Bit Replacement Audio Watermarking Using Stereo Signals

Shaoliang Meng, Wei Cao, Yixin Yan, Shengming Li
The higher educational key laboratory for Measuring & Control Technology and Instrumentations of Heilongjiang Province
shiyu_wang@126.com
doi: 10.4156/aiss.vol2.issue3.3

Abstract

We have seen an explosion of data exchange in the Internet and the extensive use of digital media. Consequently, digital data owners can quickly and massively transfer multimedia documents across the Internet. This leads to wide interest in multimedia security and multimedia copyright protection. In this paper, a blind audio watermarking scheme using stereo signals is proposed. In the proposed scheme, each bit of watermark bit stream is embedded into "non–silent" samples of stereo audio signals. The watermark bits can be embedded by replacing least significant bit or bits in higher bit layers of original signal samples. Robustness and performance of this audio watermarking scheme under different types of attacks, such as addition of noise, MP3 compression and re-sampling are investigated. Watermark extraction rates are calculated. Different audio signals including music, speech are used for the experiments. It is found that the proposed audio watermarking scheme, compared with the use of single-channel signals, is much more robust against addition of white noise.

Keywords: *Watermark, Stereo signal, Robustness*

1. Introduction

The outstanding process of digital technology has increased the ease with which multimedia data are reproduced and retransmitted. Since the advantaged of such a process are broadly available, they offer equally increasing potential to both legal and unauthorized data manipulation. Consequently, the necessity arises for copyright against unauthorized copying and redistribution, known as data piracy. The problem of copyright protection of multimedia products has attracted the interest of scientists and entrepreneurs alike. The most promising solution to this challenging task seems to lie in data hiding techniques, i.e., methods for secretly and imperceptibly embedding signals of certain characteristics in the media of interest.

Digital watermarking is a process of embedding additional secret information into digital signals such as audio, image or video signals. If the watermarked signals are copied, then the embedded information is also carried in the copy. The embedded information or watermark is usually in binary format and the quality of original signals should not be affected by the embedded watermark. The watermark bits are not apparent to any unauthorized user and they can be used to prove the ownership of watermarked signals. The watermark should be robust and easily extracted from watermarked signals even after some incidental and intentional attacks, such as addition of noise, re-sampling and MP3 compression.

For digital audio watermarking, the approach can be non-blind or blind. For non-blind watermarking technique, original signals are required for watermark extraction. On the other hand, for blind watermarking technique, original signals are not required for watermark extraction. In this paper, a blind audio watermarking technique is described.

A review of the multitude of existing multimedia watermarking techniques can be found in [1] and [2]. The obvious conclusion from reading these review papers is that research on audio watermarking is not as mature compared to image watermarking. Indeed, only few publications on audio watermarking exist, most of them in the patent literature. Among many audio watermarking schemes, the approach using least significant bit (LSB) for watermarking is a powerful method as it enables high rate of information hiding and ease of implementation. It is a simple approach in watermarking of the audio sequences, where the embedding of a watermark into a digital audio stream is performed by alternation of the LSBs, having the amplitude resolution of 16 bits per sample. But robustness of this method is very low. The watermark bits can be easily detected and changed. Watermark extraction rate is also very low when watermarked signals are subjected to different types of attacks. Instead of using LSB of mono-signal samples, it is perceived that by using higher bits of stereo-signal samples, it can be more robust against attacks. In this paper, a new watermark embedding algorithm using stereo signals is proposed and its robustness is presented.

The organization of this paper is as follows, watermark embedding algorithm is explained in Section 2. Effects of watermarking on original stereo signals are detailed in Section 3. The robustness of proposed audio watermarking scheme under different types of attacks is shown in Section 4 and 5; followed by conclusion in the last section.

