音源テーブル切替えによる実時間音響透かしとその性能評価

山本 紘太郎　　岩切 宗利

防衛大学校情報工学科
239-8686 神奈川県横須賀市走水 1-10-20

{g46073,iwak}@nda.ac.jp

あらまし　本報告では，PCM 音源テーブルを有する電子楽器を利用した実時間電子透かし法について提案し，性能を評価した結果について報告する．音響信号への電子透かしは著作権保護に有用な技術である．しかし，ライブ演奏などでは，音響の生成と同時に違法な録音が生起しうる．この場合，従来の電子透かし方式では実時間で情報を埋め込むことが難しい．本報告では，PCM 音源を利用した実時間電子透かし法を提案する．提案方式では，事前に PCM 音源テーブルへ情報を埋め込んだものを複数用意し，演奏時にそれらを切り替えながら音を生成することにより，透かし入り音声を出力する．提案方式を用いることで，実時間での透かし埋込みが可能になる．

A Watermark Availability Evaluation of Sound Wavetable Switching Method

Kotaro Yamamoto　　Munetoshi Iwakiri

Department of Computer Science, National Defense Academy
1-10-20 Hashirimizu Yokosuka Kanagawa Japan 239-8686

{g46073,iwak}@nda.ac.jp

Abstract Musical performance with digital instruments has become a common practice today. These performances are recorded as digital contents. However, the spread of digital contents causes the problem of illegal duplication, so that digital watermark has recently attracted much attention as a technology to solve this problem. In this report, we focus on a sound synthesized process in digital instruments, and propose a real-time watermark 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.

1. Introduction

In recent years, various kinds of information have been digitized as multimedia contents. Since digital contents could be duplicated easily, a copyright infraction has occurred by the illegal spreading of digital contents. Digital watermarking has been attracted much attention to solve this problem, and many approaches have been made to construct copyright management systems[1].

Audio watermarking has been proposed for multimedia contents protection, and it has been used for recording media. Most of these watermarks are achieved with a non-real-time system, and there are many methods of different approaches[2]. These systems use once recorded acoustic waveform as a cover data, therefore, it might not been suitable for embedding watermark in real-time, and it would make difficult to use it for situations like live-performance, where the illegal audio recording has easily been made. Therefore, real-time watermarking is necessary.

There are several methods to embed secret messages in real-time[3, 4]. However, these methods include DFT (Discrete Fourier Transform) or MCLT (Modulated Complex Lapped Transform), and the time-delay occurs between the host signal and watermarked signal because of the large calculation.

The purpose of this study is to achieve a real-time watermarking technology with another approach for musical performance with digital instruments (like electronic piano, electric drum and so on). 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 should be useful for the copyrights protection of real-time generated acoustic media.

Chapter 2 describes the digital instrument and its structure. Chapter 3 contains a technical proposing explanation about the embedding algorithm of inserting watermark, and Chapter 4 shows an experimental result with software simulation to evaluate the practicality of the proposed method.

2. Digital instrument

Digital instruments are classified into three types with its sound generation methods[5], PCM[1], functional generator[2] and acoustic modeling[3].

The PCM uses an acoustic signal of real instruments, but the others use pseudo waveform based on an acoustic model. Therefore, the PCM synthesis is widely used.

2.1 PCM sound synthesis

Basically, the PCM sound synthesis is simple. First, acoustic sounds of natural instruments are recorded as PCM sequences, and they are stored into digital instrument. Second, performer's operation is input from interface of the digital instrument. Finally, the PCM sequence which is matching with the performer's operation is read, and output with a little modification.

However, in some situations, the amount of data may be too large to construct instruments. For example, to make an acoustic piano requires many resources, e.g., waveforms for each key and various velocity, and since acoustic waveform sizes are flexible, the instrument has to generate acoustic synchronizing for performer's intention correctly. Generally, PCM synthesis uses "Looping[6]" and "Pitch-shift[7]" to reduce the size of a wavetable at correct sound generating.

Looping is used to expand the length of limited size waveforms, and it is functional to generate waveforms for long-term[6]. First, loop-section is set at the waveforms (i.e., start and end loop-point set). Then, the waveforms can be generated with repeating the loop-section.

If a digital instrument has insufficient storage to store full tones, the stored tones are decreased in the instrument. In this case, a necessary musical pitch has to be generated from existent waveforms in the instrument[7]. There are two types of pitch shift, one is processing in frequency-domain, and another is processing in time-domain.

In this way, PCM sequences are modified within sound synthesis procedure. And additionally, some parameters are essential to modify the source PCM sequences correctly. Therefore, PCM waveforms are collected with some parameters, and are stored as a "wavetable" in the instrument.

2.2 Wavetable

There are a number of the parameters for PCM sound synthesis in a wavetable, e.g., start and end point of the loop-section, fundamental frequency (f0) of the PCM sequence, coverage of playback frequency, and so on. They are collected with each PCM sequences, and stored in wavetable. An instrument is conformed like this, and the wavetable is loaded on the synthesizer for sound generation.

