International Journal of Security and Its Applications
Vol. 10, No. 2 (2016), pp.105-114
http://dx.doi.org/10.14257/ijsia.2016.10.2.10

Application of Audio Watermarking Procedures for Mix Music

Youngseok Lee[1] and Jongweon Kim[2*]

[1]*Dept. of Electronics, Chungwoon University, 113, Sukgol-Ro, Nam-Gu, 22100 Incheon, Korea*
[2]*Dept. of Copyright Protection, Sangmyung University, 20, Hongimum 2-Gil, 03016 Seoul, Korea*
[1]*yslee@chungwoon.ac.kr*, [2] *jwkim@smu.ac.kr*

Abstract

In this paper we applies typical four audio watermarking technologies to protect ownership or copyright of mix music which is combined two or three pre-existing song. In experimental results, applied watermarking technologies cannot be applicable to mix music and three technologies fail to extract watermarks from mix music. This results is from mix music that does not contain full watermark information. Therefore applicable watermarking technology should development for mix music. The common problem of failed technologies cannot extract the second watermark from underlying mix music.

Keywords*: Audio watermarking, Mix music, Copyright, Multiple ownership*

1. Introduction

Recent years have seen a rapid growth in the availability of digital multimedia content. A major problem faced by content providers and owners is protection of their material. They are concerned about copyright protection and other forms of abuse of their digital content [1].

Under the influence of this trend in digital music field, digital technology and the Internet have not only made infinite collections of unique art available, but they have also made it possible for people to mix and mash others' works with little difficulty and no authorization. Consequently, society is witnessing a shift away from passive involvement in culture toward a more active, participation oriented scheme. The practice of borrowing ideas to create and inspire new art has never been as prevalent as it is now. One area that is increasingly affected by this shift is music. In fact, there is an entire genre of music, commonly known as "mashups," dedicated to borrowing Music Mashups: Testing the Limits of Copyright Law as Remix Culture Published by Scholarly Commons at Hofstra Law, 2010 and mix in others' works.' A music mashup8 is a song formed by combining two or more preexisting songs.9 in Copyright Law of USA [2].

Easy access and replication, however, have led to serious problems with copyright protection for media. Therefore, media owners can use this technique to insert some information into their media for the purpose of copyright protection or ownership. A digital watermark is a kind of marker covertly embedded in a noise-tolerant digital information such as audio or image data. It is typically used to identify ownership of the copyright of such digital contents. "Watermarking" is the process of hiding digital information in digital contents. The hidden information should but does not need to contain a relation to the content. The security and enforcement of academic property rights for digital media has become an important issue [3]. This paper contributes to reply to the applicability of audio watermarking

* Corresponding Author

ISSN: 1738-9976 IJSIA

technology for multiple watermarking system such as multiple ownership protection of mix music.

2. Properties of Audio Watermarking

There are about five properties that need to be satisfied for effective application of watermarking technology. These are robustness, imperceptibility, bit rate, security and computational complexity. Some of the watermark properties are discussed below.

2.1. Robust to Signal Processing

Digital signal may undergo common signal processing operations such as Linear filtering, sample re-quantization, D/A (digital-analogue) and A/D (analogue-digital) conversion and lossy compression.

2.2. Bit Rate

This is the amount of watermark data that may be reliably embedded within the host signal per unit time or space. A higher bit rate may be desirable in some application to embed copyright information. Reliability is measured using BER (bit error rate). In order to serve the purpose of providing identification of the addresser, it is necessary to transmit a certain amount of data. For example, this might be the aircraft's tail number or the 27 bit Data Link Service Address for aircraft and ground stations used in the context of Controller Pilot Data Link Communication (CPDLC). A GPS position report, which would be advantageous in various scenarios, requires approximately 50 bit of payload data. Altogether a data rate of 100 bit/s is desired, leaving some room for potential extensions.

