International Journal of Security and Its Applications
Vol.7, No.5 (2013), pp.183-192
http://dx.doi.org/10.14257/ijsia.2013.7.5.17

# Audio Watermarking by Coefficient Quantization in the DWT-DCT Dual Domain

De Li[1], Yingying Ji[1] and JongWeon Kim[2](Corresponding author)

[1] *Department of Computer Science, Yanbian University, Yanji, China*
*[leader1223@ybu.edu.cn, ji_yingying@163.com]*
[2] *Department of Copyright Protection, Sangmyung University, Seoul, Korea*
*[jwkim@smu.ac.kr]*

***Abstract***

*Unauthorized copying and distribution of digital audio has been greatly facilitated by the wide availability of low-cost personal computers, portable devices, network access, and audio recording and editing software. One possible solution for copyright protection is audio watermarking. In this paper, we propose an effective and robust audio watermarking algorithm that employs both the discrete wavelet transform and the discrete cosine transform. The algorithm involves, first, pre-processing of the binary watermark image and then embedding it into the original audio by quantization of coefficients. Experiments on audio recordings of many different music styles confirm the robustness of the algorithm against a wide range of Stirmark attacks such as noise addition, compression, and filtering, as well as other common attacks.*

***Keywords:*** *Watermark, Quantization, DWT, DCT, Stirmark*

## 1. Introduction

With the rapid development and wide application of multimedia technology and network technology, video, audio, image, and other various multimedia works can be obtained in digital form, it brings great convenience for people's work, study and life. But at the same time it may also leads to large-scale non-licensed copy, which is likely to damage the healthy development of the publishing industry. This has been particularly true for digital audio. However, one promising means to address the issue is through the use of digital audio watermarking technology [1-3].

At present, the audio watermarking research is not at the same level of maturity as research into image and video watermarking. This is due to the fact that the human auditory system is much more sensitive than the human visual system, and that inaudibility is much more difficult to achieve than invisibility for images. Furthermore, audio signals are represented by much fewer samples per time interval, which indicates that the amount of information capacity that can be embedded robustly and inaudibly in audio files is much lower than the amount of information that can be embedded in visual files. Nonetheless, many audio watermarking techniques have been proposed in the literature in recent years. These techniques can be grouped into three categories: patchwork in the frequency domain [1], echo hiding in the time domain [4, 5], and spread-spectrum [6-8].

Just like more and more image and video watermarks employ a variety of frequency-domain transforms [9-13], the research on audio watermark is also more and more attention to the combined use of multiple frequency-domains. In this paper, we propose a quantization algorithm for audio watermarking based on the combination of the

ISSN: 1738-9976 IJSIA

discrete wavelet transform (DWT) and the discrete cosine transform (DCT). In Section 2, we introduced the proposed algorithm, including the watermark pre-processing step and the embedding and extracting steps. Section 3 describes the results of attack experiments on many samples of different music styles to examine the robustness against common attacks. Finally, our conclusions are presented in the last section.

## 2. Watermark Algorithm

Here, we employ a binary image as the watermark. First, we evenly divide the host audio into N segments and apply the DWT to each selected segment. Then, the DCT is applied to the low-frequency coefficients of segments that have been subjected to DWT. Next, the watermark information is embedded into the quantized coefficients as follows:

According to the quantized interval q, we can divide the coefficients into Class A for 1 and Class B for 0 [14] as follows:

$$\begin{aligned} A_i &= 2iq + q/2 \\ B_i &= 2iq - q/2,\ i = 0, \pm 1, \pm 2, \ldots \end{aligned} \tag{2.1}$$

If the value of watermark bit is 1, it is quantized as the midpoint of the nearest Class A; otherwise, it is quantized as the midpoint of the nearest Class B. Finally, all the segments are reassembled to obtain the watermarked audio.

### 2.1. Watermark Preprocessing

Here, we choose a meaningful binary image w of size M × N as the watermark. In order to enhance the confidentiality of the watermark image and reduce its relevance, we use the Arnold transform to scramble the watermark. This disrupts the image in a regular manner, making it appear as random noise. Consider an image of size M × M in which the gray value at position (i, j) is moved to (i', j') through the following transformation:

$$\begin{bmatrix} i' \\ j' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} (\bmod M) \tag{2.2}$$

If we know the scrambling cycle, it is very easy to descramble the image because of the reversibility of the Arnold transform.

