Audio Watermarking Based on Quantization in Wavelet Domain

Vivekananda Bhat K., Indranil Sengupta, and Abhijit Das

Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur 721 302, India
{vbk,isg,abhij}@cse.iitkgp.ernet.in

Abstract. A robust and oblivious audio watermarking based on quantization of wavelet coefficients is proposed in this paper. Watermark data is embedded by quantizing large wavelet coefficient in absolute value of high frequency detail sub-band at the third level wavelet transform. The watermark can be extracted without using the host signal. Experimental results show that the proposed method has good imperceptibility and robustness under common signal processing attacks such as additive noise, low-pass filtering, re-sampling, re-quantization, and MP3 compression. Moreover it is also robust against desynchronization attacks such as random cropping and jittering. Performance of the proposed scheme is better than audio watermarking scheme based on mean quantization.

Keywords: Audio watermarking, quantization, discrete wavelet transform (DWT).

1 Introduction

With the rapid development of the Internet and multimedia technologies in the last decade, the copyright protection of digital media is becoming increasingly important and challenging. Digital watermarking [1] has been proposed as a tool to a number of media security problems. The purpose of audio watermarking is to supply some additional information about the audio by hiding watermark data into it. This watermark data may be used for various applications such as authentication, copyright protection, proof of ownership, *etc.* Research in the area of digital watermarking has focused primarily on the design of robust techniques for Digital Rights Management (DRM).

Audio watermarking technique should exhibit some desirable properties [1]. Imperceptibility and robustness are two fundamental properties of audio watermarking schemes. The watermark should not be visible to the viewer and the watermarking process should not introduce any perceptible artifacts into the original audio signal. In other words, watermark data should be embedded imperceptibly into digital audio media. Also, the watermark should survive after various intentional and unintentional attacks. These attacks may include additive noise, re-sampling, low-pass filtering, re-quantization, MP3 compression, cropping, jittering and any other attacks that remove the watermark or confuse

R. Sekar and A.K. Pujari (Eds.): ICISS 2008, LNCS 5352, pp. 235–242, 2008.
© Springer-Verlag Berlin Heidelberg 2008

watermark reading system [2]. A trade-off should be maintained between these two conflicting properties.

In general, the audio watermarking techniques can be classified into different ways. The easier watermarking techniques were almost time domain approaches. The simplest method is based on embedding the watermark in the least significant bits (LSB's) of audio samples. Probably the most famous time domain technique proposed in [3] is based on human auditory system (HAS). However, time domain techniques are not resistant enough to MP3 compression and other signal processing attacks. For example, a simple low-pass filtering may eliminate the watermark. Transform domain watermarking [1] schemes are those based on the fast Fourier transform (FFT), and DWT, typically provide higher audio fidelity and are much robust to audio manipulations. A very few quantization based audio watermarking schemes have been proposed in the literature. Wang *et al.* [4] proposed blind audio watermarking technique based on mean quantization. Low frequency wavelet coefficients are selected to embed watermark data. From the experimental results it follows that this method is robust to common signal processing attacks. But the robustness against desynchronization attacks is not discussed in this scheme. Kalantari *et al.* [5] presented a oblivious scheme based on mean quantization in wavelet domain. The watermark data is embedded by quantizing the means of two selected wavelet sub-band. The robustness of this scheme is not discussed against the desynchronization attacks. A novel quantization based audio watermarking algorithm using discrete cosine transform is discussed in [6]. Experimental results shows that, this method has compromised audibility and robustness in better manner. Ronghui *et al.* [7] gave an semi-fragile quantization based audio watermarking scheme in wavelet domain. The watermark is embedded by changing the value of selected coefficients using quantization. This scheme has limited robustness against common signal processing. A fragile audio watermarking using adaptive wavelet packets based on quantization is reported in [8]. We present an audio watermarking algorithm in the wavelet domain based on quantization. The motivation of our work is based on the idea of image watermarking technique proposed in [9]. The important features of the proposed algorithm are: (i) The binary watermark is encrypted using Arnold transform. (ii) Watermark is embedded using quantization of wavelet coefficients. (iii) Watermark extraction is blind without using original audio signal. (iv) The proposed algorithm has better robustness compared to the scheme [4] against common signal processing and desynchronization attacks. The rest of this paper is organized as follows: A detail overview of the proposed scheme are given in section 2. The experimental results of this scheme are presented and discussed in section 3. Finally, section 4 concludes the paper.