2.3. Perceptual Quality

All suggested technologies affect the perceptual quality of the speech transmission. A too severe degradation of the sound quality would not be accepted by both the certification authorities and the intended users, and for example becoming annoying to the air traffic controllers. Ideally the participants of the communication should not notice the presence of the system. For this reason the perceptual distortion should be clearly minimized.

2.4. Watermark Security

Watermark security refers to the inability by unauthorized users to have access to the raw watermarking channel.

2.5. Computational Complexity

This refers to the processing required to embed data into the host signal and or to extract data from the signal. A discovery of methodology to satisfy these constraints will lead to a way of protecting digital audio which is the aim of this paper.

3. Watermarking Methods

In this section, the five most popular techniques for digital audio watermarking are reviewed. Specifically, the different techniques correspond to the methods for merging (or inserting) the cover data and the watermark pattern into a single signal. This section fully excerpts from [3] to help the understanding of audio watermarking to reader.

3.1. Spread Spectrum Watermarking

Spread-spectrum watermarking scheme is an example of the correlation method which embeds pseudorandom sequence and detects watermark by calculating correlation between pseudo-random noise sequence and watermarked audio signal. Spread spectrum techniques for watermarking borrow most of the theory from the communications community.

The main idea is to embed a narrow-band signal (the watermark) into a wide-band channel (the audio file). The characteristics of both audio signal *A* and watermark *W* seem to suit the model perfectly. In addition, spread spectrum techniques offer the possibility of protecting the watermark privacy by using a secret key to control the pseudorandom sequence generator. Spread spectrum techniques allow the frequency bands to be matched before embedding the message. This is why spread spectrum techniques are valuable not only for robust communication but for watermarking as well.

There are two basic approaches to spread spectrum techniques: *direct sequence* and *frequency hopping*. In both of these approaches the idea is to spread the watermark data across a large frequency band, namely the entire audible spectrum.

3.2. Amplitude Modulation

This method, also known as *least significant bit (LSB) substitution*, is both common and easy to apply in both steganography and watermarking (5) as it takes advantage of the quantization error that usually derives from the task of digitizing the audio signal. As the name states, the information is encoded into the least significant bits of the audio data. There are two basic ways of doing this: the lower order bits of the digital audio signal can be fully substituted with a pseudorandom (PN) sequence that contains the watermark message *m*, or the PN-sequence can be embedded into the lower order bit stream using the output of a function that generates the sequence based on both nth bit of watermark message and nth sample of audio file. The major disadvantage of this method is its poor immunity to manipulation. Encoded information can be destroyed by channel noise, re-sampling, *etc.*, unless it is encoded using redundancy techniques. In order to be robust, these techniques reduce the data rate, often by one to two orders of magnitude. Furthermore, in order to make the watermark more robust against localized filtering, a pseudorandom number generator can be used to spread the message over the cover in a random manner.

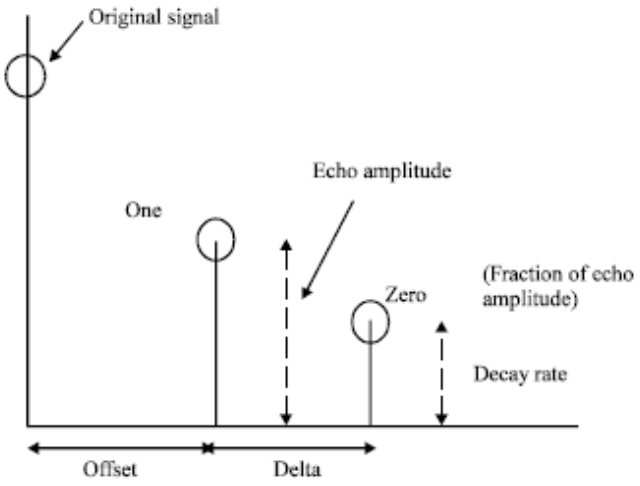


Figure 1. Echo Hiding Watermarking Scheme

3.3. Replica Modulation

Original signal can be used as an audio watermark. Echo hiding is a good example. Replica modulation also embeds part of the original signal in frequency domain as a watermark. Thus, replica modulation embeds replica, i.e., a properly modulated original signal, as a watermark. Detector can also generate the replica from the watermarked audio and calculate the correlation. The most significant advantage of this method is its high immunity to synchronization attack.