[1]Collected natural instrument sounds are used for sound generation.

[2]UG (Unit Generator) with a factor of acoustic-signal-processing. The UG has many functional processor or software algorithms.

[3]Acoustic models based on the structure of natural instruments to generate sound.

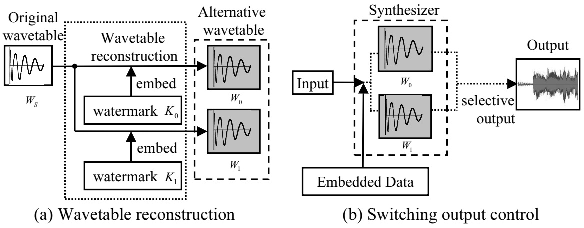


Figure 1: Systematic flow of the proposed scheme

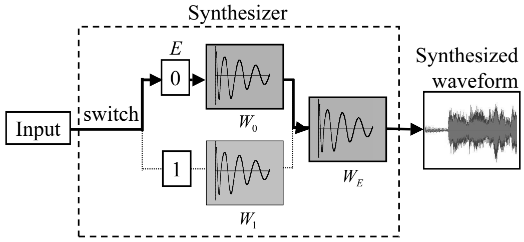


Figure 2: Switching wavetable ($E = 0$)

3. Proposed scheme

In section 2., it was described that 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.

3.1 Wavetable switching method

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 generate a number of alternative wavetables, so that these marker signals can be observed from the output waveform.

Thereupon, we construct the instruments that waveforms are able to be switched at each output, if its selection is controlled by special key, and then the special components can reveal any information as a watermark. **Figure 1** shows the systematic flow of the proposed scheme.

3.1.1 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.

3.1.2 Output control based on watermarks

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 2** shows an example of the insertion processes of a watermark with this wavetable switching at "$E = 0$".

3.1.3 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 3** 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 can confirm

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

3.1.4 Feature of the proposed method

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

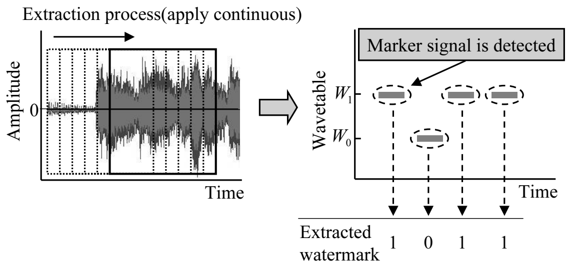


Figure 3: Processing flow of extraction

3.2 Robustness of the 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 are attacked for removal of watermarks from outside.

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.

We used "Spread Differential Method (SDM)" as the watermarking based on a frequency hopping method[8] in this study. Especially, so as to limit the influence to human auditory system, it was used with psychoacoustic analysis (MPEG 1 Audio psychoacoustic model 1 for layer 2[9]).

4. An implementation and evaluations

The proposed method was implemented as software simulation, and was evaluated from two perspectives, extraction and robustness.

4.1 Experimental system

This study was performed with a PCM synthesizer implemented to a software "TiMidity++[4]". We used a GUS patch (Gravis UltraSound patch) which is distributed by freepats[5] as a PCM soundsource. **Figure 4** shows the structure of the practical system, and the each part is shown as below.

[4]http://sourceforge.net/projects/timidity/
[5]http://freepats.opensrc.org/

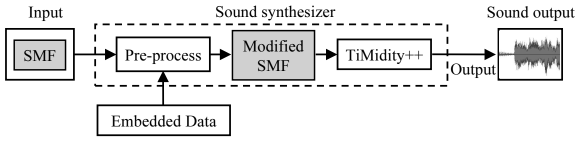


Figure 4: Structure of experimental system

Input.

Standard MIDI File (SMF)[10] is used as sound control codes in this simulation.

Sound synthesizer.

SMF is modified with a watermark as a preprocess (addition of program change, and the channel number of note messages is rewritten), and output acoustic signal is generated by TiMidity++ that is controlled by the modified SMF.

Sound output.

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

Modifying SMF at pre-process in "Sound synthesizer" is equivalent to the Wavetable Switching 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.

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.

4.2.1 Synthesis conditions

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.

Table 1: Experimental instruments

Sample name	Patch file name
Piano	000_Acoustic_Grand_Piano.pat
Guitar	026_Jazz_Guitar.pat
Flute	073_Flute.pat
Trumpet	056_Trumpet.pat

These waveforms were registered in TiMidity++ as different instruments.

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.

