

# Audio Watermarking Based on PCM Technique

Ranjeeta Yadav  
Department of ECE  
SGIT,  
Ghaziabad, INDIA

Sachin Yadav  
Department of CSE  
SGIT,  
Ghaziabad, INDIA

Jyotsna Singh  
Department of ECE  
NSIT,  
New Delhi, INDIA

## ABSTRACT

Digital instruments come to be used in Desk Top Music (DTM), live performance, etc. These performances are recorded as digital contents, and circulated actively through network and electronic media. Digital Water mark technology paid much attention to solve the problem of illegal distribution & duplication when the digital contents are spread over electronic media. In this paper, we focus on a sound synthesized process in digital instruments, and propose a audio watermarking based on wavetable switching method. Certain watermarks are embedded in wavetables that are included in our digital instruments, and the insertion of secret messages is actualized with wavetable switching. Additionally, embedded watermarks can be extracted from the acoustic signal. The proposed method is able to achieve a real-time watermark, i.e., both musical performance and the insertion of watermark can be actualized.

**Keywords:** Audio Watermark, Digital Instrument, Wavetable switching, MIDI

## 1. INTRODUCTION

Audio watermarking has been proposed for the protection of multimedia contents, and it has been used for recording media such as MPEG 1 Audio Layer III (MP3) and Microsoft Windows Media Audio. Most of these watermarks are achieved with a non-real-time system, and there are many methods of different approaches in this type of system, e.g., LSB (Least Significant Bit) substitution methods are most fundamental techniques for information hiding[5, 7], and amplitude modulation or phase shift methods in frequency domain are powerful tools for acoustic watermarking[2, 4]. This type of system uses previously recorded acoustic waveforms as cover data, and therefore, it might not be suitable for embedding watermark in real-time, and it would make difficult to use it for situations like live-performance, where the illegal recording of acoustic sound has easily been made. It is a serious problem, and therefore, real-time watermarking is required. Although there are several methods to embed secret messages in real-time, they have another problem. (6)We focused on a situation of live performance, and watermarking has been achieved by amplitude modulation in frequency domain, with several composition methods in real-time. A real-time watermarking device that uses MCLT (Modulated Complex Lapped Transform) for embedding, with hardware implementation (8). This approach tries to reduce the time-delay with embedding procedure. That is, in the conventional approach of real-time embedding, unnecessary delay occurs when sound outputs with embedding procedure, and the delay could not be avoided. This is the problem that remains to be solved.

In this study, we focused on digital instrument to solve this problem. Digital instrument (like electronic piano, electric drum and so on) has been used for musical performances in practice, and digital instrument has a characteristic structure in its sound synthesis process. Additionally it was considered that characteristics of the synthesis could be used for real-time watermarking. The purpose of this study is to achieve a real-time watermarking technology with another approach for musical performance with digital instruments. In the proposed method, watermarks are embedded in the wavetable of digital instruments, and embedded data are extracted from the playback acoustic signal of digital instrument. Therefore, watermarks can be embedded in the output acoustic signal that is synthesized from the wavetable in real-time musical performance, and it can be useful for the copyrights protection of real-time generated acoustic media.

## 2. PROPOSED SCHEME

### 2.1 Instrument Watermarking

PCM sound-sources are used in synthesizing acoustic signals with a wavetable. Although the PCM waveform of a wavetable has been modified for sound output during the sound generation, the synthesized PCM acoustic sound maintains the frequency spectra of an original wavetable. We focused on the characteristics of sound generation, and an attempt was made to develop a watermarking method using digital instruments. The basic scheme of our technical proposal shows in Figure 1

**Step 1:** Watermark has been embedded in PCM sources of wavetable.

**Step 2:** Playing music with the instrument that includes a watermarked wavetable and the output acoustic signal is recorded.

**Step 3:** Watermark includes all the output acoustic signals.

In this way, watermarks are automatically inserted into music in real-time.

