Analysis of D/A and A/D Conversions in Quantization-based Audio Watermarking

Shijun Xiang[12] and Jiwu Huang[12]
(Corresponding author: Jiwu Huang)

School of Information Science and Technology, Sun Yat-Sen University[1]
Guangzhou 510275, P. R. China. (Email: isshjw@zsu.edu.cn)
The key laboratory of Information Security Technology, Sun Yat-Sen University[2]
Guangzhou 510275, P. R. China.

(Received July 16, 2005; revised and accepted Aug. 12, Sept. 5, & Oct. 4, 2005)

Abstract

Due to its advantages, quantization-based embedding has been introduced into audio watermarking to improve robustness performance. Existing audio watermarking algorithms often focus on a given attack. However, in some transmission environments, digital audio files may suffer from the different attacks. For example, effects of D/A and A/D conversions (denoted as DA/AD in this paper) on audio watermarking may be modeled as modification of signal energy, phase changes, and noises corruption. These effects present an important challenge to audio watermarking. Extensive experiments show that phase changes of audio signals caused by the DA/AD may be represented as some extent of temporal scaling. Furthermore, we analyze the performance of quantization-based audio watermarking against these attacks in the DA/AD, and present corresponding calculation expressions of *BER* (Bit Error Rate) in case of Gaussian noise and modification of audio amplitude, and then investigate the influences of temporal scaling caused by the DA/AD on audio watermarking. As a conclusion, audio watermarking quantization-based is very susceptible to the DA/AD.

Keywords: Audio watermarking, temporal scaling, quantization-based embedding, DA/AD

1 Introduction

According to IFPI (International Federation of the Phonographic Industry) [10], audio watermarking, at a certain data payload or data embedding capacity of more than 20 bps and under the imperceptibility constraint (Signal-to-Noise Ratio, *SNR*, should be higher than 20 dB), should be able to resist the most common signal processing manipulations and attacks, such as temporal scaling (stretching by 10%), additive and multiplicative noise corruption, MP3 compression, re-sampling, re-quantizing, DA/AD.

Most of the recent audio watermarking algorithms can be grouped into two categories: additive and quantization-based watermarking schemes. The additive scheme [6] embeds the watermark in time domain [1] or frequency domain [5], while in the quantization-based scheme [3, 8, 11, 12, 22, 24], the original signals are quantized by different quantizers varied with watermark information to embed watermarks rather than the simple addition to the original signal. In the extraction, the information can be recovered according to the distances between incoming data and different quantization results. Due to its merits, such as blind extraction and better robustness performance etc., quantization-based watermarking becomes a dominant method.

There are mainly two types of attacks on watermarked audio: 1) Modify the amplitude of audio signal. It results in the lost of parts of hided information. These attacks include noise corruption, amplitude scaling, re-sampling, and MP3 compression, etc.; 2) Destroy synchronization of the watermark in time domain. This kind of attack is more effective than corrupting watermarked audio amplitude directly, such as time scaling, shifting, cropping.

The previous works on quantization-based audio watermarking mainly focused on a proposed attack, such as temporal scaling [12], amplitude modification [11], MP3 compression [8], synchronization attack analysis [24], and then proposed a corresponding watermark technique.

Existing audio watermarking embedding and detection strategies often depend on digital channels like CD, MP3 and IP network transmission. From the view point of applications, however, robustness of audio watermarking against the DA/AD is an important issue [14]. For instances, in many applications [4, 7, 13, 17, 21], watermarks are required to survive in analog environments, in which the DA/AD may be involved. Secret data is proposed to be transmitted via analog telephone channel in [4] as well as for many applications. For instance, hidden watermark signal may be used to identify pirated music through speaker and PC soundcard and for broadcast music [7, 13] and live concert [21] monitoring. Embedded audio data can also be used to control toys [17].

It is noted that it was recently claimed [4, 7, 9, 13, 16, 17, 18, 21] that the watermarks embedded are robust to the attacks including the DA/AD. However, the *BER*s in these reports are either not reported [4, 13, 17, 21] or rather high [7, 9, 16, 18]. Furthermore, there are no any technical descriptions on how to resist the DA/AD. Specifically, none of them have reported how to resist the changes caused by the DA/AD in details. It is also noted that there are no test functions in StirMark Benchmark for Audio [19, 23] that have been designed for evaluating the robustness of audio watermarking to the DA/AD. In summary, to the authors' best knowledge, among all of the literature on audio watermarking, only [20] has discussed the effects caused by the DA/AD on audio signals, but failed to deduct the main degradations during the DA/AD.

According to the references [2, 15, 20], the serious degradation of audio signal caused by the DA/AD includes modification of amplitude and phase simultaneously, which still puzzles the watermark extraction. Therefore, the DA/AD is considered a challenging issue for audio watermarking [20].

As a dominant watermark scheme, it is necessary to analyze the performance of quantization-based audio watermarking to the DA/AD attack by investigating the main distortion caused by the DA/AD.