2. Watermark embedding algorithm

Numerous steganography techniques have been designed and implemented, which is due the fact that multimedia objects have a highly redundant representation. In multimedia compression techniques data redundancy is used in order to decrease the number of bits per sample needed for perceptually transparent representation of multimedia content. On the other hand, data redundancy is used in steganography techniques to embed additional data into the cover object, without causing perceptual distortion. As already noted, one of the most common and simplest steganographic method is the LSB coding method that embeds a message in the least significant bits of the host audio. As the number of used LSBs increases or the depth of the modified LSB layer becomes larger, the risk of making the embedded message statistically detectable increased and perceptual transparency of stego objects is decreased. Therefore, there is a limit for the number of bits in each sample of host audio that can be used to embed messages.

For LSB watermarking method[3-8], only the least significant bits of each mono-signal samples are replaced by watermark bits and bits in higher bit layers remain unchanged. In the watermark embedding algorithm proposed by N. Cvejic and T. Seppänen[5,6], bits in higher bit layers are used for watermark embedding. To minimize the changes in original sample values due to embedding watermark in higher bit layers, bits in other bit layers of the same sample have also to be changed. This mono-signal watermarking method is only robust when watermarked signals are subjected to low level noise. Other types of attacks, such as addition of high level noise, re-sampling and MP3 compression may completely destroy the watermark information[9-13].

The proposed audio watermarking scheme improves above mono-signal watermarking method by introducing watermarking threshold and making use of stereo signals. The result is a slight amplitude modification of each audio sample in a way that does not produce any perceived effect[14-19].

The watermarking threshold is calculated using the following equation:

$$\text{Watermarking threshold} = \frac{1}{2^{n-(w+4)}} \tag{1}$$

Where n stands for the number of bits per sample and w stands for the watermark embedded bit layer.

The samples whose values are higher than the threshold are considered as “non-silent” samples and they are used to embed watermark bits. The samples whose values are lower than the threshold are considered as “silent” samples and they are not used for watermark embedding. This procedure is suggested as the original values of “silent” samples are very small. Embedding watermark in “silent” samples will significantly change the values and cause much additional noise, the quality of original audio signals will be severely affected, hence “silent” samples are not used in watermark embedding process. After calculation of the threshold, let each “non-silent” sample value of original stereo signals be represented in 16-bit binary format, $a_{16},a_{15},a_{14},a_{13},a_{12},a_{11},a_{10},a_9,a_8,a_7,a_6,a_5,a_4,a_3,a_2,a_1,$ where a_1 is the bit in the 1st bit layer; a_2 is the bit in the 2nd bit layer and so on. A watermark bit stream is first generated and ith bit layers of “non-silent” samples are used as watermark embedded bit layer. If the first bit of the stream is bit “1”, and a_i of the first “non-silent” sample of left-channel signal is also bit “1”, no action is taken. If a_i is not “1”, watermark embedding process of left-channel signal is performed according to the following procedures,

$a_i = 1$
if $a_{i-1}=1$ then $a_{i-1}\, a_{i-2} \ldots a_1 = 00\ldots0$
if $a_{i-1}=0$ and
if $a_{i+1}=1$ then $a_{i+1}=0$ and $a_i a_{i-1} \ldots a_1 = 11\ldots1$
else if $a_{i+2}=1$ then $a_{i+2}=0$ and $a_{i+1} a_i \ldots a_1 = 11\ldots1$
else if $a_{i+3}=1$ then $a_{i+3}=0$ and $a_{i+2} a_{i+1} \ldots a_1 = 11\ldots1$
.
.
else if $a_{16}=1$ then $a_{16}=0$ and $a_{15} a_{14} \ldots a_1 = 11\ldots1$

If the first bit of the stream is bit "0", and a_i of the first "non-silent" sample of right-channel signal is also bit "0", no action is taken. If a_i is not "0", watermark embedding process of right-channel signal is performed according to the following procedures,