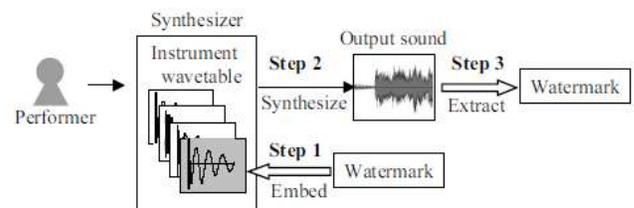


Figure 1: Outline of the proposed method

## 2.2 Wavetable Switching

Certain marker signals are embedded in a waveform on a wavetable, and embedded signals can be extracted from the output acoustic signal by the proposed method. Additionally, a number of marker signals are embedded in a certain wavetable, and number of alternative wavetables are generated, so that these marker signals can be observed from the acoustic waveform.

Thereupon, we construct the instruments that waveforms can be switched at each output, if its selection is controlled by a special key, and then the special components can reveal any information as a watermark. Figure 2 shows the systematic flow of the proposed scheme.

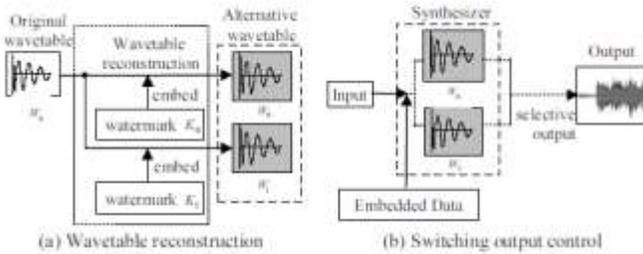


Figure 2: Systematic flow of the proposed scheme

### 2.2.1 Wavetable Reconstruction

In the proposed method, alternative wavetables have been used to reveal information, and in this process, a number of wavetables have been generated from the source wavetable by embedding marker signal. Two alternative wavetables  $W_0$  and  $W_1$  are prepared from one source wavetable  $W_S$  for watermarking as 1 bit expression.

### 2.2.2 Output control based on embedded data

Secret messages would be embedded as described in the following. First, a control code as each input is received at the interfaces of instruments, a bit  $E$  is taken from a watermark message at the same time. Second, the output waveform of the Wavetable is selected to conform

$$W_E = \begin{cases} W_0 & (E = 0) \\ W_1 & (E = 1). \end{cases} \quad (1)$$

Finally, acoustic signals are synthesized using  $W_E$ . Figure 3 shows an example of the insertion processes of a watermark with this wavetable switching at “ $E = 0$ ”.

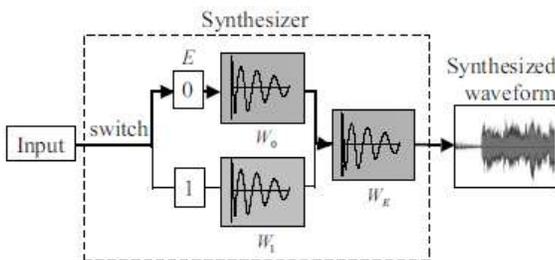


Figure 3: Switching wavetable ( $E = 0$ )

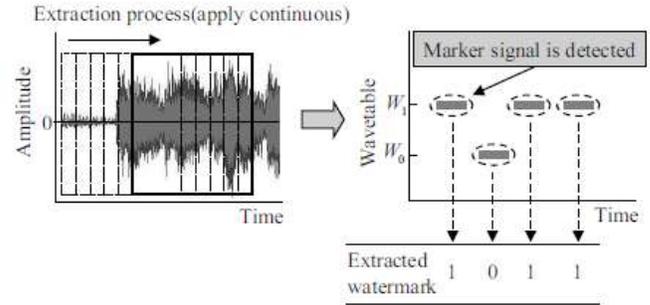


Figure 4: Processing flow of extraction

### 2.2.3 Watermark extraction

In the previous process, the embedded data are expressed as wavetable alternation, and we have to distinguish each watermarked wavetable with time-shift and continuous analysis to extract embedded data. Figure 4 shows a continuous extraction procedure for identification of the embedded watermarks. If any marker signal is detected, the wavetable  $W_d$  which is used in ‘wavetable switching’ can be determined, and extracted data  $E$  confirm