2 Quantization Based Watermarking Scheme

The basic idea in the DWT for a one dimensional signal is the following. A signal is split into two parts, usually high frequency detail sub-band and low frequency approximate sub-band using wavelet filter. The low frequency part is split again

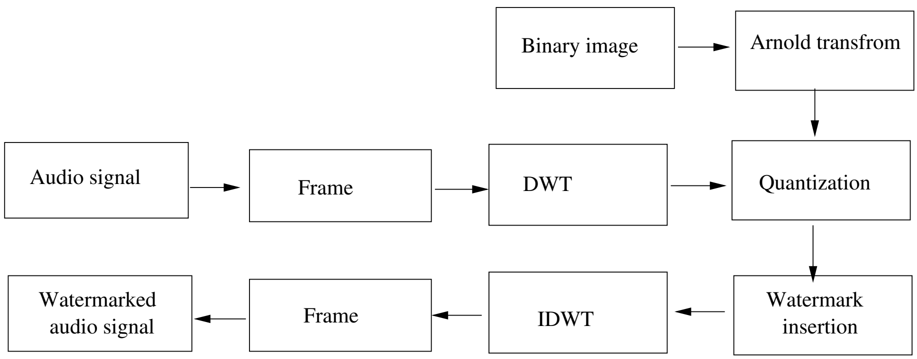


Fig. 1. Watermark embedding scheme

into two parts of high and low frequencies. This process is repeated finite number of times. The original signal is restored back similarly using inverse DWT (IDWT). The Arnold transform [10] is used for watermark permutation to ensure security and robustness of the proposed scheme. The watermark is embedded by quantizing the largest wavelet coefficient in absolute value of the high frequency detail sub-band. The detail steps involved in the watermark embedding and extraction are outlined below.

2.1 Embedding Scheme

The block diagram of the watermark embedding algorithm is shown in Fig. 1. The watermark embedding process involves the following steps:

Step 1: The watermark data P is a binary image of size $N \times N$ and is given by $P = \{p(n_1, n_2) : 1 \leq n_1 \leq N, 1 \leq n_2 \leq N, p(n_1, n_2) \in \{0, 1\}\}$.

Step 2: The watermark is pre-processed before embedding into the host signal. All the elements P will be scrambled by applying Arnold transform. Due to the periodicity of the Arnold transform, the watermark can be recovered easily after transformation. Let $(n_1, n_2)^T$ is the coordinate of the watermark image's pixel, then $(n_1^{'}, n_2^{'})^T$ is the coordinate after the transform. Arnold transform can be expressed as

$$\begin{bmatrix} n_1^{'} \\ n_2^{'} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \quad (\mathrm{mod}\ N) \tag{1}$$

We should convert the two-dimensional scrambled watermark into the one-dimensional sequence in order to embed it into the audio signal. The corresponding one-dimensional sequence is given by: $W = \{w(k) = p(n_1^{'}, n_2^{'}) : 1 \leq n_1^{'} \leq N, 1 \leq n_2^{'} \leq N, k = (n_1^{'} - 1) \times N + n_2^{'}, 1 \leq k \leq N \times N\}$.

Step 3: The original audio is first divided into non-overlapping frames, with a frame size of 2048 samples. The number of frames is equal to the size of the watermark data.

Step 4: The audio frame is decomposed into 3-level wavelet transform using 4-coefficient Daubechies wavelet (db4) filter. The wavelet coefficients after decomposition are given by $cA3, cD3, cD2$, and $cD1$, where $cA3$ is the low frequency approximate coefficients of the audio and $cD3, cD2, cD1$ are the high frequency detail coefficients of the audio.

Step 5: Select the largest coefficient $cD3_{max}(m)$ in absolute value of $cD3$ to embed watermark using quantization.