In this paper, we first introduce the quantization-based audio watermarking model, and then discuss the attacks existing in different kinds of transmission environments. For instance, the attacks in the DA/AD include noise corruption, amplitude, and phrase change. Based on extensive experiments, it is fount that phrase changes caused by the DA/AD may be modeled as temporal scaling. Furthermore, we analyze the performance of quantization-based audio watermarking against the attacks cause by the DA/AD, and present the expressions of *BER* against Gaussian noise and amplitude change attacks. We also investigate the effects of the temporal scaling on audio watermarking. Finally we draw the conclusions that quantization-based audio watermarking is susceptible to the DA/AD processing.

The rest of this paper is organized as follows. Section 2 introduces the quantization-based watermarking model and analysis the possible attacks in some transmission environments. In Section 3, we investigate the interferences caused by the DA/AD by conducting a great deal of experiments. The performance of quantization-based audio watermarking against these attacks is discussed in Section 4. Section 5 presents the experimental results to illustrate the expressions of *BER* in case of Gaussian noise corruption and the modification of audio amplitude. Finally, the conclusions are drawn in Section 6.

2 Transmission Environments

2.1 Quantization-based Watermarking Model

Let vector $X = \{x_1, x_2, \ldots, x_N\}^T$ denote the cover-signal (such as digital image, audio, etc.), and perform linear transform to X:

$$Y = \{y_1, y_2, \ldots, y_N\}^T = UX$$

where U is $N \times N$ dimensions unitary matrix. DFT, DCT and DWT are special forms of unitary transformation. If U is an identity matrix, then the information is hidden in the time domain of the cover-signal.

Let $W = \{w_i\}, i = 1, 2, \ldots, M$ denote the to-be hidden binary sequence. Use the key to choose M coefficients from Y, denoted as $C = \{c_1, c_2, \ldots, c_M\}^T$. Accordingly, one information bit is embedded into a transform coefficient by using quantization watermarking scheme. Quantization watermarking embedding is introduced to hide information into digital image in [22]. As a dominant method, quantization watermark scheme was verified feasible for digital multimedia [3], such as digital audio information hiding [8, 11, 12, 24]. The quantization embedding formula may be expressed as follows:

$$\begin{cases} c_i' = c_i - Mod(c_i, S) + T_1, & \text{if } c_i \geq 0 \text{ and } w_i = 1 \\ c_i' = c_i - Mod(c_i, S) + T_2, & \text{if } c_i \geq 0 \text{ and } w_i = -1 \\ c_i' = c_i - Mod(c_i, S) - T_1, & \text{if } c_i < 0 \text{ and } w_i = 1 \\ c_i' = c_i - Mod(c_i, S) - T_2, & \text{if } c_i < 0 \text{ and } w_i = -1 \end{cases} \quad (1)$$

where the different values of the thresholds of T_1 and T_2 indicate the differences of embedding "1" or "-1". S is quantization step, namely the embedding strength. The value of S should be as large as possible under the imperceptibility constraint. c_i and c_i' are the i^{th} coefficient in the original signal and the watermarked signal, respectively. $Mod(c_i, S)$ denotes the remainder of c_i divided by S. Usually, $T_1 = 3S/4$, $T_2 = S/4$. Inverse transformation is performed after the watermarked signal is generated.

In watermark extraction, the same unitary transformation is performed. After c_i' is figured out by the key, in [22] the following formula is utilized to extract watermark:

$$w_i' = \begin{cases} 1 & \text{if } Mod(c_i', S) \geq (T_1 + T_2)/2 \\ -1 & \text{if } Mod(c_i', S) < (T_1 + T_2)/2 \end{cases} \quad (2)$$

where w_i' is the extracted watermarked bit.

In [8, 24], Equations (1) and (2) were simplified by assigning $T_1 = 3S/4$, $T_2 = S/4$, as shown in Equations (3) and (4):

$$c_i' = \begin{cases} \lfloor (c_i/S) \rfloor \cdot S + 3S/4 & \text{if } w_i = 1 \\ \lfloor (c_i/S) \rfloor \cdot S + S/4 & \text{if } w_i = -1 \end{cases} \quad (3)$$

$$w_i' = \begin{cases} 1 & \text{if } c_i' - \lfloor (c_i'/S) \rfloor \cdot S \geq S/2 \\ -1 & \text{if } c_i' - \lfloor (c_i'/S) \rfloor \cdot S < S/2 \end{cases} \quad (4)$$

where $\lfloor\ \rfloor$ indicates the floor function, c_i and c_i' are the i^{th} coefficient or sample point of the original audio and