$$E = \begin{cases} 0 & (W_d = W_0) \\ 1 & (W_d = W_1). \end{cases} \quad (2)$$

As remarked above, the proposed method uses alternative wavetables to reveal embedded data, and information embedding would be processed in each tone. Therefore, embedding payload of the proposed method depends on the number of alternative wavetables, i.e., if the number of alternative wavetables is  $2n$ ,  $n$  bit data can be embedded in each tone.

In the proposed scheme, a marker signal is embedded in a waveform beforehand, and watermarks can be embedded only by switching of alternative wavetables in real-time. Therefore, the time-delay is negligibly small as compared with that of the conventional method.

## 2.3 Consideration about marker signal

In the proposed scheme, the embedded marker signal must be detected from an output waveform. However, there exist some Problems to detect the marker signal. It is because the inner waveform modification is held during sound synthesis procedures and output waveforms can be attacked for removal of watermarks from outside.

### 2.3.1 Inner waveform modifications.

There exist amplitude modification, pitch-shift, and looping during the sound synthesis process. Looping and pitch-shift are used to reduce the size of storage as described in section 3, and therefore, a waveform is transformed into an output acoustic waveform in these processes.

### 2.3.2 Watermark removal attacks.

Generally, watermarked waveforms are modulated to remove the embedded watermarks, and there are many types of attacks, e.g., amplitude modulation, adding noise, linear data compression, frequency modulation, application of band-pass filter, and so on.

The marker signal ought to be robust against these modifications of attacks, and additionally, the marker signal must be imperceptible to human auditory systems. Therefore, it is desirable to use a certain watermarking technique as a marker signal.

## 2.4 An implementation and evaluation

The proposed method is implemented as software simulation, and is evaluated from three perspectives, correct extraction, robustness and sound quality.

### 2.4.1 Experimental system

This study was performed with a PCM synthesizer implemented to a software called “TiMidity++”.

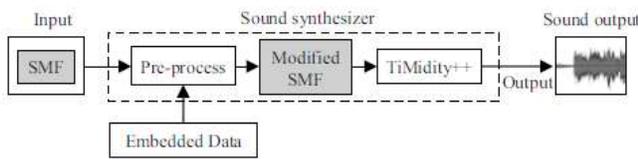


Figure 5: Structure of experimental system

Figure 5 shows the structure of the practical system, and the each part is shown as below.

#### 2.1.1.1 Input.

Standard MIDI File (SMF) [3] is used as sound control codes in this simulation.

#### 2.1.1.2 Sound synthesizer.

Sound is generated through next two procedures.

Step 1: SMF is modified with a watermark as a pre-process (addition of program change, and the channel number of note messages is rewritten).

Step 2: Output acoustic signal is generated by TiMidity++ that is controlled by the modified SMF.

#### 2.1.1.3 Sound output.

Acoustic waveform is stored as output signals in TiMidity++. Digital sound recording is implemented as this process.

Modifying SMF at the Step 1 of the process (2) is equivalent to the waveform modification in this system. The characteristic of TiMidity++ is used in the process of embedding watermarks effectually.

The characteristic of TiMidity++ is that instruments can be registered arbitrary in configuration. Instruments can be changed by the channel number of note message. It means that information is embedded at the pre-process stage by modifying SMF, i.e., switching channel of instrument is equivalent to wavetable switching.

### 2.4.2 Extraction test

Some correct extraction tests were held whether the proposed method worked correctly or not. The purpose of this test is to clarify the basic effectiveness of the wavetable switching method.

