A Novel Robust Zero-Watermarking Scheme Based on Discrete Wavelet Transform

Yu Yang[1, 2] Min Lei[1, 2]
1. Information Security Center, Beijing University of Posts and Telecommunications, Beijing, China
2. National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing, China
Email: yangelm@vip.sina.com, byleimin@tom.com

Huaqun Liu
School of Information and Mechanical Engineering, Beijing Institute of Graphic Communication, Beijing, China

Yajian Zhou[1, 2] Qun Luo[1]
1. Information Security Center, Beijing University of Posts and Telecommunications, Beijing, China
2. National Engineering Laboratory for Disaster Backup and Recovery, Beijing University of Posts and Telecommunications, Beijing, China

Abstract—**In traditional watermarking algorithms, the insertion of watermark into the original signal inevitably introduces some perceptible quality degradation. Another problem is the inherent conflict between imperceptibility and robustness. Zero-watermarking technique can solve these problems successfully. But most existing zero-watermarking algorithm for audio and image cannot resist against some signal processing manipulations or malicious attacks. In the paper, a novel audio zero-watermarking scheme based on discrete wavelet transform (DWT) is proposed, which is more efficient and robust. The experiments show that the algorithm is robust against the common audio signal processing operations such as MP3 compression, re-quantization, re-sampling, low-pass filtering, cutting-replacement, additive white Gaussian noise and so on. These results demonstrate that the proposed watermarking method can be a suitable candidate for audio copyright protection.**

Index Terms—**zero-watermarking, discrete wavelets transform, copyright protection**

I. Introduction

Recently, the rapid development of the Internet has increased multimedia services, such as electronic commerce, pay-per-view, video-on-demand, electronic newspapers, and peer-to-peer media sharing. As a result, multimedia data can be obtained quickly over high speed network connections. However, the authors, publishers, owners and providers of multimedia data are reluctant to grant the distribution of their documents in a networked environment because the ease of intercepting, copying and redistributing electrical data in their exact original form encourages copyright violation. Therefore, it is crucial for the future development of networked multimedia systems that robust methods are developed to protect the intellectual property right of data owners against unauthorized copying and redistribution of the material made available on the network. Classical encryption systems do not completely solve the problem of unauthorized copying because once encryption is removed from a document, there is no more control of its dissemination.

Digital watermarking is a good approach for providing copyright protection of digital contents (for example text, word, image, audio, video) [1-14]. This technique is based on direct embedding the watermarking information into the digital contents. The embedding operations should not be introduced any perceptible distortions ideally, there must be no perceptible difference between the watermarked and the original version. That is to say the watermark data should be embedded imperceptibly into the audio media. Apart from imperceptibility, capacity and robustness are two fundamental properties of audio watermarking schemes. To make tradeoff between imperceptible and robust, it’s a common method to incorporate the human auditory system (HAS) into a watermarking system. The watermark should be extractable after various intentional and unintentional attacks. These attacks may include additive noise, re-sampling, MP3 compression, low-pass filtering, re-quantization, and any other attack which removes the watermark or confuse the watermark extraction system. Considering a trade-off between capacity, transparency and robustness is the main challenge for audio watermarking applications.

We focus on the audio watermarking scheme in this paper. In traditional audio watermarking techniques, either in spatial domain, transform domain, or dual