Step 6: Quantize the largest coefficient as follows. The following quantization function Q is used during the embedding and extraction process.

$$Q(cD3_{max}(m)) = \begin{cases} 0 \text{ if } & \lfloor \frac{cD3_{max}(m)}{\Delta\ 2^3} \rfloor \text{ is even} \\ 1 \text{ if } & \lfloor \frac{cD3_{max}(m)}{\Delta\ 2^3} \rfloor \text{ is odd} \end{cases} \tag{2}$$

where Δ is a user defined positive real number called quantization parameter and $\lfloor\ \rfloor$ is the floor integer function.

If $Q(cD3_{max}(m)) = w(k)$, then no change is made to the coefficient. If $Q(cD3_{max}(m)) \neq w(k)$ then

$$cD3_{max}(m) = \begin{cases} cD3_{max}(m) - \Delta\ 2^3 \text{ if } & cD3_{max}(m) > 0 \\ cD3_{max}(m) + \Delta\ 2^3 \text{ if } & cD3_{max}(m) \leq 0 \end{cases} \tag{3}$$

Step 7: Inverse DWT is applied to the modified wavelet coefficients to get the watermarked audio signal.

2.2 Extraction Scheme

The watermark can be extracted without using original audio signal. The extraction algorithm is given below:

Step 1: The input watermarked audio signal is segmented into non-overlapping frames, with a frame size of 2048 samples.

Step 2: The frame is transformed into 3-level wavelet transform using db4 filter.

Step 3: Select the largest coefficient $cD3'_{max}(m)$ in absolute value of 3^{rd} level high frequency detail coefficient $cD3'$.

Step 4: The extracted watermark is given by, $w'(k) = Q(cD3'_{max}(m))$.

Step 5: The extracted watermark is de-scrambled using inverse Arnold transform to obtain the original binary watermark image.

3 Experimental Results and Comparison

Audio files used in the experiment are 16 bit mono audio signal in WAVE format sampled at 44100 Hz sampling rate. A plot of a short portion of the jazz audio signal and its watermarked version is shown in Fig. 2. The embedded watermark is a Crown binary logo image of size 36×36 shown in Fig. 3. The binary

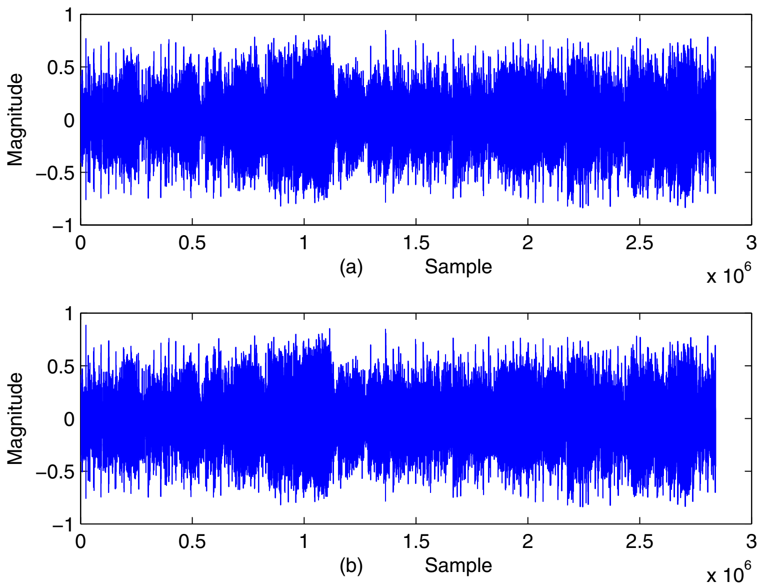


Fig. 2. (a) Jazz audio signal (b) Watermarked jazz audio signal

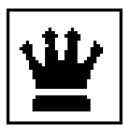


Fig. 3. Binary watermark

watermark is encrypted using Arnold transform. The signal to noise ratio (SNR) for evaluating the quality of watermarked signal is given by the equation:

$$SNR = 10\log_{10}\frac{\sum_{a=1}^{M} Z^2(a)}{\sum_{a=1}^{M}[Z(a) - Z^*(a)]^2} \tag{4}$$

where Z and Z^* are original audio signal and watermarked audio signal respectively.