### 2.4.2.1 Synthesis conditions

#### i. Experimental instruments and embedding conditions.

The proposed method was evaluated with four instruments (wavetables), that is, acoustic-piano, jazz-guitar, flute and trumpet (see Table 1). Additionally, watermarks with  $K_0 = 188$  and  $K_1 = 27$  were embedded in waveforms of each instrument respectively, and  $W_0$  and  $W_1$  were generated. These waveforms were registered in TiMidity++ as different instruments.

#### ii. Experimental phrase

A phrase was used for this test. It has five notes with no pitch-shift, and all notes were existent in the wavetable (see Figure 6).



Figure 6: Experimental phrase

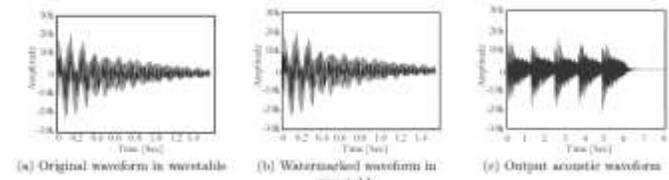


Figure 7: Results of waveform reconstruction and sound output (Piano)

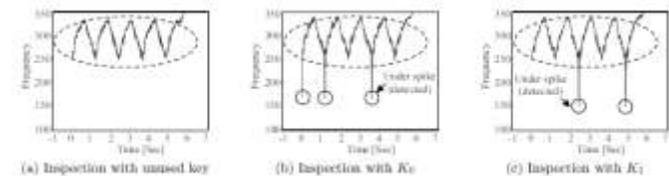


Figure 8: Inspection results of Piano

#### iii. Sound synthesis conditions

Sound synthesis with various volumes was held, that is, amplitudes of the output waveforms were 20, 40, 60, 80 and 100 percent of the original amplitude in wavetables. In this test, the bits of “00101” were embedded with wavetable switching.

### 2.4.2.2 Result and considerations.

Figure 7-(a) is an original waveform of the piano instrument, and Figure 7-(b) is a watermarked waveform in the instrument which is embedded with key  $K_0 = 188$ . Figure 7-(c) is the output acoustic waveform played with the embedded instrument (sampling ratio is 44.1 kHz, and signed 16 bit quantization). In the waveform, the amplitude was modified to 60 percent of the original waveform.

TABLE 1. Experimental instruments

Sample name	Patch file name	Included waveforms
Piano	000_Acoustic_Grand_Piano.pat	10
Guitar	026_Jazz_Guitar.pat	13
Flute	073_Flute.pat	4
Trumpet	056_Trumpet.pat	6

Additionally, Figure 8-(a), (b) and (c) show results of inspection after analyzing Figure 7-(c) output waveform at an unused key, watermarked key  $K_0$  and  $K_1$ . From comparison of the under peak with Figure 8-(a) and (b), remarkable spikes are observed at time  $t = 0; 1:2$  and  $3:6$ [sec] with key  $K_0$  (Figure 8-(b) ). And from comparison of the under peak with Figure 8-(a) and (c) in a similar way, spikes are observed at time  $t = 2:4$  and  $4:8$ [sec] with  $K_1$  (Figure 8-(c) ). These spikes show that the embedded data "00101" are extracted correctly with "Piano".

The results with all instruments are shown in Table 2. From the results, some considerations are made as below.

First, watermarks were extracted correctly in performance with all instruments, and the result shows that the proposed method works effectively as real-time watermarking.

Second, embedded bits were extracted on all amplitude in this test, and therefore, it can be said that the proposed method can work correctly in sound synthesis of amplitude modification. This method is robust with at least 20 percent amplification from experimental results because applying the psychoacoustic model has strengthened the embedding intensity.

Third, a marker signal was embedded in the heading of waveform in the proposed method, and the influences of looping can be avoided. It is because, in general, the loop section is located near the tail of the waveform.