doi:10.4304/jmm.7.4.303-308

domain [15-16], the embedding of watermark into the original audio inevitably introduces some audible quality degradation. Another problem is the inherent conflict between the imperceptibility and robustness. Then, zero-watermarking technique was proposed by some researchers to solve these problems [17–25]. Instead of embedding watermark into the original signal, the zero-watermarking approach just constructs a binary pattern based on the essential characteristics of the original signal and uses them for watermark recovery. In [21], the property of the natural images that the vector quantization indices among neighboring blocks tend to be very similar was utilized to generate the binary pattern. A scheme that combined the zero-watermarking with the spatial-domain-based neural networks was proposed in [22], in which the differences between the intensity values of the selected pixels and the corresponding output values of the neural network model were calculated to generate the binary pattern. Some low-frequency wavelet coefficients were randomly selected from the original image by chaotic modulation and used for character extraction in [23]. And in [24], two zero-watermarks were constructed from the original image. One was robust to signal process and central cropping, which was constructed from low-frequency coefficients in discrete wavelet transform domain and the other was robust to general geometric distortions as well as signal process, which was constructed from DWT coefficients of log-polar mapping of the host image. There are novel zero-watermarking algorithms for audio now. An efficient and robust zero-watermarking technique for audio signal is presented in [25]. The multi-resolution characteristic of discrete wavelet transform (DWT), the energy compression characteristic of discrete cosine transform (DCT), and the Gaussian noise suppression property of higher-order cumulant are combined to extract essential features from the original audio signal and they are then used for watermark recovery. Simulation results demonstrate the effectiveness of the scheme in terms of inaudibility, detection reliability, and robustness. However, all these zero-watermarking techniques are designed for still image and their robustness against some signal processing manipulations or malicious attacks is not satisfying. In this paper, a novel robust zero-watermarking technique for audio signal is proposed.

The paper is organized as follows. Section II introduces basic concepts for discrete wavelet transform. Section III describes a novel robust zero-watermarking scheme. Section IV exhibits the experimental results illustrating that the proposed method can get good robustness. A brief conclusion can be available in Section V.

II. DISCRETE WAVELET TRANSFORM

Wavelet transform is a time domain localized analysis method with the window's size fixed and convertible. It supplies high temporal resolution in high frequency part, and high frequency resolution in low frequency part of signals. It can distill the information from signal effectively.

For audio, we need convert the 1-dimensional audio into 2-dimensional. Wavelet transform decomposes an audio into a set of band limited components which can be reassembled to reconstruct the original audio without error. Since the bandwidth of the resulting coefficient sets is smaller than that of the original audio, the coefficient sets can be down sampled without loss of information. Reconstruction of the original signal is accomplished by up sampling, filtering and summing the individual sub bands. For 1-D audio, we convert the 1-D audio into 2-D signal, and then apply DWT corresponds to processing the audio by 2-D filters in each dimension. The filters divide the input audio into four non-overlapping multi-resolution coefficient sets, a lower resolution approximation audio (LL1) as well as horizontal (HL1), vertical (LH1) and diagonal (HH1) detail components. The sub-band LL1 represents the coarse-scale DWT coefficients while the coefficient sets LH1, HL1 and HH1 represent the fine-scale of DWT coefficients. To obtain the next coarser scale of wavelet coefficients, the sub-band LL1 is further processed until some final scale N is reached. When N is reached we will have 3N+1 coefficient sets consisting of the multi-resolution coefficient sets LLN and LHx, HLx and HHx where x ranges from 1 until N. Figure 1 show the representation of 2-levels of wavelet transform.

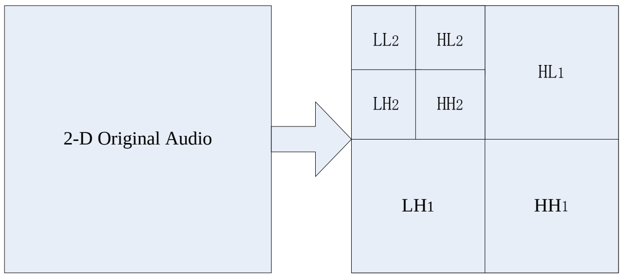


Figure 1: Representation of 2-Levels of Wavelet Transform

Due to its excellent spatio-frequency localization properties, the DWT is very suitable to identify the areas in the original audio where a watermark can be embedded effectively. In particular, this property allows the exploitation of the masking effect of the human visual and auditory system such that if a DWT coefficient is modified, only the region corresponding to that coefficient will be modified. In general most of the audio energy is concentrated at the lower frequency coefficient sets LLx and therefore embedding watermarks in these coefficient sets may degrade the audio significantly. Embedding in the low frequency coefficient sets, however, could increase robustness significantly.