The SNR of all selected audio files are above 20 dB. This ensures fact that watermarked audio signal is quite similar to original audio signal. Which means our watermarking method has good imperceptibility. We know that there is a trade-off between imperceptibility of the watermark with different value of the quantization parameter Δ. A large value of the parameter makes the watermark robust, but it will destroy quality of original signal. A small value of the parameter allows us to achieve good imperceptibility, but it is fragile to attacks. The following attacks were performed to test the robustness and effectiveness of our scheme.

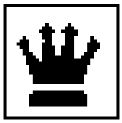
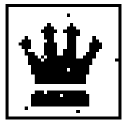
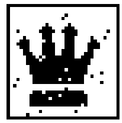
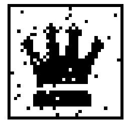
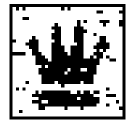
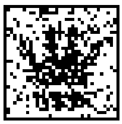
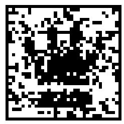
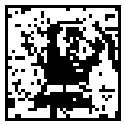
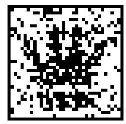
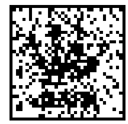
Watermark					
NC	1	0.9932	0.9847	0.9729	0.9050
Watermark					
NC	0.8143	0.8116	0.8062	0.7899	0.7831

Fig. 4. Extracted watermark with various NC

(i) Additive white Gaussian noise (AWGN): White Gaussian noise is added so that the resulting signal has a SNR of 30 dB.
(ii) Re-sampling: The watermarked signal originally sampled at 44.1 kHz is re-sampled at 22.05 kHz, and then restored by sampling again at 44.1 kHz.
(iii) Low-pass filtering: The low-pass filter used here is a second order Butterworth filter with cut-off frequency 11025 Hz.
(iv) Re-quantization: The 16-bit watermarked audio signals have been re-quantized down to 8 bits/sample and back to 16 bits/sample.
(v) MP3 Compression: The MPEG-1 layer 3 compression with 64 kbps is applied.
(vi) Cropping: Segments of 500 samples were randomly removed and replaced with segments of the signal attacked with filtering and additive noise.
(vii) Jittering: Jittering is an evenly performed form of random cropping. We removed one sample out of every 2000 samples in our jittering experiment.

The similarity measurement between extracted watermark and original watermark can be evaluated using normalized correlation (NC), which is defined as follows:

$$NC(Y, Y^*) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} Y(i,j)Y^*(i,j)}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} Y(i,j)^2}\ \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} Y^*(i,j)^2}} \tag{5}$$

where Y and Y^* are original and extracted watermarks respectively, i and j are indexes of the binary watermark image.

The bit error rate (BER) is used to find the percentage of the extracted watermark, it is defined by:

$$BER = \frac{B}{N \times N} \times 100\% \tag{6}$$

where B is the number of erroneously detected bits.

Table 1. NC values and BER of extracted watermark from different types of attacks

Audio file	Type of attack	Proposed scheme NC	Scheme [4] NC	Proposed scheme BER (%)	Scheme [4] BER(%)
Classical	Attack free	1	1	0	0
	AWGN	1	0.9729	0	4
	Re-sampling	1	0.9958	0	1
	Low-pass filtering	1	0.9932	0	1
	Re-quantization	1	0.9847	0	2
	MP3 64 kbps	0.9050	0.8909	13	14
	Cropping	1	0.9889	0	2
	Jittering	1	0.9905	0	1
Country	Attack free	1	1	0	0
	AWGN	1	0.9916	0	1
	Re-sampling	1	0.9905	0	1
	Low-pass filtering	1	0.9858	0	2
	Re-quantization	1	0.9852	0	2
	MP3 64 kbps	0.8116	0.7831	26	29
	Cropping	1	0.9905	0	1
	Jittering	1	0.9968	0	0
Jazz	Attack free	1	1	0	0
	AWGN	1	0.9789	0	3
	Re-sampling	1	1	0	0
	Low-pass filtering	1	0.9958	0	1
	Re-quantization	1	0.9937	0	1
	MP3 64 kbps	0.8143	0.7899	26	28
	Cropping	1	0.9963	0	1
	Jittering	1	0.9958	0	1
Pop	Attack free	1	1	0	0
	AWGN	1	0.9920	0	1
	Re-sampling	1	0.9942	0	1
	Low-pass filtering	1	0.9911	0	1
	Re-quantization	1	0.9764	0	4
	MP3 64 kbps	0.8062	0.6324	26	46
	Cropping	1	0.9905	0	1
	Jittering	1	0.9968	0	0