Finally, in this implementation, embedding with wavetable switching was applied only in the note that compliant waveforms were present in wavetable, and watermarks were not embedded in other notes. It is because the SDM provides little robustness against pitch scale modification, that is, marker signal detection from the waveform that is generated with "pitch-shift" process is difficult. Future tasks are compliant to "pitch-shift". The robustness against pitch-scale modification might be necessary in the marker signal for this task.

### 2.4.3 Robustness

For the evaluation of robustness as watermarking technique, some attacks were held to a watermarked acoustic waveform by

TABLE 2. Extraction results of extraction test

Amplitude [%]	20	40	60	80	100
Piano	○	○	○	○	○
Guitar	○	○	○	○	○
Flute	○	○	○	○	○
Trumpet	○	○	○	○	○

the proposed method, and it was checked whether extraction of watermarks could be succeeded or not. Target waveforms were output waveforms of the performance with experimental phrase by each instrument, and the playback amplitude was 60 percent of the original waveforms.

Experimental attacks are shown in the followings.

#### 2.4.3.1 Linear data compression (MP3, AAC)

The ratios of 64kbps, 96kbps, 128kbps and 192kbps were used in MPEG1 Audio layer 3 (MP3). The ratios of 96kbps, 128kbps and 192kbps were used in MPEG2 AAC (AAC).

#### 2.4.3.2 Down sampling (Down).

The sampling ratio was down-sampled to 22.05 kHz.

#### 2.4.3.3 Reduction of quantization bit rate (Quant).

The resolution of samples was modulated to 8bit.

#### 2.4.3.4 Adding white-noise (Noise).

Continuous 53:8dB SPL (Sound Pressure Level) white-noise was added. A normalization of the reference level of 96 dB SPL is performed in such a way that the maximum value of sample corresponds to 96dB.

The experimental results are shown in Table 3. The feature common to all observations is the robustness against addition of white-noise, amplitude modification and down sampling. Robustness against addition of white noise and amplitude modification were good enough in all instrument. Embed components were controlled with a key in the proposed method. Because of this frequency hopping, there is robustness essentially against attacks of passing a band-pass filter.

But in contrary, the proposed method would not have robustness against down sampling, and this is because the frequency spectra over 11.025 kHz were erased with the down sampling procedure, hence SDM did not work correctly. The robustness can be considered to improve by the multiplexing of embedding with a limited bandwidth.

The robustness against linear data compression showed a different tendency with instruments. In piano and guitar, the robustness was good in comparison with others, and the robustness was good enough except the compression with low bit-ratio.

TABLE 3. Robustness of the proposed method

Sample	(a) Robustness against linear data compression				(b) Robustness against other attacks			
	MP3 [kbps]		AAC [kbps]		Sample	Down	Quant	Noise
Piano	×	○	○	○	×	○	○	○
Guitar	×	○	○	○	×	○	○	○
Flute	×	×	○	○	×	○	○	○
Trumpet	×	×	×	×	×	○	○	○

This is because the cut-off frequency of low-pass filter was lower in low bit-ratio compression. In contrast, there was little robustness against linear data compression in flute and trumpet.

Although it is caused by the difference in tuned sounds and sustained sounds, the detailed research is a future task.

#### 2.4.4 Sound quality.

In this study, the sound-source identification test was made by an ABX double blind test with 9 raters whose ages were twenties and thirties. The ABX test assumes that acoustic waveform without embedding (played by original instruments) is labeled as A, and performance with embedding of the proposed method (played by watermarked instruments) is labeled as B. First, these performances are shown to the raters. Then, A or B is shown as the performance X to each rater randomly, and the raters evaluate X as either A or B. The accuracy ratio of X has a biased-value from 50 % when A and B can be distinguished clearly by listening, and accuracy ratio is defined as 50 % when these performances cannot be distinguished. However, even when A and B cannot be identified clearly, the accuracy ratio has a biased-value because the number of trials is limited. In this study, significance of the accuracy ratio was investigated by the  $X^2$  test.