Note that the watermark image w is a two-dimensional signal whereas the audio signal is only one-dimensional. Therefore, it is necessary to convert the image into a one-dimensional signal as follows:

$$V = \{v(k) = w(i, j), 1 \le i \le M, 1 \le j \le N, k = i + j \times (M-1)\} \tag{2.3}$$

Thus, the element v(k) of the sequence V denotes the pixel w(i, j) of watermark w.

### 2.2. Proposed Algorithm

As illustrated in Figure 1, the procedure for watermark embedding is as follows.

Step 1) Divide the original audio I of length LEN into segments Q (Q = LEN / piece) evenly where "piece" is the length of every segment.

Step 2) Extract all the even-numbered segments of the original audio and renumber them. Then apply the H-level DWT to all the renumbered segments.

Step 3) Apply the DCT to the low-frequency DWT coefficients.

Step 4) Embed the preprocessed watermark image (of size M × N). Check the number of the renumbered segments: if the number is even, then embed only one bit into the DC component of the even-numbered segment by quantization; if the number is odd, then embed $\lceil (M \times N \times 2)/(Q/2) \rceil - 1$ bits into the AC components of the odd-numbered segment. After embedding the watermark, the inverse DWT and inverse DCT are applied to each segment.

Step 5) Repeat step 4) until all the watermark information is embedded.

Step 6) Reassemble the audio segments to obtain the watermarked audio.

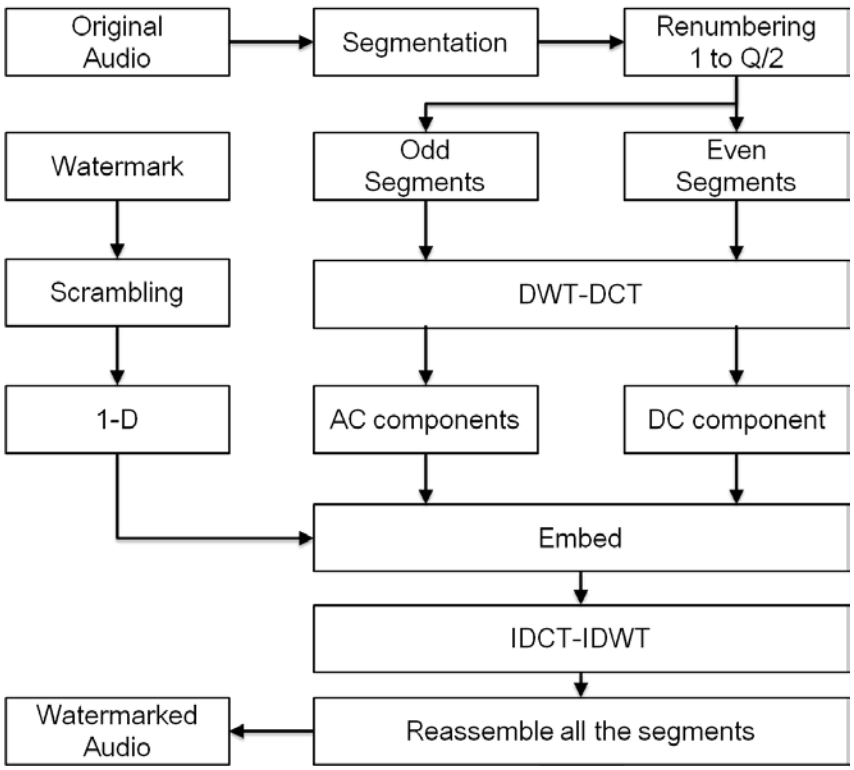



**Figure 1. Illustration of Embedding Algorithm**

After processing all the steps, we can get the watermarked audio. The watermark extracting scheme is the inverse of the embedding procedure. That is blind extraction. As illustrated in Figure 2, the procedure is as follows:

Step 1) Read the watermarked audio *I'* of length *LEN'*, and divide it evenly into segments *Q (Q = LEN' / piece)* and numbered each segment.

Step 2) Extract all the even-numbered segments and renumber them. Apply the H-level DWT and then apply the DCT on the low-frequency coefficients for all the renumbered segments.

Step 3) Extract the DC component of the even-numbered segments and the bits of the AC components of the odd-numbered segments.