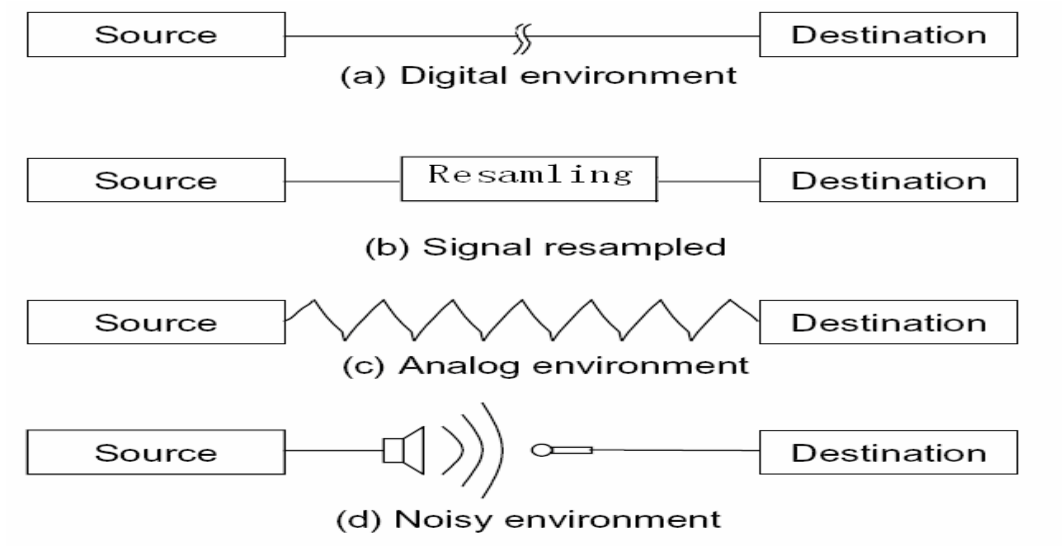

Figure 1: Transmission environments of digital audio [15]

the watermarked audio, respectively. The relationship between *mod* and *floor* functions is

$$Mod(c_i, S) = \begin{cases} c_i - \lfloor (c_i/S) \rfloor \cdot S & \text{if } (c_i \geq 0) \\ c_i + (\lfloor (-c_i/S) \rfloor + 1) \cdot S & \text{if } (c_i < 0). \end{cases}$$

2.2 Transmission Environment Analysis

The digital audio can be transmitted in various environments in practical applications. Some possible scenarios are described in [2, 15], as shown in Figure 1.

The first signal is transmitted through the environment in such a way that is unmodified, shown in Figure 1(a). As a result, the phase and the amplitude are unchanged. In Figure 1(b), the signal is re-sampled with a higher or lower sampling rate. The amplitude and the phase are left unchanged, but the temporal characteristics are changed. The third case, in Figure 1(c), is to convert the signal and transmit it in the analog form. In this case, even if the analog line is considered clear, the amplitude, the phase and the sampling rate are changed. The last case (see Figure 1(d)) is when the environment is not clear, the signal being subjected to nonlinear transformations, resulting in phase changes, amplitude changes, echoes, etc.

In the term of signal processing, watermark is weak signal embedded into a strong background like the digital audio, so the variety of carriers will influence the watermark directly. Therefore, audio watermarking may suffer from the attacks similar to the cover signal. In Figure 1(a), audio watermark is not infected; In Figure 1(b), re-sampling attacked the audio watermarking, which had been settled by many algorithms; Even it is considered no noise corruption in Figure 1(c), audio watermarking still suffer from the effects of DA/AD; Figure 1(d) showed the worst environment, various interferences attacked simultaneity.

3 Investigation of the DA/AD

In [2, 15, 20], the influences caused by the DA/AD on audio are mentioned, but no one of them has pointed out the property of the degradations caused by the DA/AD. Specially, there are no any technical or experimental descriptions to indicate the effects of the DA/AD. Accordingly, it is very necessary to further investigate these effects in order to better represent the DA/AD attack.

3.1 Test Scenario

Due to its degradation on audio watermarking, the DA/AD is taken as an important challenge in [20]. In order to investigate the effects of the DA/AD, we design the test model shown in Figure 2 to simulate the transmission environment in Figure 1(c).

Four 16-bit signed mono audio file in WAVE format denoted by *march.wav*, *drum.wav*, *flute.wav* and *dialog.wav* respectively are used for test, as shown in Table 1. *dialog.wav* is a daily dialog while others are music which involve the different frequency properties.

3.2 The Watermarked Audio in DA/AD

During DA/AD, digital audio signal will suffer from wave distortion (due to the modification of signal energy followed by noise corruption) and the phase change (appeared as temporal scaling).

3.2.1 Wave Magnitude Distortion

Based on many experiments, we observe that the audio amplitudes are modified during the DA/AD, and the modifications rely on the volume played back, the performance of soundcard.

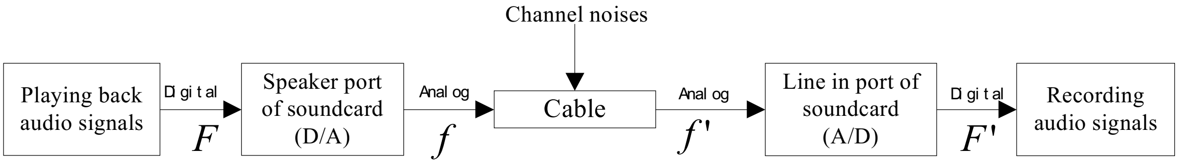

Figure 2: Test Model for DA/AD

Table 1: To-be tested audio files

File name	Sampling Rate	Length	Properties
march.wav	8 and 44.1 (kHz)	56 (Sec)	Composition of both low-frequency and high-frequency
drum.wav	8 and 44.1 (kHz)	56 (Sec)	Mainly composed of low- frequency
flute.wav	8 and 44.1 (kHz)	56 (Sec)	Well-proportioned frequency distribution
dialog.wav	8 and 44.1 (kHz)	56 (Sec)	Daily dialog