As one of replica modulation, echo hiding embeds data into an original audio ignal by introducing an echo in the time domain such that

$$x(n) = s(n) + \alpha s(n - d). \quad (1)$$

For simplicity, a single echo is added above (see Figure 1). Binary messages are embedded by echoing the original signal with one of two delays, either a d_0 sample delay or a $d_1$sample delay. Extraction of the embedded message involves the detection of delay d. Autocepstrum or cepstrum detects the delay d. Cepstrum analysis duplicates the cepstrum impulses every d samples. Echo hiding is usually imperceptible and sometimes makes the sound rich. Synchronization methods frequently adopt this method for coarse synchronization. Disadvantage of echo hiding is its high complexity due to cepstrum or autocepstrum computation during detection. On the other hand, anybody can detect echo without any prior knowledge. In other words, it provides the clue for the malicious attack. This is another disadvantage of echo hiding.

3.4. Dither Method

Dither is a noise signal that is added to the input audio signal to provide better sampling of that input when digitizing the signal. As a result, distortion is practically eliminated, at the cost of an increased noise floor.

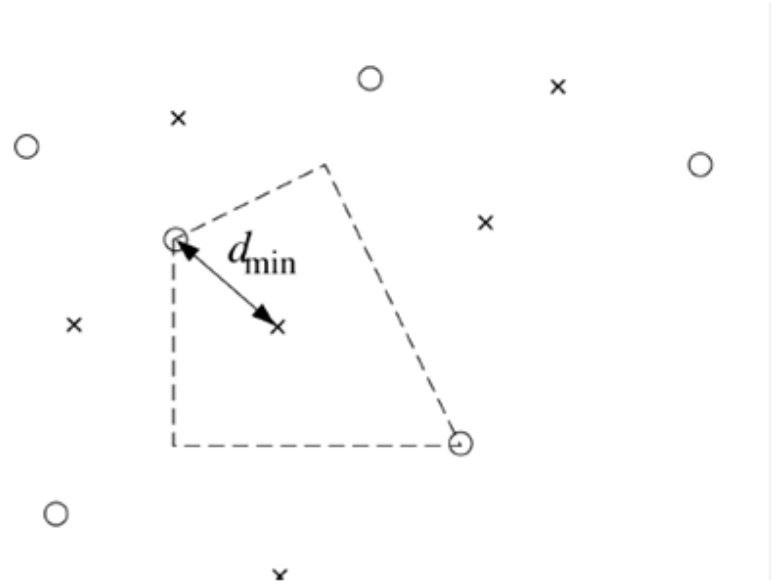


Figure 2. Quantization Index Modulation Lattice. The signal is Quantized to the Nearest 'o' or 'x' Depending on the Watermark bit. The Robustness is Determined by the Minimum Distance between any Two 'o' and 'x'

To implement dithering, a noise signal is added to the input audio signal with a known probability distribution, such as Gaussian or triangular. In the particular case of dithering for watermark embedding, the watermark is used to modulate the dither signal. The host signal (or original audio file) is quantized using an associated dither quantizer. This

technique is known as quantization index modulation (QIM).

Chen and Wornell [4, 5, 6] present a class of watermarking methods that inherent host interference rejection. The two core methods are quantization index modulation (QIM) and dither modulation (DM). They are based on lattice coding. The host signal amplitude or any other representation of the watermark. A graphical view of this technique is shown in Figure 2. Here, the points marked with X's and O's belong to two different quantizer, each with an associated index; that is, each one embedding a different value. The distance d_{min} can be used as an informal measure of robustness, while the size of the quantization cells (one is shown in the figure) measures the distortion on the audio file. If the watermark message m=1, then the audio signal is quantized to the nearest X. If m=2 then it is quantized to the nearest O.