III. AUDIO WATERMARKING ALGORITHM

A. Embeddbing Process

Let A be the original audio signal and let $W=\{w(i,j) \mid i=1,2,...,M_1,\ j=1,2,...,M_2\}$ be the binary-valued image watermark to be embedded. Now we describe the watermark embedding procedure. (Figure 2)

Step 1: Dimension reduction of watermark image. Because the audio signal is one- dimensional and watermarking W is two-dimensional. Therefore the watermarking should be reduced into one-dimensional signal:

$$w=\{w(i)=w(m_1,m_2)\big|0\le m_1\le M_1, 0\le m_2\le M_2, i=M_1\times M_2\}$$

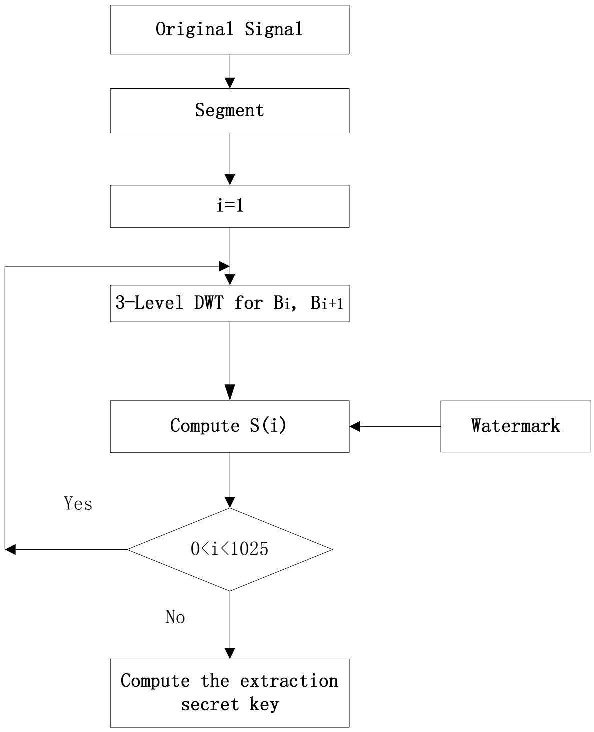


Figure 2: Embedding Process

Step 2: Let the original audio $A=(a_1,a_2,...,a_N)$ with N PCM (pulse-code modulation) samples be segmented into $M=\lfloor N/512\rfloor$ blocks B_i, where $M=M_1\times M_2, i=1,2,...,M$. Each block includes 512 samples.

Step 3: for *i=1: M*

3.1). The 3-level discrete wavelet packet decomposition is applied to the audio block B_i, and can get the approximate sub-band coefficients $C_i=\{c_i(1),c_i(2),...,c_i(64)\}$, which has 64 samples.

3.2). Computes

$$t(i)=\sum_{j=1}^{64}c_i(j)\times c_i(j)\times j,$$

$$t'(i)=\sum_{j=1}^{64}c_i(j)\times c_i(j)$$

According to the approximate sub-band coefficients $C_i=\{c_i(1),c_i(2),...,c_i(64)\}$. And gets:

$$s'(i)=\lfloor t(i)/t'(i)\rfloor.$$

3.3). Computes $s(i)=s'(i)(\mathrm{mod}\,2)$.

3.4). End

Step 4: We can get the watermarking extraction secret key $k(i)=w(i)\oplus s(i)$ through computing the exclusive or operation between $s(i)$ and watermarking information $w(i)$.

B. Extraction Process

The watermark recovery procedure can be carried out as follows (Figure 3).