The Fig. 4. shows the some of the extracted watermarks after the above attacks with different NC values. Performance comparison of our method with the audio watermarking scheme based on the mean quantization [4] for the above attacks is summarized in the Table 1. Moreover our scheme has high NC and low BER as compared to the scheme [4] for the above mentioned attacks, which clearly shows the efficiency of our approach. The proposed scheme has NC above 0.8 for all the above attacks, that ensures extracted watermark is recognizable and acceptable. For example, the NC value of our scheme under MP3 compression attack at 64 kbps for pop audio file is 0.8062, where as the NC value for the scheme [4] is 0.6324 under the same attack. This result shows that our scheme is better than the scheme [4]. Also for the desynchronization attacks such as random cropping and jittering the results of our approach is better than [4]. We observe that the performance is mainly depends on the audio files and watermark logo image used in the experiment. Audio files such as classical, country, jazz, and pop are good host signals for our audio watermarking scheme. These results prove that the proposed method has superior performance.

4 Conclusions

A blind audio watermarking technique resistant to signal processing and desynchronization attacks was presented in this paper. Arnold transform is used to ensure the security and robustness of the watermark data. This technique yields impressive results both in terms of image quality and robustness against various signal processing attacks. Watermarked audio signal maintains more than 20 dB SNR, which assures the imperceptibility of our scheme. From the robustness tests, we can see that the algorithm is robust to common signal processing and desynchronization attacks. Experimental results shows that performance of the proposed scheme is much better than audio watermarking scheme based on mean quantization.

References

1. Cox, I., Miller, M., Bloom, J., Fridrich, J., Kalke, T.: Digital watermarking and steganography, 2nd edn. Morgan Kaufmann, San Francisco (2007)
2. Cvejic, N., Seppanen, T.: Digital audio watermarking techniques and technologies. Information Science Reference, USA (August 2007)
3. Basia, P., Pitas, I., Nikolaidis, N.: Robust audio watermarking in the time domain. IEEE Transactions on Multimedia 3(2), 232–241 (2001)
4. Lanxun, W., Chao, Y., Jiao, P.: An audio watermark embedding algorithm based on mean-quantization in wavelet domain. In: 8th International Conference on Electronic Measurement and Instruments, pp. 2-423–2-425 (2007)
5. Kalantari, N., Ahadi, S., Kashi, A.: A robust audio watermarking scheme using mean quantization in the wavelet transform domain. In: IEEE International Symposium on Signal Processing and Information Technology, pp. 198–201 (2007)
6. Zhou, Z., Zhou, L.: A novel algorithm for robust audio watermarking based on quantification DCT domain. In: Third International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIHMSP 2007), vol. 1, pp. 441–444 (2007)
7. Tu, R., Zhao, J.: A novel semi-fragile audio watermarking scheme. In: Proceedings of the 2nd IEEE Internatioal Workshop on Haptic, Audio and Visual Environments and Their Applications (HAVE 2003), pp. 89–94 (2003)
8. Quan, X., Zhang, H.: Perceptual criterion based fragile audio watermarking using adaptive wavelet packets. In: Proceedings of the 17th International Conference on Pattern Recognition (ICPR 2004), vol. 2, pp. 867–870 (2004)
9. Kundur, D., Hatzinakos, D.: Digital watermarking for telltale tamper proofing and authentication. Proceedings of the IEEE 87(7), 1167–1180 (1999)
10. Voyatzis, G., Pitas, I.: Applications of toral automorphisms in image watermarking. In: Proceedings of International Conference on Image Processing, vol. 1, pp. 237–240 (1996)