$a_i = 0$
if $a_{i-1} = 0$ then $a_{i-1}\, a_{i-2} \ldots a_1 = 11\ldots1$
if $a_{i-1} = 1$ and
if $a_{i+1}=0$ then $a_{i+1}=1$ and $a_i a_{i-1} \ldots a_1 = 00\ldots0$
else if $a_{i+2}=0$ then $a_{i+2}=1$ and $a_{i+1} a_i \ldots a_1 = 00\ldots0$
else if $a_{i+3}=0$ then $a_{i+3}=1$ and $a_{i+2} a_{i+1} \ldots a_1 = 00\ldots0$
.
.
else if $a_{16}=0$ then $a_{16}=1$ and $a_{15} a_{14} \ldots a_1 = 00\ldots0$

The same watermark embedding procedures are applied to the second bit of the stream and so on until the last "non-silent" sample of stereo signals.

3. Effects of watermarking on original signal

From above watermark embedding process, it can be seen that besides watermark embedded bit layer, other bit layers also have a lot of changed bits. Hence the bit layer used for watermarking cannot be exactly identified and watermark cannot be easily changed or removed. The method is applied to a stereo audio signal and Fig. 1 shows the percentage of changed bits in different bit layers as a result of embedding watermark in the 5^{th} bit layers of "non-silent" samples.

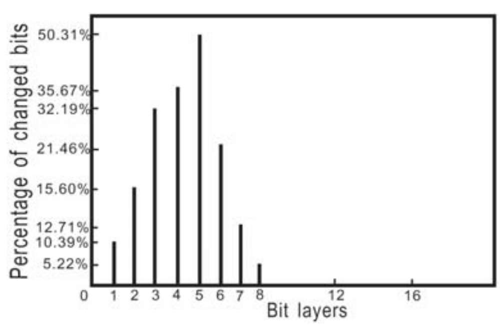

Figure 1. Percentage of changed bits as embedding in the 5^{th} bit layer

When watermark bits are embedded in original stereo signals, the original sample values are changed. This is equivalent to adding noise to original signals. An experiment is conducted to investigate signal-to-noise ratio (SNR) of watermarked signals due to embedding watermark bits in different bit layers. The experiment is conducted using a 15 second stereo music signal sampled at 44,100 samples per second with each sample having 16 bits. The proposed watermarking scheme is applied to the music signals. A watermark bit stream is embedded in Layer 1 to Layer 5 of "non-silent" samples and generate 5 different watermarked stereo signals. The mono-signal watermarking method is also applied to one single-channel signal of the music signal. The same watermark bit stream is embedded in Layer 1 to Layer 5 of each mono-signal samples and generate 5 different watermarked mono signals.

Mathematically, SNR for the watermark embedded in time-domain is calculated as follows,

$$SNR = 10\log_{10}\left\{\sum_n x^2(n) / \sum_n \left|x^2(n) - y^2(n)\right|\right\} \quad (2)$$

where $x(n)$ stands for a sample value of original signals and $y(n)$ stands for a sample value of watermarked signals. The SNRs of all 5 watermarked stereo signals and 5 watermarked mono signals are shown in Table 1, 2.

Table 1. SNR(dB) of watermarked stereo signals

Signal-to-Noise Ratios	Embedded Bit Layer of Watermarked Signals				
	1st	2nd	3rd	4th	5th
SNR (dB)	72.09	70.15	67.38	59.36	46.25

Table 2. SNR(dB) of watermarked mono signals

Signal-to-Noise Ratios	Embedded Bit Layer of Watermarked Signals				
	1st	2nd	3rd	4th	5th
SNR (dB)	66.52	61.49	56.81	44.37	28.53

It can be observed that each SNR of different embedded bit layer for the proposed watermarking scheme is much higher than that for mono-signal watermarking method. For mono-signal watermarking method, the SNR of 5^{th} embedded bit layer is 28.53dB. However, informal listening test shows that the quality of the watermarked audio signal is already affected. When the proposed scheme is used, additional noise level due to embedding watermark is significantly reduced. The quality of watermarked signals is also improved.