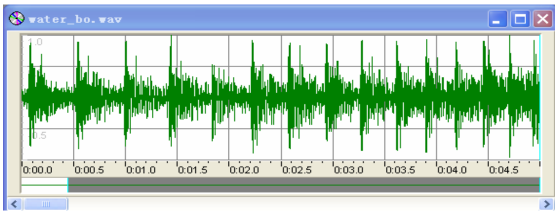

Figure 3: The original audio

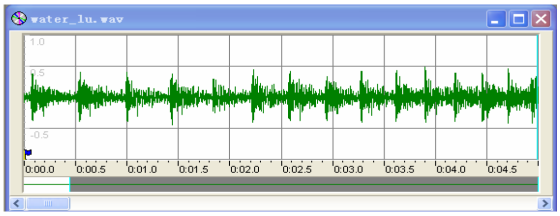

Figure 4: The recorded audio

Figures 3 and 4 show waves of the original audio and the corresponding recorded audio by using the soundcard *Sound Blaster Live5.1*. Compared to the original audio, the amplitude of the recorded audio is obviously modified. The modification of the amplitude varies with different soundcards. Of course, noise during the DA/AD, including quantization noise, will also introduce some distortion.

In the test model in Figure 2, five different soundcards are used to test the audio files in Table 1. Where *Sound Blaster Live5.1* is a civil sound blaster and *TERRATEC AUREON Xfire 1723* is the professional one. *ICON StudioPro7.1* is used for professional recording. *Audio PCI* and *Nightingale Pro 6* are common PC sound blaster. The results are shown in Tables 2 and 3.

3.2.2 Temporal Scaling

From Tables 2 and 3 we have the following observations during the DA/AD:

1) During the DA/AD conversions at a fixed sampling rate, linear temporal scaling occurs and the scaling factor is varied with the different soundcards. Taking audio files sampled at 8 kHz as an example, the soundcard *Studio Pro 7.1* reduces 70 samples while *Audio 2000 PCI* increases 102 samples every 10s. We also test some other kinds of audio like pop music etc., with the length of more than one hour. The experimental results show that such temporal scaling occurs repeatedly during the DA /AD.

2) The number of modified samples is related to the sampling rate of audio signals even the soundcard is unchanged. For example, the samples increase 1079 and 14 every 10s at 8 kHz and 44.1 kHz, respectively, with the same soundcard *Nightingale Pro 6*. As for other common sampling rates of 11.025 kHz, 16 kHz, 22.05 kHz, 32 kHz, 48 kHz, this phenomenon still exists but the modified samples are a little different.

The analysis above indicates that the modification of phrase described in [2, 15, 20] may be represented as temporal scaling. The scaling factor is only related to the sampling rate and the performance of the soundcard.

Consequently, the main degradation caused by the DA/AD on audio watermarking may be concluded as follows.

1) The modification of wave magnitude. Distortions are brought to the audio watermarking by energy change followed additional noise.

2) Temporal scaling. This will lead to the synchronization problem in detection.

In Section 4, we will investigate their degradations on quantization-based audio watermarking.

Table 2: The number of increased or reduced samples caused by DA/AD in sampling rate of 8 kHz

Soundcard / Time	***Sound Blaster Live5.1***	***AUREON Xfire 1723***	***Audio 2000 PCI***	***Nightingale Pro 6***	***Studio Pro 7.1***
10s	Reduce: 1	Increase: 1000	Increase: 102	Increase: 1079	Reduce: 70
20s	Reduce: 2	Increase: 2000	Increase: 204	Increase: 2158	Reduce: 140
30s	Reduce: 3	Increase: 3000	Increase: 306	Increase: 3237	Reduce: 210
40s	Reduce: 4	Increase: 4000	Increase: 408	Increase: 4316	Reduce: 280
50s	Reduce: 5	Increase: 5000	Increase: 510	Increase: 5395	Reduce: 350

Table 3: The number of increased or reduced samples caused by DA/AD in sampling rate of 44.1 kHz

Soundcard / Time	***Sound Blaster Live5.1***	***AUREON Xfire 1723***	***Audio 2000 PCI***	***Nightingale Pro 6***	***Studio Pro 7.1***
10s	Reduce: 6	Increase: 0	Increase: 0	Increase: 14	Reduce: 0
20s	Reduce: 12	Increase: 0	Increase: 0	Increase: 28	Reduce: 0
30s	Reduce: 18	Increase: 0	Increase: 0	Increase: 42	Reduce: 0
40s	Reduce: 24	Increase: 0	Increase: 0	Increase: 56	Reduce: 0
50s	Reduce: 30	Increase: 0	Increase: 0	Increase: 70	Reduce: 0

4 Performance Analysis

From the discussions above, we know that the watermark suffers from noise, amplitude changes and temporal scaling during the DA/AD. In order to study the performance of quantization-based audio watermarking in the DA/AD, the possible distortions against Gaussian noise, amplitude change and temporal scaling attacks are discussed in this section.

4.1 Gaussian Noise Corruption

Gaussian noise is possibly introduced during transmission. Let $F' = \{f'_i | i = 1, 2, \ldots, N\}$ denote the watermarked audio, then the corrupted audio can be expressed as $\{f^*_i\} = \{f'_i\} + \{n_i\}$, where $\{n_i\}$ is Gaussian noise obeying $N(0, \sigma^2)$. The version of n_i after the unitary transformation may be expressed as:

$$d_i = c^*_i - c'_i$$

where $\{c^*_i\}$ and $\{c'_i\}$ are the transform coefficients of $\{f'_i\}$ and $\{f^*_i\}$ respectively.