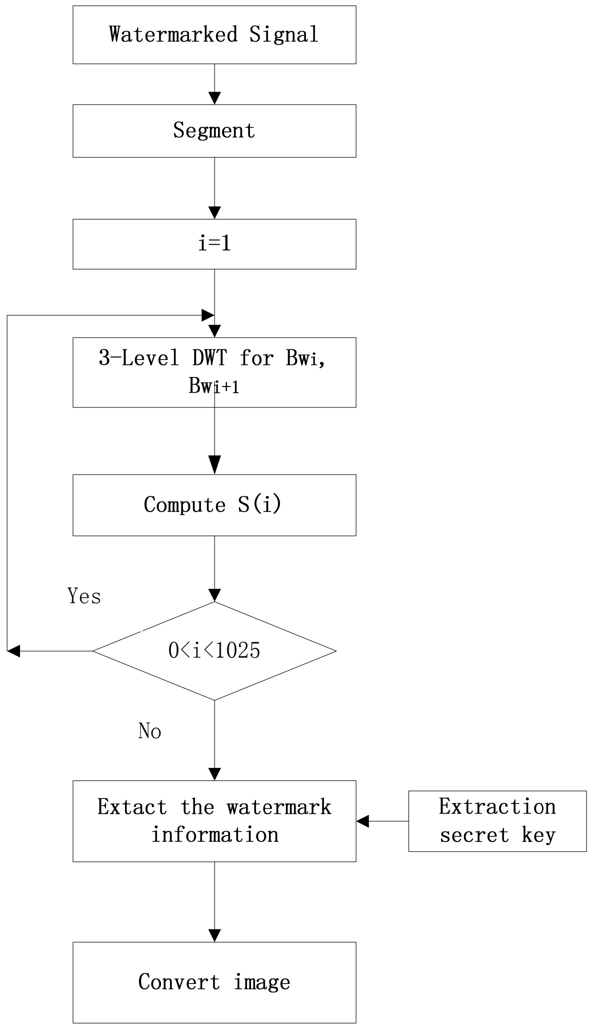


Figure 3: Extraction Process

Step 1: Let the watermarked audio $A'=(a'_1,a'_2,...,a'_N)$ be segmented into $M=\lfloor N/512\rfloor$ blocks B'_i, where $M=M_1\times M_2, i=1,2,...,M$. Each block includes 512 samples.

Step 2: for *i=1: M*

2.1). The 3-level discrete wavelet packet decomposition is applied to the audio block B'_i, and can get the approximate sub-band coefficients $C'_i=\{c'_i(1), c'_i(2),...,c'_i(64)\}$ which has 64 samples.

2.2). Computes

$$T(i)=\sum_{j=1}^{64}c'_i(j)\times c'_i(j)\times j\ ,$$

$$T'(i)=\sum_{j=1}^{64}c'_i(j)\times c'_i(j)$$

according to the approximate sub-band coefficients $C'_i=\{c'_i(1),c'_i(2),...,c'_i(64)\}$. And gets

$$S'(i)=\lfloor T(i)/T'(i)\rfloor .$$

2.3). Computes $S(i)=S'(i)(\bmod 2)$.

2.4). End

Step 3: We can get the watermarking information $w'(i)=k(i)\oplus S(i)$ through computing the exclusive or operation between $S(i)$ and he watermarking extraction secret key $k(i)$.

Step 4: We can get the binary-valued image watermark according to the extracted watermarking information.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

In this experiment, two binary stamp image with size 32×32 (i.e., $M_1=M_2=32$), displayed in Fig. 5, are taken as the original watermark And we evaluate the performance of our proposed watermarking method for different types of 16 bit mono audio signals sampled at 44.1 KHz as shown in Fig. 4. The sound files are: (a) Pop, (b) Speech, (c) Classic. Each audio file contains 524,288 samples with duration of 11 seconds.