In the experiment, X was shown so that X could not be distinguished as A or B. One cycle of judging X from A and B assumed one trial, and 5 times of trials were carried out. In evaluation, X was selected at every trial randomly. A monitor headphone ATH-PRO5V (made from Audio Technica) was used for listening of raters. The evaluation results of all raters were summarized, and tested as a result of 45 trials in the  $X^2$  test.

The results are shown in Table 4. From the results, there is no significant difference between watermarked waveforms and original waveforms with 10 % of significant level. It can be said that the raters in this experiment were not able to distinguish A or B in the result. That is, information can be embedded by the proposed method without degrading quality of performance. It is because embedding procedure of the proposed method is based on the psychoacoustic analysis.

TABLE 4.  $\chi^2$  test result of the proposed method

sample	correct[%]	$\chi^2$	p
Piano	51.1	0.022	0.881
Guitar	62.2	2.689	0.101
Flute	46.7	0.200	0.655
Trumpet	60.0	1.800	0.180

### 3. CONCLUSION

In this paper, we proposed an information hiding method as a real-time watermarking technique in musical acoustic signals.

Our watermark is inserted in the wavetable of a sound synthesizer separately, and the watermarked waveform was generated automatically with a selective output of the instruments in real-time. Therefore, the proposed method can be useful technique for rights management technology in such situations as illegal recording on live performance.

### 4. REFERENCES

- [1] A. Duncan, and D. Rossum.: Fundamentals of pitch-shifting, Preprint 2714 (A-1). Presented at the 85th Convention of the Audio Engineering Society. New York: Audio Engineering Society, 1988.
- [2] W. Bender, D. Gruhl, and N. Morimoto, Techniques for data hiding, IBM Systems Journal, vol. 35, no. 3/4, pp. 131-136, 1996.
- [3] AMEI: MIDI 1.0 Specification, Ritto Music, 1998.
- [4] N. Yajima, and K. Oishi, Digital watermarks for audio signals based on amplitude and phase coding, IEICE technical report. Communication systems, 104(720):53{59, May 2005. (in Japanese.)
- [5] M. Iwakiri, and K. Matsui, Visualizing technique of digital watermarks to audio data, Transactions of IPSJ, 41(6):1840{1847, June 2000. (in Japanese.)
- [6] R. Tachibana, Audio Watermarking for Live Performance, in Proc. of Security and Watermarking of Multimedia Contents V, SPIE vol. 5020, pp. 32-43, Santa Clara, USA, January 2003.
- [7] D. Gruhl, A. Lu, and W. Bender, "Echo Hiding," in Information Hiding: First International Workshop, Proceedings, vol. 1174 of Lecture Note in Computer Science, Springer, pp7-21, 1996.
- [8] J. J. G. Hernandez, M. Nakano and P. M. Hector, Real-Time AudioWatermarking System Prototype, Multimedia, 2006. ISMapos;06. Eighth IEEE International Symposium on Volume , Issue, pp. 792 - 793, Dec. 2006.
- [9] M. Iwakiri, and K. Matsui, Watermarking Technique for Digital Sound Data Using the Reactivity Regulation Scheme, Transactions of IPSJ, 43(8) pp.2519-2528, 2000. (in Japanese.)
- [10] ISO/IEC: Information technology { coding of moving picture and associated audio for digital storage media at up to about 1.5 mbit/s - Part 3: Audio, Tech. Rep. 11172-3, ISO/IEC, 1993.
- [11] C. Roads, The Computer Music Tutorial, MIT Press, 1996.
- [12] J. Pan, H. Huang, and L. C. Jain, Information Hiding and Applications, Springer, Berlin-Heidelberg, Germany, 2009.
- [13] M. Collins, In\_nity: DSP sampling tools for Macintosh, Sound on Sound 9(1): 44-47, 1993.