By Equations (3) and (4), we know that if $d_i = (c^*_i - c'_i) \notin (kS - S/4, kS + S/4)$, then the watermark can not be extracted correctly. According to the theory of unitary transforms, $\{d_i\}$ still obey Gaussian distribution $N(0, \delta^2)$ if $d_i \in R$. Consequently, the BER in case of Gaussian nose can be computed as:

$$\begin{aligned} BER &= 1 - \sum_{k \in Z} \int_{kS-S/4}^{kS+S/4} \frac{1}{\sqrt{2\pi}\delta} \cdot e^{\frac{-(x-a)^2}{2\delta^2}} \cdot dx \\ &\approx \int_{|x|>S/4} \frac{1}{\sqrt{2\pi}\delta} \cdot e^{\frac{-(x-a)^2}{2\delta^2}} \cdot dx \\ &= \frac{2}{\sqrt{2\pi}\delta} \cdot \int_{S/4}^{+\infty} e^{-\frac{x^2}{2\delta^2}} \cdot dx \qquad (5) \end{aligned}$$

Equation (5) indicates that the robustness of watermark is mostly determined by the quantization step S under the same power of noise corruption and is independent of the original audio.

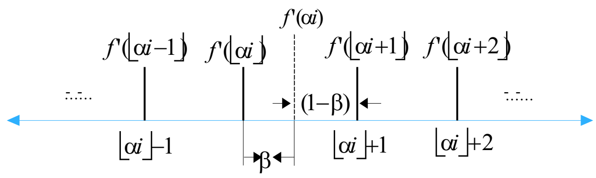

Figure 5: The scheme of the temporal scaling

4.2 Amplitude Modification Attack

Assume that the amplitude of the audio sample $\{f_i | i = 1, 2, \ldots, N\}$ is changed to $\{|\beta_i \cdot f_i\}$ after the DA/AD, where β_i is the amplitude scaling factor of the i^{th} sample. In terms of unitary transformation, the amplitude of coefficients $Y = \{y_i | i = 1, 2, \ldots, N\}$ will be changed to $\{|\lambda_i \cdot y_i\}$, where λ_i is the corresponding scaling factor of the i^{th} coefficient y_i.

Let $C' = \{c'_j | j = 1, 2, \ldots, M\}$ denote the transform coefficients of the watermarked audio, so the amplitude modification on the coefficients is calculated as $d_i = (\lambda_i -$

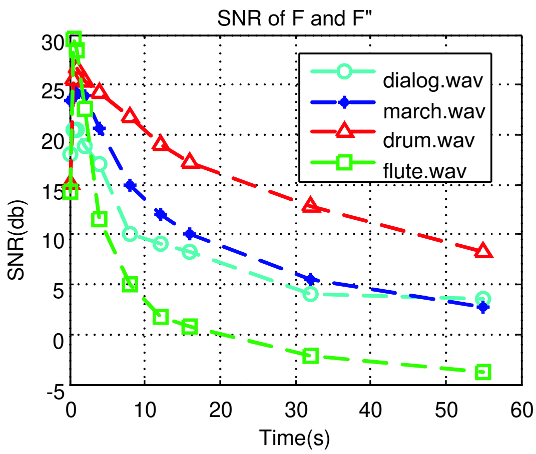

Figure 6: SNR curves between F and F''

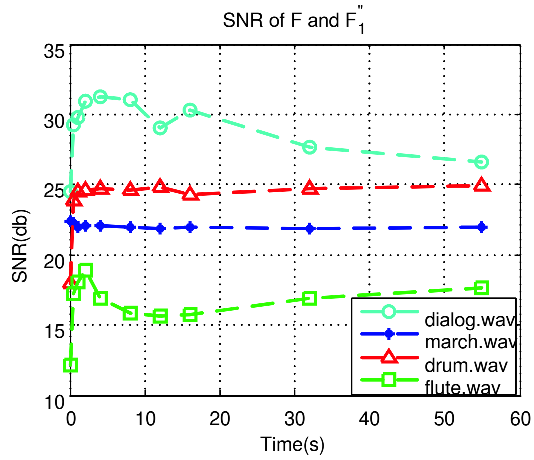

Figure 7: SNR curves between F and F''_1

$1) \cdot c'_i$. In case of $d_i > kS + S/4$ ($k \in Z$), the watermark bit can not be extracted exactly. In other words, if Equation (6) is true, errors will occur:

$$(\lambda_i - 1) \cdot c'_i > kS + \frac{S}{4} \quad (6)$$

where $k = \lfloor (\lambda_i - 1) \cdot c'_i / S + 0.25 \rfloor$. c'_i is the i^{th} transform coefficient. If the amplitude modification is linear, then $\beta = \beta_i$, From the property of unitary transformation, we have $\beta = \lambda_i$. When $\beta \neq 1$, we rewrite Equation (6) as:

$$(\beta - 1) \cdot c'_i > kS + \frac{S}{4}.$$

Furthermore, we have

$$\begin{cases} c'_i > \frac{S}{(\beta-1)}(k + 0.25) & \text{if } \beta > 1 \\ c'_i < \frac{S}{(\beta-1)}(k + 0.25) & \text{if } 0 < \beta < 1 \end{cases} \quad (7)$$

where $k = \lfloor (\beta - 1) \cdot c'_i / S + 0.25 \rfloor$. β is amplitude change factor. Note that the modification of the amplitude become smaller, namely the value of β goes to 1, so $(\beta - 1) \cdot c'_i$ go to 0. The smaller number of the coefficients satisfying Equation (7) indicates a lower BER. In the case of $\beta = 1$, no attack occurs and BER is zero according to Equation (6). Therefore, the BER of the watermark against amplitude modification attack can be estimated as:

$$BER = R/M$$

where R denotes the number of coefficients satisfying Equations (6) or (7) and M is the number of the coefficients selected for embedding watermark bits.

By Equations (6) and (7), the mainly influences caused by amplitude modification attack on the BER of quantization-based audio watermarking are as follows,

1) The embedding strength. The larger of the size, the stronger ability of the watermark to resist the attack.

2) The extent of the audio amplitude modification.

3) The type of the audio. For the same amplitude attack, $(\lambda_i - 1) \cdot c'_i$ keep unchanged, thus BER does not rely on the audio type. As for amplitude scaling attack, $(\lambda_i - 1) = (\beta - 1)$ a constant, the stronger of the signal energy c'_i, the larger $(\beta - 1) \cdot c'_i$, then the weaker ability to resist amplitude scaling attack.

4.3 Temporal Scaling Attack

Let $F = \{f(i)\}$, $F' = \{f'(j)\}$ denote the watermarked audio and the audio performed the DA/AD respectively. L' and L are the number of samples in the and F. Temporal scaling occurs with a scaling factor of α when $L' \neq L$, where $\alpha = (L' - 1)/(L - 1)$. Based on the property of continuous function, if $f(t)$, $0 \leq t < L$, is the analog version of F, it can be denoted as $f(\alpha t)$ after temporal scaling with the factor α. This means the position of sample $f(i)$ is shift to $f(\alpha i)$. In case of α is not an integer, the sample $f'(\alpha i)$ cannot be directly used in F' because each sample point must be integer (refer to Figure 5). Normally, the nearest sample points in F' are utilized to approximately evaluate $f'(\alpha i)$.

Since the positions of the samples in time domain are modified by the temporal scaling, the transform coefficients c'_i are changed accordingly. From Equation (4), we know that errors may occur if we extract the watermark directly. Once the amount of the coefficient c'_i modified is over $(kS + S/4)$, the corresponding watermark bit w'_i cannot be extracted correctly.

We use SNR to measure the effects of temporal scaling on the watermarked audio.

$$\begin{aligned} SNR &= -10 \log_{10}\left(\frac{\| F - F'' \|^2}{\| F \|^2}\right) \\ F'' &= F' / \left(\sum_{i=0}^{N-1} |f'_i| / \sum_{i=0}^{N-1} |f_i|\right) \end{aligned} \quad (8)$$

where $N = \min(L, L')$, $F'' = \{f''_i\}$ denotes the normalized version of F' by Equation (8), while f_i, f'_i and f''_i

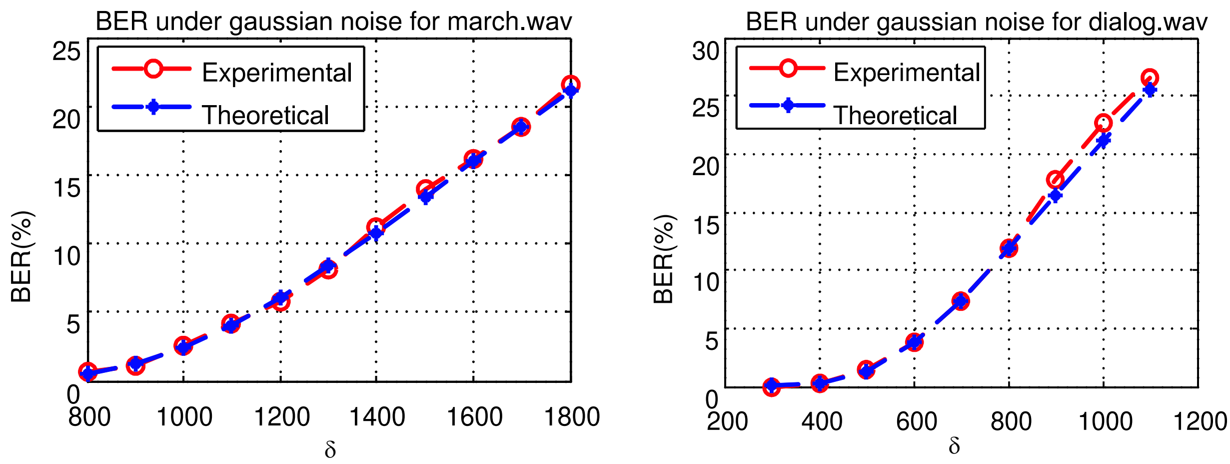

Figure 8: *BER* against Gaussian noise $n(0, \delta^2)$

denote the amplitude of the i^{th} sample in F, F' and F'' respectively. Since the amplitude may be modified after the DA/AD, F' should be normalized before calculating *SNR*s.