4. Robustness under attack of addition of noise

To assess the robustness of proposed audio watermarking scheme under attack of noise, white noise with different magnitude are added to the 5 watermarked stereo signals and 5 watermarked mono signals mentioned above, and obtain watermarked signals with resulting SNRs of 20dB and 50dB for each of these watermarked signals. The watermark bits can be extracted by reading the bits in the embedded bit layers. Percentages of correct extracted bits are calculated for each of these watermarked signals. The watermark extraction rates are tabulated in Table III, IV and plotted in Fig. 2, 3.

Table3. Watermark extraction rates(%) for watermark signals with 20dB resulting SNR

Watermark Extraction Rate (%)		Audio Watermarking Method	
		Stereo-signal	Mono-signal
Embedded Bit Layer	1st	50.16	49.97
	2nd	57.82	49.99
	3rd	61.96	50.11
	4th	69.55	50.21
	5th	73.47	50.32

Table 4. Watermark extraction rates(%) for watermark signals with 50dB resulting SNR

Watermark Extraction Rate (%)		Audio Watermarking Method	
		Stereo-signal	Mono-signal
Embedded Bit Layer	1st	50.39	50.23
	2nd	69.13	62.81
	3rd	75.65	65.73
	4th	81.97	71.50
	5th	91.52	73.36

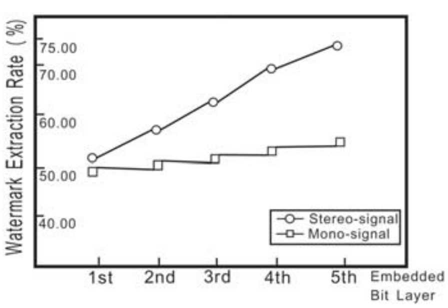

Figure 2. Watermark Extraction Rates of Table 3

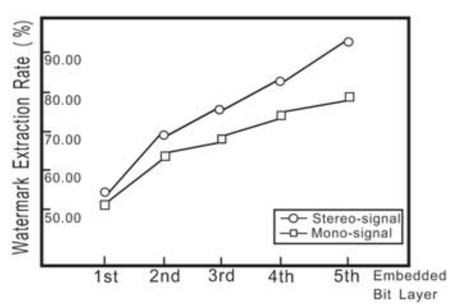

Figure 3. Watermark Extraction Rates of Table 4

When no white noise is added, the watermark extraction rates are 100% for both methods. When the level of added white noise is high such that the resulting SNR is low, the extraction rates for mono-signal method are very low. For example, when the added noise level is high such that the resulting SNR is only 20dB, the watermark extraction rates for mono-signal method are around 50%, which is the level of random guess for binary data. But for the proposed stereo-signal method, watermark extraction rates are significantly increased across 5 embedded bit layers. A rate of 73.47% can be achieved when 5^{th} bit layer is embedded. When added white noise level is reduced, for example, the resulting SNRs of watermarked signals are 50dB, the extraction rates are increased. When 5^{th} bit layer is embedded, the rate for proposed stereo-signal method is 91.52%, which is 18.16% higher than that for mono-signal watermarking method.

Similar results are obtained when the proposed watermarking scheme is applied to a stereo speech signal and they are shown in Table V. This shows that the robustness of proposed scheme is not significantly affected by the types of audio signals.

Table 5. Watermark extraction rates(%) for a speech signal

Watermark Extraction Rate (%)		Embedded Bit Layer of Watermarked Signals				
		1st	2nd	3rd	4th	5th
Resulting SNR(dB)	No white noise	100	100	100	100	100
	20	50.22	55.61	62.72	70.36	72.79
	50	50.28	69.04	73.24	83.65	90.07

5. Robustness under MP3 compression and resampling attacks

Experiments are carried out to test the robustness of proposed audio watermarking scheme when watermarked stereo signals are under attack of MP3 compression. The watermarked audio signal is compressed to MP3 format. After compression, the sampling rate of watermarked signal is reduced to 1,400 samples per second and the number of samples is also reduced. Then the MP3 format is converted back to the original format. After conversion, the number of samples is increased to original value. The above 5 watermarked stereo signals without any white noise are used again in these experiments. The watermark exaction rates are calculated and they are given in Table 6.