3.4. Amplitude Escaling Method

The decoder tries to estimate and subsequently correct the amplitude scaling that occurred on the channel. Looking at the histogram of the received data, the quantization based watermark embedding process leaves significant marks there, which allow the estimation of the channel's amplitude scaling. Shterev [6] presents a probability density model of the received data. The model shows that the probability density function (PDF) of the received data has characteristic peaks and discontinuities, which allow the extraction of the amplitude scaling by using Fourier nalysis or Maximum Likelihood estimators.

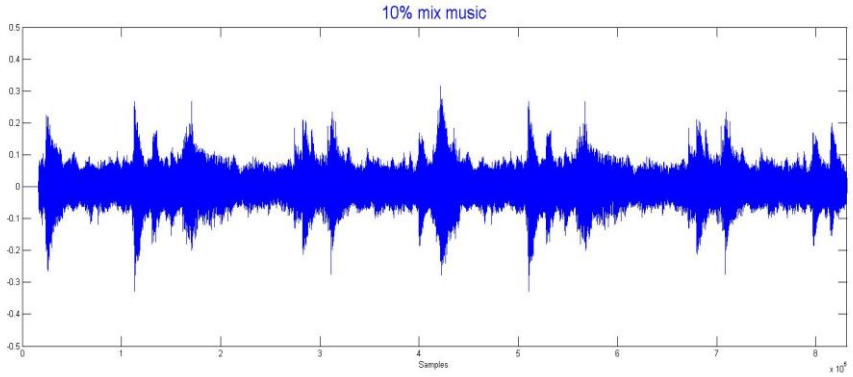
A reliable decoding of the watermark highly depends on the accuracy of the estimation of the scaling factor. As one needs a rather big number of samples to build a meaningful histogram, problems arise when the scaling factor is not constant over the analysis frame. Additionally, the current estimation algorithms are known to not work well for non-linear amplitude scaling. A general approach to face the amplitude scaling problem is the embedding of the watermark in a signal domain which is not sensitive to amplitude scaling. For example the fundamental frequency, the pitch or the duration of a speech phoneme are inherently robust to amplitude scaling [7].

Further research is necessary though to achieve sufficient data rates. Another strategy used to counter the volumetric scaling problem is the encoding of the watermark data with codes that are resilient to the attack. Among many others, [8] and [9] present watermarking schemes based on Trellis coding. Trellis coding is used to compress and clean communications signals to allow greater data rates and robustness. By integration of a convolutional code with a bandwidth efficient modulation, significant coding gain can be achieved compared to un-coded schemes, without sacrificing data rate or requiring more channel bandwidth. A short introduction to signal coding theory and a thorough overview on Trellis coding can be found in [10]. Its extension and application to watermarking as dirty paper trellis codes is shown in [11].

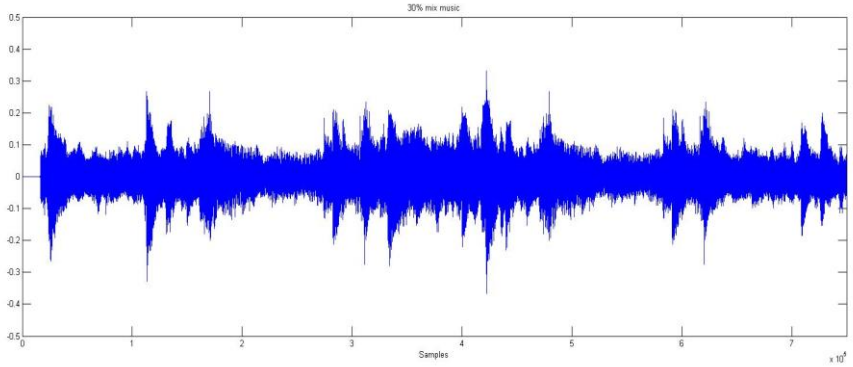
The modified Trellis codes take the host signal into consideration not only at the embedding, yet already at the watermark encoding stage, this further increasing the possible data-rate. For decoding Trellis-coded signals, usually a Viterbi-decoder [12] is used, which considers the entire watermark message instead of single bits. As the detection is based on relative correlation values, the mechanism is inherently robust to amplitude scaling.