Taking *Sound Blaster Live5.1* soundcard as an example, we perform the DA/AD conversions on the audio signal sampled at 8 kHz, and calculate *SNR*s of F verse F''. The results are shown in Figure 6. F_1'' is the version that F'' is performed linear interpolation processing. The *SNR*s curves of F verse F_1'' are calculated as Figure 7.

Figure 6 shows that *SNR*s of F verse F'' decrease quickly. It means that temporal scaling changes the locations of audio samples, and thus results in serious desynchronization distortion.

Figure 7 shows that the *SNR*s of some different audio signals F'' with respect to F keep stabilization, indicating that interpolation processing can effectively alleviate the degradation caused by the temporal scaling in the DA/AD.

It is noted that the quantization-based audio watermarking is susceptive to temporal scaling caused by the DA/AD, even minor scaling may cause large distortion. As for the other soundcards in Table 2 or other sampling frequency, the same conclusion may be drawn.

5 Experimental Results

In order to demonstrate the validity of results shown in Sections 4.1 and 4.2, we test the audios in Table 1 sampled at 8 kHz denoted as *march.wav* and *dialog.wav*, with the embedding strength of 9000 and 5000 respectively. The watermarks are embedded into low frequency sub-band of DWT (with the wavelet base db1) coefficients. Figures 8 and 9 show the test results. The simulation results for other kinds of audios and other sampling frequencies like sampled at 44.1 kHz are similar.

Figure 8 shows that the experimental and theoretical results of *BER* after Gaussian noise corruption agree with one another very well, which indicated the correctness of Equation (5).

Plots in Figure 9 show the *BER* under the amplitude scaling attack. The excellent agreement between experimental and theoretical results confirms the validity of Equation (7).

6 Conclusions

Compared with some given attacks, the DA/AD presents more challenges on audio watermarking due to it involving energy modification and temporal scaling simultaneously. The main contributions of this paper are as follows,

1) Investigated the possible degradation of audio watermarking caused by the DA/AD, and modeled the main degradations as noise corruption, temporal scaling, and amplitude modification.

2) Analyzed the performance of quantization-based audio watermarking against audio amplitude modification and deducted the expression of *BER*. The experiment results demonstrate its validity.

3) Pointed out that the phase change of audio caused by the DA/AD may be represented as the temporal scaling based on the extensive experiments. The scaling factor relies on the soundcard performance and the sampling rate of audio files. Furthermore, we analyzed the effects of temporal scaling on audio watermarking.

4) Both the theoretical and experimental results show that quantization-based audio watermarking is susceptible to amplitude modification and temporal scaling, even minor scaling occurring in time domain or wave magnitude will cause large watermark distortion. These conclusions mean that quantization

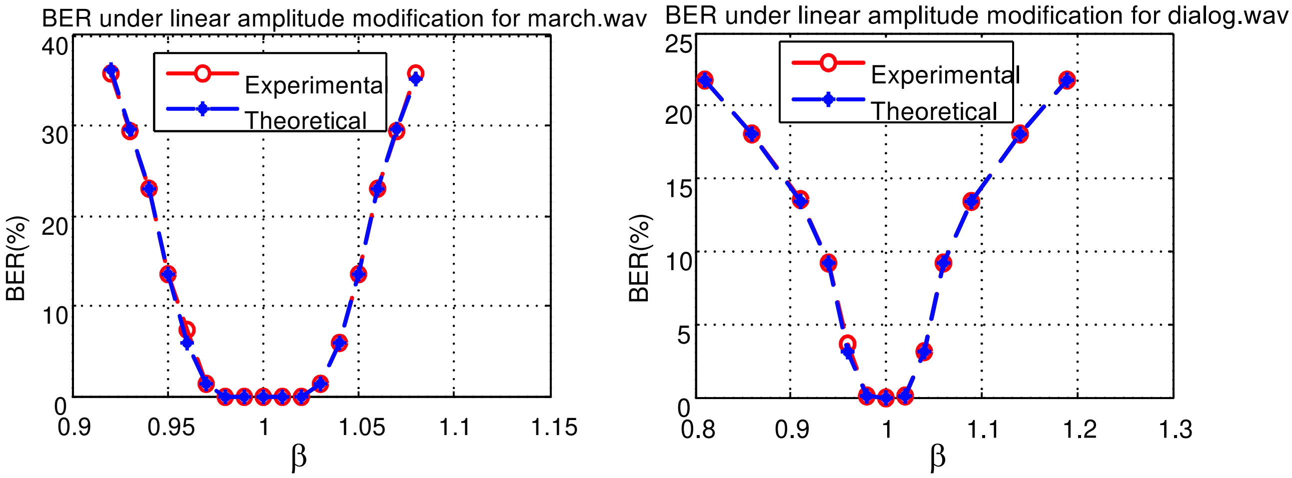

Figure 9: *BER* against linear amplitude scaling (β)

watermark embedding is sensitive to the DA/AD and other effective watermark embedding should be addressed for the audio watermark to resist the DA/AD.

Acknowledgements

We thank the support of NSFC (60325208, 60133020), NSF of Guangdong (04205407).