Table 6. Watermark extraction rates(%) under attack of MP3 compression

ExtractionRate (%) for MP3 Compression	Embedded Bit Layer of Watermarked Signals				
	1st	2nd	3rd	4th	5th
Extraction Rate (%)	49.80	49.82	49.83	50.06	50.11

MP3 compression is lossy[20]. The extraction rates are only about 50%. As 50% detection of binary data is equivalent to random guess, the watermark information is completely destroyed. It can be concluded that MP3 compression completely destroys the watermark information and the proposed watermarking scheme is not robust under attack of MP3 compression.

Experiments are also carried out to test the robustness of proposed audio watermarking scheme when watermarked stereo signals are under attack of re-sampling. The watermarked stereo signals are first re-sampled at a different frequency and then re-sampled back at 44.1 kHz which is the original sampling frequency. After re-sampling, the length of audio signal and the number of samples are not changed. The watermark exaction rates are calculated and they are shown in Table 7.

Table 7. Watermark extraction rates(%) under attack of re-sampling

Watermark Extraction Rate (%)		Embedded Bit Layer of Watermarked Signals				
		1st	2nd	3rd	4th	5th
Resampling Frequency (kHz)	14.70	48.99	49.54	49.83	49.71	50.11
	22.05	49.61	49.73	50.17	50.66	50.70
	66.15	49.17	49.84	49.96	50.30	50.43
	88.20	50.03	53.29	55.64	58.52	60.79

From above table, it can be seen that the watermark extraction rates are round 50% except when re-sampling frequency is 88.2 kHz which is an integer multiple of original sampling frequency, hence it can be concluded that the proposed scheme is not robust under attack of re-sampling unless the re-sampling frequency is an integer multiple of original sampling frequency.

Subjective quality evaluation of the watermarking method has been done by listening tests involving ten persons. Three of them had basic or medium level music education or are active musicians. In the first part of the test, the participants were repeatedly presented with unwatermarked and watermarked audio clips in random order and were asked to determine which one is the watermarked one (blind audio watermarking test). Values near 50% show that two audio clips (original audio sequence and watermarked audio signal) cannot be discriminated by people that participated in the listening tests.

6. Conclusions

Standard LSB watermarking scheme is a powerful method of information hiding and the complexity of implementation is low. However, if the least significant bit layer is used for watermarking, robustness is very low. Even if higher bit layers of each mono-signal samples are used for watermark embedding, it is only robust when low level white noise is added to the watermarked signals. Some other types of attacks, such as addition of high level noise, MP3 compression and re-sampling, will completely destroy the watermark information.

In this paper, a blind audio watermarking scheme is proposed. The scheme embeds watermark bits into "non-silent" samples of stereo signals. The noise level due to embedding watermark is significantly reduced. The proposed scheme is very robust against addition of white noise. Even the added noise level is high, the watermark extraction rate can still reach about 73% when 5^{th} bit layer is embedded, which is 23% higher than that when mono-signal watermarking method is used.

The proposed audio watermarking scheme, however, as expected of bit replacement watermarking schemes, is not robust against MP3 compression and re-sampling unless the re-sampling frequency is an integer multiple of original sampling frequency.

7. Acknowledgement

This work was supported in part by the China Postdoctoral Science Foundation (20090450953).