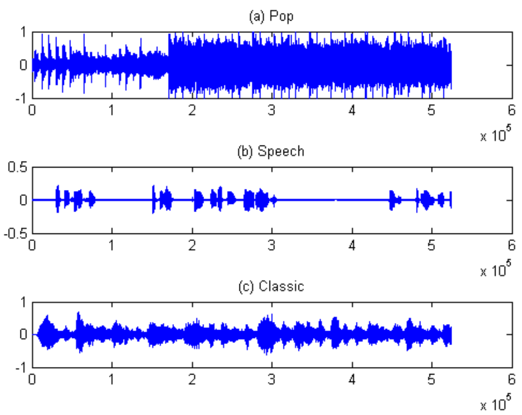


Figure 4: The signal of original audio

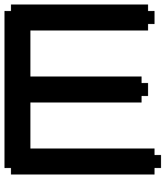


Figure 5: Watermarks with size 32×32

In order to evaluate the quality of watermarked audio, the following signal-to-noise radio (SNR) equation is used:

$$SNR=10\log_{10}\frac{\sum_{n=1}^{N}Y^2(n)}{\sum_{n=1}^{N}[Y(n)-Y'(n)]^2}$$

where $Y(n)$ and $Y'(n)$ are original audio signal and watermarked audio signal respectively.

In order to evaluate the robustness of the watermarked algorithm, the following normalized coefficient (NC) and bit error rate (BER) are employed respectively:

$$NC(W,W^*)=\frac{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2}W(i,j)W^*(i,j)}{\sqrt{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2}W(i,j)^2}\times\sqrt{\sum_{i=1}^{M_1}\sum_{j=1}^{M_2}W^*(i,j)^2}}$$

$$BER=(\sum_{i=1}^{M_1}\sum_{j=1}^{M_2}|W(i,j)-W^*(i,j)|)/(M_1\times M_2)$$

where $W(i,j)$ and $W^*(i,j)$ are the original watermark and extract watermark respectively, and $M_1\times M_2$ is the size of watermark.

And, the bit error rate (BER) was employed to measure the robustness of our algorithm,

$$BER=\frac{B}{M_1\times M_2}\times 100\%$$

where B is the number of erroneously extracted bits and $M_1\times M_2$ is the size of watermark.

A. Imperceptibility Anallysis

One of the main requirements of audio watermarking techniques is inaudibility of the embedded watermark. For the proposed scheme, this requirement is naturally achieved because the watermark is embedded into the secret key but not the original audio signal itself. Actually, the watermarked audio is the identical to the original one.

B. Robusness Test and Analysis

In order to test the robustness of our proposed method, eight different types of attacks is as following:

(1) Noise addition: 20 dB additive white Gaussian noise (AWGN) is added to the watermarked audio signal

(2) Re-sampling: The watermarked signal originally sampled at 44.1kHz is re-sampled at 22.050kHz, and then restored by sampling again at 44.1 kHz.

(3) Low-pass filtering: cut-off frequency is 11.025kKz.

(4) Re-quantization: the 16 bit watermarked audio signal is quantized down to 8 bits/sample and again re-quantized back to 16 bits/sample.

(5) MP3 compression: MPEG-1 layer 3 compression with 64 kbps is applied to the watermarked signal.

(6) MP3 compression: MPEG-1 layer 3 compression with 32 kbps is applied to the watermarked signal.

(7) MP3 compression: MPEG-1 layer 3 compression with 128 kbps is applied to the watermarked signal

(8) Noise reduction: Preset shape is Hiss removal.

(9) Equalization.

Table 1 show the NC and BER of the proposed watermarking method in terms of robustness against several kinds of attacks applied to three different types of watermarked audio signal “Speech”, “Classic” and “Pop” respectively.

Table 1 The result of attacks

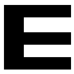
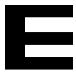
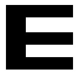
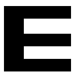
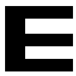
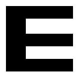
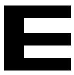
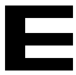
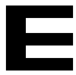
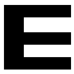
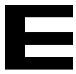
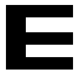
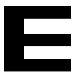
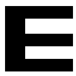
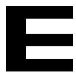
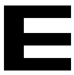
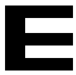
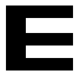
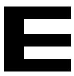
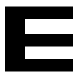
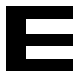
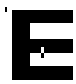
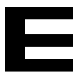
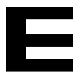
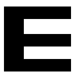
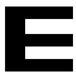
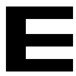
Audio Attack	Speech			Classic			Pop		
	NC	BER	Extracted image	NC	BER	Extracted image	NC	BER	Extracted image
1	1	0		1	0		1	0	
2	1	0		1	0		1	0	
3	1	0		1	0		1	0	
4	1	0		1	0		1	0	
5	1	0		1	0		1	0	
6	1	0		1	0		1	0	
7	1	0		1	0		1	0	
8	0.9991	0.001		1	0		1	0	
9	1	0		1	0		1	0	