References

[1] P. Bassia, I. Pitas, N. Nikolaidis N, "Robust audio watermarking in the time domain," *IEEE Transactions on Multimedia*, vol. 3, pp. 232-241, 2001.

[2] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for data hiding," *IBM Systems Journal*, vol. 35, pp. 313-336, 1996.

[3] 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, pp. 1423-1443, 2001.

[4] S. Chen and H. Leung, "Concurrent data transmission on analog telephone channel by data hiding technique," in *Proceedings of 2004 IEEE International Symposium on Consumer Electronics*, pp. 295-298, 2004.

[5] I. J. Cox, J. Kilian, T. Leighton and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, pp. 1673-1687, 1997.

[6] I. J. Cox, M. L. Miller, and J. A. Bloom, *Digital Watermarking*, Academic Press, 1999.

[7] J. Haitsma, M. V. D. Veen, T. Kalker, and F. Bruekers, "Audio watermarking for monitoring and copy protection," in *Proceeding of ACM Multimedia Workshops*, pp. 119-122, 2000.

[8] J. Huang, Y. Wang, and Y. Q. Shi, "A blind audio watermarking algorithm with self-synchronization," in *Proceeding of IEEE International Symposium on Circuits and Systems*, Arizona, USA, vol. 3, pp. 627-630, 2002.

[9] W. N. Lie and L. C. Chang, "Robust and high-quality time-domain audio watermarking subject to psychoacoustic masking," in *Proceeding of IEEE International Symposium on Circuits and Systems*, vol. 2, pp. 45-48, 2002.

[10] S. Katzenbeisser and F. A. P. Petitcolas, Eds., *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House, Inc., 2000.

[11] S. Kim and K. Bae, "Robust estimation of amplitude modification for scalar costa scheme based audio watermark detection," in *Proceeding of International Workshop on Digital Watermarking*, Seoul, pp. 121-133, 2004.

[12] M. Mansour and A. Tewfik, "Audio watermarking by time-scale modification," in *Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, pp. 1353-1356, 2001.

[13] T. Nakamura, R. Tachibana, and S. Kobayashi, "Automatic music monitoring and boundary detection for broadcast using audio watermarking," in *Proceeding SPIE*, vol. 4675, pp. 170-180, 2002.

[14] C. I. Podilchuk and E. J. Delp, "Digital watermarking: Algorithms and applications," *IEEE Signal Processing Magazine*, vol. 18, no. 4, pp. 33-46, July 2001.

[15] Richard Popa, *An Analysis of Steganographic Techniques*, Ph. D Thesis, pp. 26-27, 1998.

[16] J. Seok, J. Hong, and J. Kim, "A novel audio watermarking algorithm for copyright protection of digital audio," *ETRI Journal*, vol. 24, no. 3, pp. 181-189, 2002.

[17] R. K. Sharma and S. Decker, "Practical challenges for digital watermarking applications," *Eurasip Journal on Applied Signal Processing*, vol. 2002, no. 2, pp. 133-139, Feb. 2002.

[18] S. Shin, O. Kim, J. Kim, and J. Choil Jonguk, "A robust audio watermarking algorithm using pitch scaling," in *Proceeding of IEEE Workshop on Digital Signal Processing*, vol. 2, pp. 701-704, 2002.

[19] M. Steinebach, F. A. P. Petitcolas, F. Raynal, J. Dittmann, C. Fontaine, S. Seibel, N. Fates, and L. C. Ferri, "StirMark benchmark: audio watermarking attacks," in em Proceeding of International Conference on Information Technology: Coding and Computing (ITCC'01), pp. 49-54, 2001.

[20] M. Steinebach, A. Lang, J. Dittmann, and C. Neubauer, "Audio watermarking quality evaluation: Robustness to DA/AD processes," in *Proceeding of International Conference on Information Technology: Coding and Computing*, pp. 100-103, 2002.

[21] R. Tachibana, "Audio watermarking for live performance," in *Proceeding SPIE*, vol. 5020, pp. 32-43, 2003.

[22] M. J. Tsai, K. Y. Yu, and Y. Z. Chen, "Joint wavelet and spatial transformation for digital watermarking," *IEEE Transactions on Consumer Electronics*, vol. 46, no. 1, pp. 241-245, 2000.

[23] http://www.petitcolas.net/fabien/watermarking/stirmark/.

[24] S. Wu, J. Huang, D. Huang, and Y. Q. Shi, "Efficiently self-synchronized audio watermarking for assured audio data transmission," *IEEE Transactions on Broadcasting*, vol. 51, pp. 69-76, 2005.

Shijun Xiang received the B.E. degree from Chang'an University, China, in 1997, and M.S. degree from Guizhou University, China, in 2000, and is currently working toward the Ph.D. degree in the School of Information Science and Technology, Sun Yat-Sen University, Guangzhou, China. His research interests include digital audio watermarking, data hiding.

Jiwu Huang received his B.S. degree from Xidian University, China, in 1982, M.S. degree from Tsinghua University, China, in 1987, and Ph.D. degree from Institute of Automation, Chinese Academy of Science in 1998. He is currently a professor with the School of Information Science and Technology, Sun Yat-Sen University, P.R. China. He serves as a member of IEEE CASS Technical Committee of Multimedia Systems and Applications. His current research interests include multimedia security and data hiding.