4. Experimental Results

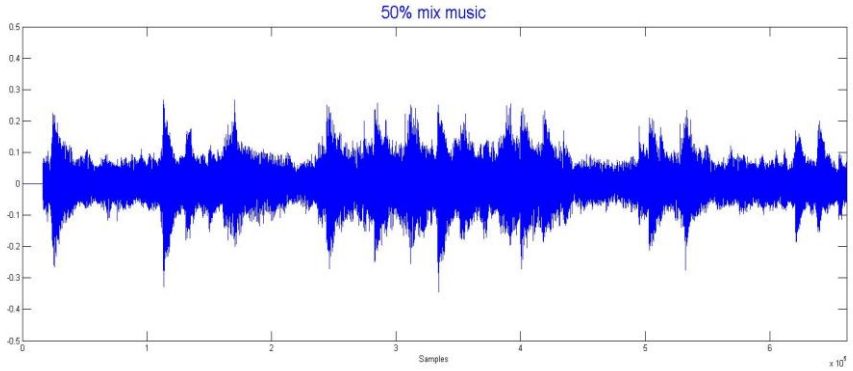
To our experiment, we prepared 3 songs that have overlapping of 10%, 30%, and 50% as Figure 1. The songs used to overlap are the same as the former and the latter. The sampling rate of each song is 44.1 KHz that is mono- CD audio quality.



(a) 10% overlapped



(b) 30% overlapped



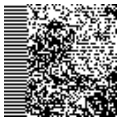
(c) 50% overlapped

Figure 3. 10%, 30% and 50% Overlapped Mix Music that Combine to Two Songs. The Sampling Rate of Each Song is 44.1 KHz, CD Audio Quality

To test the applicability of existing audio watermarking technology, four watermarking techniques, one discrete cosine (DCT) domain watermarking algorithm, three discrete wavelet transform (DWT) domain watermarking algorithm are applied. DCT algorithm embeds 64 x 64 watermark image in low frequency element of 2 x 2 DCT block. The first DWT domain watermarking algorithm

embeds 24 x 24 watermark image in middle frequency bands. The second DWT domain watermarking algorithm embeds 64 x 64 watermark image in middle frequency bands by different embedding process from the first DWT algorithm and the last DWT algorithm is the implementation of dither method in DWT domain. The four watermarking technologies are applied to 10% overlapped mix music. Table 1 shows the experimental results of watermark extraction described four watermarking algorithms.

Table 1. Extracted Watermarks from 10% Overlapped Mix Music

Technologies	Extracted watermark (former)	Extracted watermark (latter)
DCT1		
DWT1		
DWT2		
DWT3		

As seeing in Table 1, the last DWT domain watermarking algorithm which is based on dither method uniquely extracts watermarks from mix music that combine two songs. And the first DWT domain watermarking algorithm perfectly extracts the first watermark from the first song and the second watermark with some bit errors. Other algorithms fail to extract the second watermark from the second song of underlying mix music. The common of failed algorithm to extract watermark is that the first watermark is normally extracted but the second watermark is not. Even two algorithms, DCT1 and DWT2 in Table 1 never extract the second watermark. In case of DWT1 algorithm, the second watermark is not extracted from 20% overlapped mix music as in Table 2. The results of two failed algorithms, DCT1 and DWT2, and the result at Table 2 imply that the typical audio watermarking technology applied a non- mix music *i.e.*, the typical music may not be apply the mix music consistently. As the copyright law changes based on the changes of the popular music genre by modern music trend and/or birth of new music genre, the field of audio watermarking technology for supporting the copyright law should change.