Step 4) According to the quantization algorithm, determine whether the watermark information extracted from the DC and AC components is a 1 or 0.
Step 5) Repeat steps 3) and 4) until all the watermark information has been extracted.
Step 6) Transform the one-dimensional watermark information into a two-dimensional watermark image, and obtain the original watermark by Arnold descrambling.

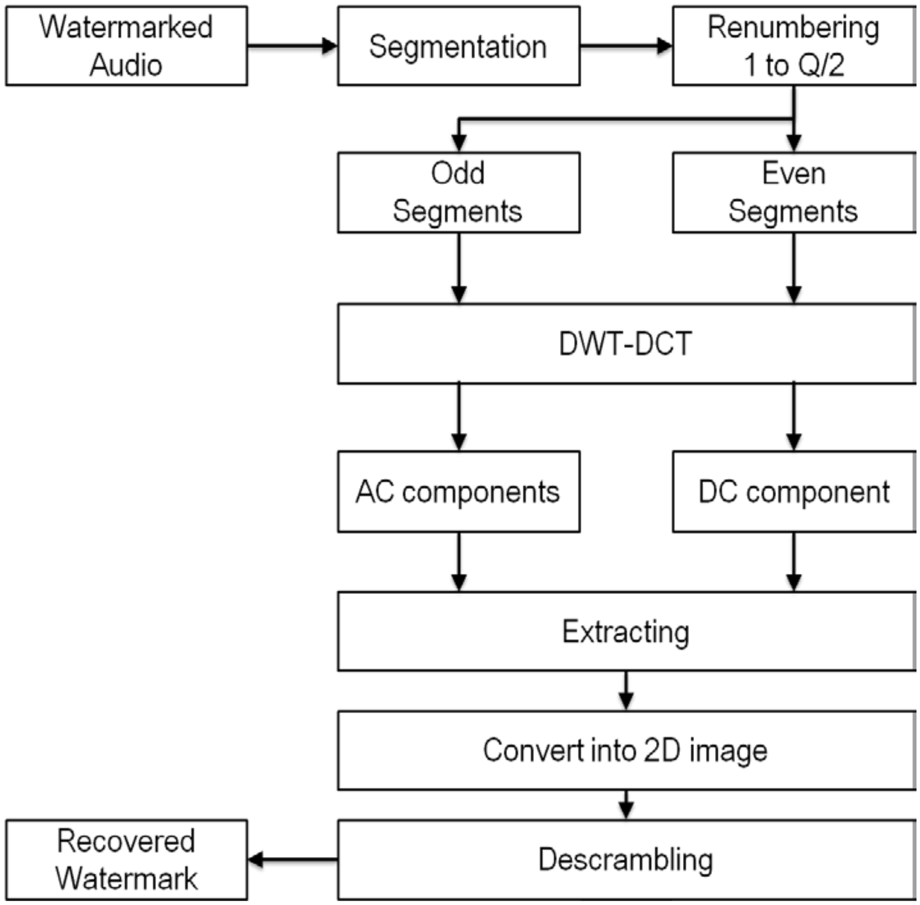



**Figure 2. Illustration of Extraction Algorithm**

## 3. Experimental Results and Analysis

In the experiment, we chose 10–s-long mono audio signal with a sampling frequency of 44.1 kHz and a resolution of 16 bits. A binary image which is made by text “YD” of size 64 × 64 pixels was used as the watermark. For the DWT, we adopted the common wavelet "Daubechies-1" and level H = 3.

We tested our algorithms on audio samples of different music styles.

As is shown in Table 1, the average SNR was 31.38dB, and the watermark was extracted exactly with a BER of 0 and NC of 1.

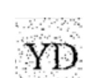
 

**Table 1. Test on Different Music Styles**

| Audio content | SNR(dB) |
|---|---|
| Light music | 30.32 |
| Hip hop | 32.31 |
| Jazz | 32.55 |
| R&B | 31.94 |
| Rock | 30.03 |
| Classical | 31.70 |

Next, to evaluate the performance of watermark objectively, we subjected the watermarked audio signal to some common attacks using the Stirmark Benchmark, including compression, noise addition, and low-pass filtering; the experimental results are shown in Table 2. From the results, we see that the watermark can be correctly extracted and identified clearly despite these attacks.