8. References

[1] F.Hartung and M. Kutter, "Multimedia watermarking techniques",Proc. IEEE, vol.87, no.7,pp.1-97-1107, July 1999.
[2] M.D. Swanson, M. Kobayashi, and A.H..Tewfik, "Multimedia data embedding and watermarking technologies", Proc, IEEE, vol. 86, no.6, pp.1064-1087, June 1998.
[3] Wu C.P., Su P.C. & Kuo C.J. "Robust Audio Watermarking for Copyright Protection". SPIE's 44th Annual Meeting Advanced Signal Processing Algorithms, Architectures, and Implementations IX (SD39), 1999.
[4] N. Cvejic and T. Seppänen, "Increasing robustness of LSB audio steganography by reduced distortion LSB coding," Journal of University Computer Science, vol 11, pp56, 2006.
[5] N. Cvejic, and T. Seppänen, "Digital Audio Watermarking Techniques and Technologies," Information Science reference, New York, pp.1-11, 2007.
[6] Hsieh C. T. & Tsou P. Y. "Blind Cepstrum Domain Audio Watermarking Based on Time Energy Features". 4th Int. Conf. on Digital Signal Processing, pp.705-708,2002.
[7] Garcia R.A. "Digital Watermarking of Audio Signals Using a Psychoacoustic Auditory Model and Spread Spectrum Theory". 107th Convention, Audio Engineering Society, preprint 5073, 1999.
[8] Kim H. J., Choi Y.H., Seok J.W. & Hong J.W. "Audio Watermarking Techniques: Intelligent Watermarking Techniques". Chapter 8, pp.185-218, 2004.
[9] Li W., Xue X.Y. & Li X.Q. "Localized Robust Audio Watermarking in Regions of Interest". ICICS-PCM 2003.

[10] Neubauer C. & Herre J. "Advanced Audio Watermarking and its Application". 109th AES Convention, AES preprint 5176, 2000.
[11] Pohlmann, K.C. (1991). Advanced Digital Audio, Carmel, 1991. ISBN: 0672227681.
[12] Swanson M.D., Zhu B., Tewfik A.H. & Boney L. Robust audio watermarking using perceptual masking. Signal Processing, Vol.66, No.3, pp.337-355, 1998.
[13] Vladimir, B. & Rao K.R. "An Efficient Implementation of the Forward and Inverse MDCT in MPEG Audio Coding". IEEE Signal Processing Letters, Vol.8, No.2, 2001,
[14] Xu C., Wu J., Sun Q. & Xin K. "Applications of Watermarking Technology in Audio Signals". Journal Audio Engineering Society, Vol. 47, No. 10, 1999,
[15] Erkucuk. s.,Krishnan S.and Zeytinoglu M., "A Robust Audio Watermark Representation Based on Linear Chirps", IEEE Transactions on Multimedia, Vol.8, Issue 5, pp.925-936,2006.
[16] Xiangyang Wang, Wei Qi and Panpan Niu, "A New Adaptive Digital Audio Watermarking Based on Support Vector Regression", IEEE Transactions on Audio, Speech and Language Processing, Vol. 15, Issue 8, pp.2270-2277, 2007.
[17] Malik H, Ansari R and Khokhar A, "Robust Audio Watermarking Using Frequency-selective Spread Spectrum", IET Information Security, Vol.2, Issue 4, pp.129-150,2008.
[18] Kalantari N.K, Akhaee M.A, Ahadi S.M and Amindavar H, "Robust Multiplicative Patchwork Method for Audio Watermarking", IEEE Transactions on Audio, Speech and Language Processing, Vol. 17, Issue 6, pp.1133-1141, 2009.
[19] Shijun Xiang and Jiwu Huang, "Histogram-Based Audio Watermarking Against Time-Scale Modification and Cropping Attacks", IEEE Transactions on Multimedia, Vol. 9, Issue 7, pp.1357-1372, 2007.
[20] Chen O.T.-C and Chia-Hsiung Liu, "Content-Dependent Watermarking Scheme in Compressed Speech with Identifying Manner and Location of Attacks", IEEE Transactions on Audio, Speech and Language Processing, Vol. 15, Issue 5, pp.1605-1616, 2007.