Table 2. Extracted Watermarks from 10% Overlapped Mix Music by DWT1 Algorithm

Technologies	Extracted watermark (former)	Extracted watermark (latter)
DWT1		

4. Conclusions

In this paper, we evaluated the performance of four typical audio watermarking technologies applied mix music. The experimental result showed that only one audio watermarking technology based on dither method in the DWT domain perfectly extracted the watermark from mix music. The common problem of other audio watermarking technologies cannot extract the second watermarking from the underlying mix music. This mean that the advent of mix music requires a new watermarking techniques. Our future works is improving the typical watermarking technology to be suitable for the mix music. Especially we are concentrating on the development of an audio watermarking technology to be able to survive at the overlapped areas of mix music.

Acknowledgments

This research project was supported by Ministry of Culture, Sports and Tourism (MCST) and from Korea Copyright Commission in 2015.

References

[1] K. K. Parhi and T. Nishitani Editors, “Digital Signal Processing in Multimedia Systems,” Marcell Dekker Inc. **(1999).**
[2] H. Emily, “Music Mashups: Testing the Limits of Copyright Law as Remix Culture Takes Society by Storm,” Hofstra Law Review, vol. 39, Iss. 2, **(2010).**
[3] Komal V. Goenka1, Pallavi K. Patil, “Overview of Audio Watermarking Techniques,” International Journal of Emerging Technology and Advanced Engineering, vol. 2, **(2012)**, pp. 67-70.
[4] B. Chen, “Design and Analysis of Digital Watermarking, Information Embedding, and Data Hiding Systems,” PhD thesis, Massachusetts Institute of Technology, June **(2000)**.
[5] B. Chen and G. W. Wornell. Quantization index modulation, “A class of provably good methods for digital watermarking and information embedding”, IEEE Transactions on Information Theory, vol. 47, no. 4, **(2001)**.
[6] B. Chen and G. W. Wornell,”Quantization index modulation methods for digital watermarking and information embedding of multimedia,” Journal of VLSI Signal Processing, vol. 23, **(2001)**, pp. 7–33.
[7] I. D. Shterev, I. L. Lagendijk, and R. Heusdens, “Statistical amplitude scale estimation for quantization-based watermarking,” Proceedings of SPIE, Security, Steganography and Watermarking of Multimedia Contents VI, **(2004)**.
[8] E. Esen, A. A. Alatan, and M. Askar, “Trellis coded quantization for data hiding,” Proceedings of the IEEE EUROCON, **(2003)**.
[9] M. L. Miller, G. J. Doërr, and I. J. Cox, ”Applying informed coding and embedding to design a robust, high capacity watermark,” IEEE Transactions on Image Procesing, vol. 13, no. 6, **(2004)**.
[10] M. Celik, G. Sharma, and A. M. Tekalp, “Pitch and duration modification for speech watermarking,” Proceedings of the 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing, **(2005)**.
[11] C. B. Schlegel and L. C. Pérez Editors “Trellis and Turbo Coding,” IEEE Press Series on Digital & Mobile Communication. Wiley-IEEE Press, **(2004)**.
[12] M. L. Miller, G. J. Doërr, and I. J. Cox, “Dirty paper trellis codes for watermarking,” Proceedings of the 2002 IEEE International Conference on Image Processing, **(2002)**.

Authors

Youngseok Lee, he received the Ph.D. degree from University of Seoul, major in signal processing in 1998. He is currently a professor of Dept. of Electronic Engineering at Chungwoon University in Korea. His research interests are in the areas of copyright protection technology, biometric authentication and machine learning system.

Jongweon Kim (Corresponding Author), he received the Ph.D. degree from University of Seoul, major in signal processing in 1995. He is currently a professor of Dept. of Contents and Copyright at Sangmyung University in Korea. He has a lot of practical experiences in the digital signal processing and copyright protection technology in the institutional, the industrial, and academic environments. His research interests are in the areas of copyright protection technology, digital rights management, digital watermarking, and digital forensic marking.

.