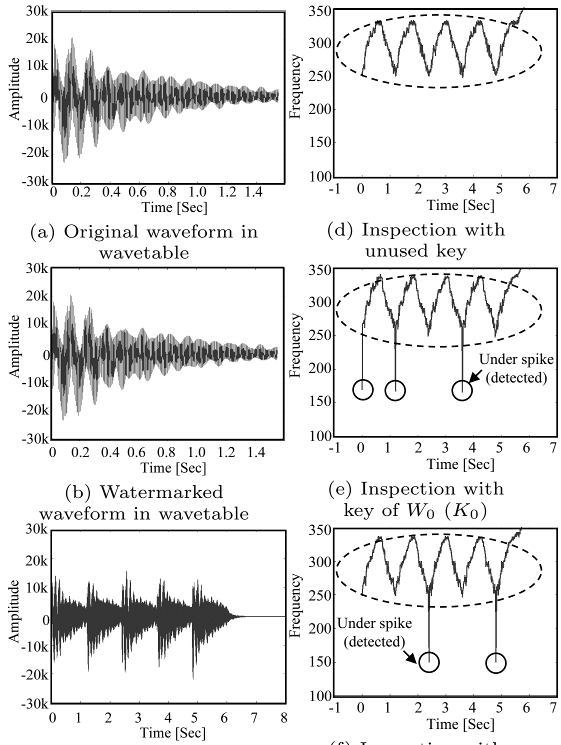
Sound synthesis conditions.

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.

4.2.2 Result and considerations

Figure 5-(a) is an original waveform of the piano instrument, and Figure 5-(b) is a watermarked waveform in the instrument which is embedded with key $K_0 = 188$. Figure 5-(c) is the output acoustic waveform played with the embedded instrument (sampling ratio is 44.1kHz, and signed 16bit quantization). In the waveform, the amplitude was modified to 60 percent of the original waveform. Additionally, Figure 5-(d), (e) and (f) show results of inspection after analyzing Figure 5-(c) output waveform at an unused key, watermarked key K_0 and K_1. Comparison of the under peak with Figure 5-(d), (e) and (f), the spikes observed with key K_0 at time $t = 0, 1.2$ and 3.6[sec], and those observed with K_1 at time $t = 2.4$ and 4.8[sec], and correctly extracted embedded data “00101” with “Piano”. Similarly, watermarks were extracted correctly in all amplitudes with each instrument. From the results, some considerations are made as below.

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



(a) Original waveform in wavetable
(b) Watermarked waveform in wavetable
(c) Output acoustic waveform
(d) Inspection with unused key
(e) Inspection with key of W_0 (K_0)
(f) Inspection with key of W_1 (K_1)

Figure 5: Results of waveform reconstruction and watermark extraction (Piano)

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 the embedding intensity has been strengthened by applying the psychoacoustic model. Finally, 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 the loop section is located near the tail of the waveform in general.

4.3 Robustness

For the evaluation of robustness as watermarking technique, some attacks which is shown in **Table 2** were held to a watermarked acoustic waveform by 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

Table 2: Evaluation items for robustness

(a) Linear Data Compression

Encoding method	Ratio [kbps]
MPEG1 Audio Layer 3 (MP3)	64, 96, 128, 192
MPEG2 AAC (AAC)	96, 128, 192

(b) Other Attack

Attack method	intensity
Down sampling (Down)	22.05 kHz
Reduction of quantization (Qua)	16 bit to 8 bit
Adding white noise (Noise)	53.8 dB SPL

Table 3: Robustness of the proposed method

(a) Robustness against linear data compression

Sample	MP3 [kbps]				AAC [kbps]		
	64	96	128	192	96	128	192
Piano	×	○	○	○	×	○	○
Guitar	×	○	○	○	×	○	○
Flute	×	×	○	○	×	×	×
Trumpet	×	×	×	×	×	×	×

(b) Robustness against other attacks

Sample	Down	Qua	Noise
Piano	×	○	○
Guitar	×	○	○
Flute	×	○	○
Trumpet	×	○	○

percent of the original waveforms.

The experimental results are shown in **Table 3**. The commonly-observed feature 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 against attacks of passing a band-pass filter essentially.

But in contrary, the proposed method would not have robustness against down sampling, and this is because the frequency spectra over 11.025kHz were erased with the down sampling procedure, and 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 the different tendency with instruments. In piano and guitar, the robustness was good in comparison with others, and the robustness was good enough except compression with low bit-ratio. 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 future task.

5. 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.

References

[1] S. Katzenbeisser, and F. A. P. Petitcolas, *Information Hiding Technique for steganography and digital watermarking*, Artech House, 2000.

[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] R. Tachibana, Audio Watermarking for Live Performance, in Proc. of Security and Watermarking of Multimedia Contents V, SPIE vol.5020, pp.32-43, 2003.

[4] J. J. G. Hernandez, M. Nakano and P. M. Hector, Real-Time Audio Watermarking System Prototype, Multimedia, 2006. ISMapos;06. 8th IEEE International Symposium on Volume , Issue, pp.792-793, 2006.

[5] C. Roads, *The Computer Music Tutorial*, MIT Press, 1996.

[6] M. Collins, Infinity: DSP sampling tools for Macintosh, *Sound on Sound* 9(1), pp.44-47, 1993.

[7] A. Duncan, and D. Rossum.: Fundamentals of pitch-shifting, Preprint 2714 (A-1). Presented at the 85th Convention of the Audio Engineering Society. Audio Engineering Society, 1988.

[8] 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.)

[9] 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.

[10] AMEI: *MIDI 1.0 Specification*, Ritto Music, 1998.