**Table 2. Experimental Results for Proposed Algorithm**

| Evaluation | SNR (dB) | NC | Extracted watermark |
|---|---|---|---|
| No attacks | 30.32 | 1 |  |
| Compression | 19.74 | 1 |  |
| Noise addition (900) | 14.20 | 0.98 |  |
| Low-pass filter | 23.63 | 0.97 |  |

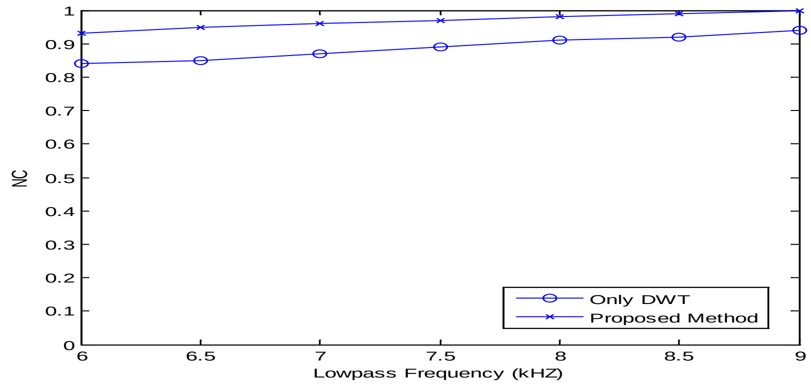



**Figure 3. Comparison of Different Low-pass Frequencies**

The robustness of the audio samples of different music styles to the various attacks is described in Table 3. From these results, we conclude that the proposed algorithm is robust to these common attacks.

Figure 3 compares the different low-pass frequencies between samples subjected to the proposed method and those subject to another method that only employs DWT. It is obvious that the robustness of our proposed method is superior to that of methods that only use DWT.

Table 4 lists the experimental results for the typical Stirmark attacks such as brumn addition, noise addition, compression, dyn_noise, exchange, extra stereo, and low-pass filtering.

**Table 3. Robustness for Different Music Styles**

| Audio content | SNR(dB)/NC | | |
|---|---|---|---|
| | Compression | Noise addition | Filtering |
| Light music | 19.74 | 14.20 | 23.63 |
| | 1 | 0.98 | 0.97 |
| Classical | 17.53 | 19.68 | 31.10 |
| | 1 | 0.98 | 0.91 |
| R&B | 27.95 | 16.11 | 27.17 |
| | 1 | 0.98 | 0.92 |
| Hip hop | 27.36 | 16.40 | 22.37 |
| | 1 | 0.98 | 0.93 |
| Rock | 21.37 | 19.94 | 21.29 |
| | 0.99 | 1 | 1 |
| Jazz | 14.27 | 22.57 | 24.73 |
| | 0.97 | 1 | 0.96 |

**Table 4. Results for Typical Stirmark Attacks**

| Name of attack | BER | NC | SNR(dB) |
|---|---|---|---|
| Add_brumm(100) | 0 | 1 | 29.25 |
| Add_noise(900) | 0 | 1 | 19.95 |
| Add_sinus | 0.36 | 0.78 | 14.92 |
| Amplify | 0.49 | 0.68 | 6.02 |
| Compressor | 0.02 | 0.99 | 21.37 |
| Dyn_noise | 0.30 | 0.82 | 18.67 |
| Echo | 0.50 | 0.68 | 1.75 |
| Exchange | 0 | 1 | 12.45 |
| Extra_stereo(70) | 0 | 1 | 30.01 |
| Lsb_zero | 0 | 1 | 30.02 |
| Nothing | 0 | 1 | 30.01 |
| Rc_lowpass | 0 | 1 | 21.29 |
| Smooth | 0 | 1 | 20.85 |
| Stat | 0.05 | 0.98 | 18.25 |
| Zero_cross | 0.01 | 0.99 | 25.29 |

## 4. Conclusions

We proposed a hybrid DCT and DWT audio watermarking algorithm. Experimental results demonstrate that it is robust to common attacks such as compression, noise addition, and low-pass filtering. It is effective for recordings of many music styles. Moreover, the proposed algorithm outperforms methods that only employ DWT in some of the attacks, although the embedding capacity is decreased. However, the algorithm could be improved by modifying the approach for selecting the coefficients in which the watermark information is embedded.

