariadi, “Segmentation of Identical and Simultaneously Played Traditional Music Instruments using Adaptive,” vol. 20, no. 3, pp. 88–92, 2009.
[4] A. I. Technology, “HIGH PERFORMANCE GAMELAN ANALYZER USING,” vol. 71, no. 1, 2015.
[5] N. Ishida, “The textures of Central Javanese gamelan music: Pre-notation and its discontents,” _Bijdr. tot Taal-, Land- en Volkenkd._, vol. 164, no. 4, pp. 475–499, 2008.
[6] Y. K. Suprapto, S. Member, and M. H. Purnomo, “Saron Sound Separation from Gamelan Ensemble using Cross Correlation for Music Notation,” no. xx, pp. 1–8.
[7] L. Ayers and C. W. Bay, “The Hong Kong University of Science & Technology Department of Computer Science.”
[8] D. P. Wulandari and Y. K. Suprapto, “Noise Cancellation in Gamelan Signal by Using Least Mean Square Based Adaptive Filter,” pp. 1–5.