Table 2 Comparison between our algorithm and other algorithm

Attack	NC(our)	NC([25])	NC([26])	NC([27])
Equalization	1	0.97	—	0.998
MP3 compression(32kbps)	1	0.99	—	0.999
Low-pass filtering	1	1	0.9971	0.824
Noise addition	1	1	0.9668	0.999
Re-sampling	1	1	—	0.981

The Table 2 compares our algorithm and other algorithms, we can easily get the robustness of our algorithm is more stronger than other algorithms.

V CONCLUSION

A novel audio zero-watermarking scheme based on discrete wavelet transform (DWT) is proposed in this paper, which is more efficient and robust. The experiments show that the proposed algorithm has good robustness against the common audio signal processing operations such as MP3 compression, re-quantization, re-sampling, low-pass filter, cutting-replacement and additive white Gaussian noise. These results demonstrate that our algorithm can be a suitable candidate for audio copyright protection.

ACKNOWLEDGMENT

This work is supported by the National Natural Science Foundation of China under Grant NO. 60973146, NO. 61170272, NO. 61170269, NO. 61003285, and Institute Level Project Funded by Beijing Institute of Graphic Communication under Grant NO. E-6-2012-23.

REFERENCES

[1] Min Lei, Research on the digital watermarking and steganalysis algorithms of audio information hiding[D].A dissertation for the degree of doctor of philosophy, Beijing University of Posts and Telecommunications. 2011:28-44.

[2] Yu Yang, Min Lei, Xiayun Jia etc.A New Digital Audio Watermarking Based on DCT-DWT-SVD. Proceedings 2011 IEEE International Conference on Information Theory and Information Security[C] August 7-9, 2011, Hangzhou China, IEEE PRESS, BEIJING,2011,906-909.

[3] Lei Min, Yang Yu Recent Advances in Audio-Based Steganalysis Research. Proceedings of the 2nd Asia-Pacific Conference on Information Network and Digital Content Security[C], August 7-9,2011, Hangzhou China, ATLANTIS PRESS, FRANCE, 2011, 208-213.

[4] LEI Min, YANG Yu, LUO Shou-shan et al. Semi-fragile Audio Watermarking Algorithm in DWT Domain [J]. China Communications, 2010, 7(4): 71-75

[5] S. C. Liu and S. D. Lin, "BCH code based robust audio watermarking in the cepstrum domain," *Journal of Information Science and Engineering*, vol. 22, pp. 535-543, 2006.

[6] M. Pooyan and A. Delforouzi, "Adaptive and Robust Audio watermarking in Wavelet Domain," in *Proceedings of International Conference on International Information Hiding and Multimedia Signal Processing (IIH-MSP 2007), vol. 2, pp. 287-290, 2007*

[7] G. Zeng and Z. Qiu, "Audio Watermarking in DCT: Embedding Strategy and Algorithm," in Proceedings *of 9th International Conference on Signal Processing (ICSP'09)*, pp. 2193-2196, 2008.

[8] Samira Lagzian, Mohsen Soryani and Mahmood Fathy, "A new robust watermarking scheme based on RDWT-SVD," *International Journal of Intelligent Information Processing*, Vol. 2, No 1, 2011.

[9] Pranab Kumar Dhar, Mohammad Ibrahim Khan and Saif Ahmad, "A new DCT-based watermarking method for copyright protection of digital audio," *International Journal of Computer Science & Information Technology*. Vol. 2, No. 5, 2010, pp. 91-101.

[10] Chen Yongqiang, Zhang yanqing, Hu Hanping and Ling Hefei, "A novel gray image watermarking scheme," *Journal of Software*, Vol. 6, No, 5, 2011, pp. 849-856.