## Acknowledgements

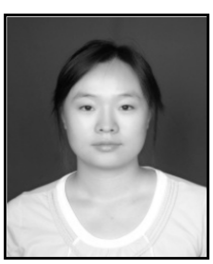
This research project was supported by the Ministry of Culture, Sports and Tourism (MCST) and the Korea Copyright Commission in 2011.

## References

[1] K. Yeo and H. J. Kim, "Modified patchwork algorithm: A novel audio watermarking scheme", IEEE Transactions on Speech and Audio Processing, vol. 11, no. 4, **(2003)**, pp. 381-386.
[2] E. Erelebi and L. Bataki, "Audio watermarking scheme based on embedding strategy in low frequency components with a binary image", Digital Signal Processing, vol. 19, no. 2, **(2009)**, pp. 265-277.
[3] P. K. Dhar and J.-M. Kim, "Digital Watermarking Scheme Based on Fast Fourier Transformation for Audio Copyright Protection", International Journal of Security and Its Application, vol. 5, no. 2, **(2011)**, pp. 33-48.
[4] P. Bassia, I. Pitas and N. Nikoladis, "Robust audio watermarking in time domain", IEEE Transactions on Multimedia, vol. 3, no. 2, **(2001)**, pp. 232-241.
[5] W. Bender, D. Gruhl and N. Morimoto, "Techniques for data hiding", IBM Systems Journal, vol. 35, no. 3-4, **(1996)**, pp. 131-336.
[6] L. Cui, S. Wang and T. Sun, "The application of binary image in digital audio watermarking", International Conference on Neural Networks and Signal Processing, Proceedings of the ICNNSP, Nanjing, China, vol. 2, **(2003)**, pp. 1497-1500.
[7] I. J. Cox, J. Kilian, F. T. Leigton and T. Shamoon, "Secure spread spectrum watermarking for multimedia", IEEE Transactions on Image Processing, vol. 6, no. 12, **(1997)**, pp. 1673-1687.
[8] D. Kirovski and H. Malvar, "Robust spread-spectrum audio watermarking", IEEE ICCASP 2001, Proceedings of ICASSP Salt Lake City, USA, **(2001)**, pp. 1345-1348.
[9] D. Li and J. Kim, "Secure Image Forensic Marking Algorithm using 2D Barcode and Off-axis Hologram in DWT-DFRNT Domain", Applied Mathematics and Information Sciences, vol. 6-2S, **(2012)**, pp. 513S-520S.
[10] J. Kim, N. Kim, D. Lee, S. Park and S. Lee, "Watermarking two dimensional data object identifier for authenticated distribution of digital multimedia contents", Signal Processing: Image Communication, vol. 25, no. 8, **(2010)**, pp. 559-576.
[11] J. Nah, J. Kim and J. Kim, "Video Forensic Marking Algorithm using Peak Position Modulation", Applied Mathematics and Information Sciences, vol. 6-3S, **(2012)**, pp. 797S-802S.
[12] O.-J. Kwon, W.-G. Kim and Y. Kim, "Binary Image Watermarking with Tampering Localization in Binary Wavelet Transform Domain", International Journal of Multimedia and Ubiquitous Engineering, vol. 7, no. 2, **(2012)**, pp. 341-346.
[13] V. Saxena, P. Khemka, A. Harsulkar and J. P. Gupta, "Performance Analysis of Color Channel for DCT Based Image Watermarking Scheme", International Journal of Security and Its Application, vol. 1, no. 2, **(2007)**, 41-46.
[14] L. Chen, Z. Yao and L. Chen, "The project of audio frequency watermarking based on odd-even quantization fight synchronous attack", The Computer Engineering and Application, vol. 44, no. 36, **(2008)**, pp. 112-114.

## Authors

**De Li** received the Ph.D. degree from Sangmyung University, major in computer science in 2005. He is currently a professor of Dept. of Computer Science at Yanbian University in China. He is also a Principal Researcher at Copyright Protection Research Institute, Sangmyung University. His research interests are in the areas of copyright protection technology, digital watermarking, and digital forensic marking.



**Yingying Ji** is a postgraduate, major in Information Security, now studying at Yanbian University in China. Her research interests are in the areas of copyright protection technology, information security, digital watermarking, and digital forensic marking.

**JongWeon Kim** received the Ph.D. degree from University of Seoul, major in signal processing in 1995. He is currently a professor of Dept. of Copyright Protection 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.