[11] Kilari Veera Swamy, B. Chandra Mohan, Y. V. Bhaskar Reddy and S. Srinivas Kumar, "Image compression and watermarking scheme using scalar quantization," *The International Journal of Next Generation Network*, Vol. 2, No. 1, 2010, pp. 37-47.

[12] Hasan M. Meral, Emre Sevinc, Ersin Unkar, Bulent Sankur, A. Sumru Ozsoy and Tunga Gungor, "Syntactic tool for text watermarking," in Proceedings of the in *International society for Optical Engineering*, 2007.

[13] Pranab Kumar Dhar, Mohammad Ibrahim Khan and Jong-Myon, Kim, "A new audio watermarking system using discrete fourier transform for copyright protection," *International journal of computer science and network security*. 6(10), 2010: 35-40

[14] Vivekananda B K, Indranil S and Abhijit D, "An Audio Watermarking Scheme Using Singular Value Decomposition and Dither-Modulation Quantization," *Multimedia Tools and Applications*, 52(2-3), 2010: 369 - 383.

[15] X.-Y. Wang and H. Zhao, "A novel synchronization invariant audio watermarking scheme based on DWT and DCT," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4835–4840, 2006.

[16] X.-Y. Wang, W. Qi, and P. Niu, "A new adaptive digital audio watermarking based on support vector regression," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 8, pp. 2270–2277, 2007.

[17] T. Sun, W. Quan, and S.-X. Wang, "Zero-watermark watermarking for image authentication," in *Proceedings of the Signal and Image Processing*, pp. 503–508, Kauai, Hawaii, USA, August 2002.

[18] G. Horng, C. Chen, B. Ceng, and T. Chen, "Neural network based robust lossless copyright protection technique,"http://www.csie.cyut.edu.tw/TAAI2002/TAAI2002PDF/Parallel.

[19] Q. Wen, T.-F. Sun, and S.-X. Wang, "Concept and application of zero-watermark," *Tien Tzu Hsueh Pao/Acta Electronica Sinica*, vol. 31, no. 2, pp. 214–216, 2003.

[20] S. Yang, C. Li, F. Sun, and Y. Sun, "Study on the method of image non-watermark in DWT domain," *Chinese Journal Image Graphics*, vol. 8A, no. 6, pp. 664–669, 2003.

[21] D. Charalampidis, "Improved robust VQ-based watermarking," *Electronics Letters*, vol. 41, no. 23, pp. 21–22, 2005.

[22] J. Sang, X. Liao, andM. S. Alam, "Neural network based zero-watermark scheme for digital images," *Optical Engineering*, vol. 45, no. 9, 2006.

[23] C. Hanqiang, X. Hua, L. Xutao, L. Miao, Y. Sheng, and W. Fang, "A zero-watermarking algorithm based on DWT and chaotic modulation," in *Independent Component Analyses*, vol. 6247 of *Proceedings of SPIE*, pp. 1–9, Orlando, Fla, USA, 2006.

[24] L. Jing and F. Liu, "Double zero-watermarks scheme utilizing scale invariant feature transform and log-polar mapping," in *Proceedings of the IEEE International Conference onMultimedia and Expo*, pp. 2118–2121, Las Vegas,Nev, USA, February 2007.

[25] Ning Chen, Jie Zhu. "A robust zero-watermarking algorithm for audio," *EURASIP Journal on Advances in Signal Processing*, 2008.

[26] W Song, J. J. Hou, Z. H. Li and L. Huang, "A novel zero-bit watermarking algorithm based on logistic chaotic system and singular value decomposition," *Acta physica sinica*, Vol. 58, No. 7, 2009, pp. 4449-4456.

[27] L. Tan, B. Wu, Z. Liu and M. T. Zhou, 'An audio information on hiding algorithm with high capacity which based on Chaotic and wavelet transform," *Acta electronic sinica*, Vol. 38, No. 8, 2010, pp. 1812-1818.