Efficient Audio Zero-Watermarking Algorithm for Copyright Protection Based on BIC and DWCM Matrix

Xueming Li, Guangjun He
College of Computer Science, Chongqing University, Chongqing, 400030, China
lixuemin@cqu.edu.cn

Abstract

A widely studied algorithm for digital audio copyright protection is based on digital watermarking technology. In this paper, to make audio watermarking both avoid the contradiction between strong robustness and transparency, and resist the attack of random large-area cropping, a novel audio zero-watermarking scheme is proposed. An audio segmentation algorithm based on Bayesian Information Criterion (BIC) is introduced to divide audio signal into many different size segments firstly. After that, each segment is performed wavelet decomposition and extracted Double-direction Wavelet Coefficients Mapping (DWCM) Matrixes. In the DWCM Matrixes Space, the audio signal is implemented watermark embedding or detection. In the procedure of extracting DWCM Matrix, we employ a novel method to map the approximate coefficients of wavelet transformation of an audio segment into a binary matrix, which is the key process of zero-watermarking algorithm proposed in the paper. The experimental results demonstrate that the proposed zero-watermark scheme has more excellent performance in terms of security, robustness, anti-cropping compared to the earlier audio watermarking schemes. This is mainly contributed by the proposed DWCM Matrix. To our best knowledge, zero-watermarking based on BIC and DWCM Matrix has not been reported before.

Keywords*: Audio Segmentation, Copyright Protection, Double-direction Wavelet Coefficients Mapping (DWCM), Anti-cropping*

1. Introduction and motivations

In the latest years, many efforts have been devoted to the problem of digital audio copyright protection by plenty of research communities. This is because that multimedia data has become a widely used carrier of information and the rapid growth of Internet has made the copyright and integrity of audio information more and more important issues. To protect digital works against illegal use and tampering, digital watermarking has been proposed to accomplish copyright protection or content integrity authentication [1-2].

In the most existing methods of digital watermarking [7-10], in order to embedding watermark information into host audio, authors generally modify more or less original audio signal through various approaches, which generate the contradiction between robustness and transparency. In such a case, a zero-watermark algorithm was advocated to apply to image authentication by Wen Quan [4, 5] for the first time. It is a novel digital watermarking technology and generates watermark information from the significant feature of host signal without modifying signal itself data. Furthermore, the zero-watermarking technology gives an effective solution to the problems of protocol attack and forged-watermark attack through registered watermark information to an authoritative authentication center. In recent, the audio zero-watermarking scheme has acquired the considerable progress [11-17]. The existing audio watermarking algorithms [14] have excellent performance in terms of resisting common signal processes. However, the main weaknesses of the existing algorithms are low security and poor ability of anti-attack, special vulnerability to random large-area cropping [11-13].

To solve the problems above, a new audio zero-watermark algorithm is proposed in this paper, which makes the audio watermarking both avoid the contradiction between strong robustness and transparency, and build up the resistance to attack of random large-area cropping for digital audio copyright protection. Before watermark embedding or detection, an audio signal must be processed through following two main steps.

- Divide the audio signal into many different size segments by an audio segmentation algorithm based on BIC [3]. As we know, the illegal attacker usually use a portion of original audio signal, so we must make sure that the imbedded watermark can be detected in a tiny segment.

- Extracted the Double-direction Wavelet Coefficients Mapping (DWCM) Matrix from each audio segment. That is, the approximate coefficients sequences of wavelet decomposition of the audio segment and its reverse segment are transformed into binary square matrixes *MM* and *reMM* respectively.

Then, in order to express the copyright information intuitively, an exclusive-or operation is done on the DWCM Matrixes and a binary copyright watermark image, the result *K* of which serves as the key of watermark detection. When detecting watermark of an audio segment, it can be obtained by exclusive-or operation on DWCM Matrixes and the key *K* of watermark detection. By filtering the results of above XOR operation, the best quality watermark of the entire audio can be acquired. Because of without participation of original audio signal, the detection is a blind procedure.

The rest of this paper is organized as follows. In section 2, we give the definition of BER. In section 3, we detail the audio segmentation algorithm based on BIC and examine the invariance of segmentation position against geometric distortions, especially large-area cropping. In section 4, we propose the approach of extracting the Double-direction Mapping Matrixes on the approximate coefficients of wavelet transform of an audio segment, and the scheme of watermark embedding and detection in detail. We show the experimental results in section 5. Finally, conclusions and discussions are drawn.

2. Bit error rate

In the communication system, the bit error rate (BER) is the number of bit errors divided by the total number of transferred bits during a studied time interval. The BER is employed to measure the similarity of two binary images with same size in this paper. That is, the BER (in percent) of two binary images $I_{N\times N}$ and $W_{N\times N}$ is defined as follow.

$$BER(I,W)=\frac{100}{N\times N}\sum_{i=1}^{N}\sum_{j=1}^{N}\begin{cases}1, & I(i,j)\neq W(i,j)\\ 0, & I(i,j)=W(i,j)\end{cases} \tag{1}$$

If the value of *BER* of two images is close to 0, the similarity between them is very high; vice versa. BER will be used to measure the resistance ability of watermarking algorithm to various attacks.

3. Audio preprocessing

Audio preprocessing includes audio segmentation and segment filtering. Bayesian Information Criterion (BIC) is employed to locally detect the single changes in the audio clip within a sliding window of variable size. Basically an audio clip is divided into several different length segments according to these change points, and the length (number of samples) of segment can be controlled by adjusting the sensitivity factor λ.

Briefly, given a sequence $O_1 \ldots O_N$ of observation vectors in the R^d space containing at most one change, the method based on the BIC for audio segmentation rests on the computation of:

$$\Delta BIC_i=\frac{N}{2}\log|\Sigma|-\frac{i}{2}\log|\Sigma_1|-\frac{(N-i)}{2}\log|\Sigma_2|-\lambda P \tag{2}$$

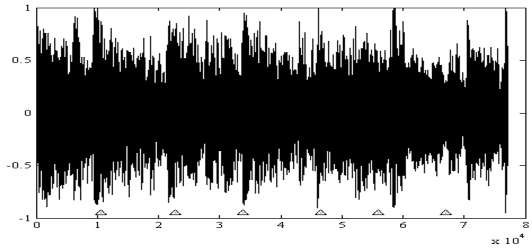
for each time index *i*, where $P=\frac{1}{2}(d+\frac{d(d+1)}{2})\log(N)$, and Σ, Σ_1 and Σ_2 are the covariance matrixes estimated on $O_1 \ldots O_i$ and $O_{i+1} \ldots O_N$ respectively. The value of *i* that maximizes ΔBIC_i is the most likely time index for a change and if $\Delta BIC_{i_{\max}}>0$ then i_{max} is confirmed to be a change. The sensitivity of the method can be tuned by adjusting the value λ to the particular task [3].

After segmenting an audio stream with BIC algorithm, every segment will be embedded or extracted watermark information with the algorithm in Section 4.

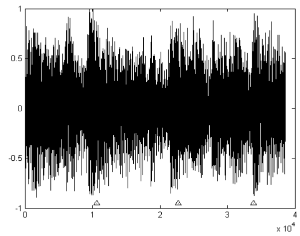
The audio physical segmentation positions (time index) based on BIC are invariable and robust nearly to lager-area cropping. In our proposed watermarking algorithm, the excellent character is utilized to resist the attack of random lager-area cropping, which is showed in Figure 1.

The audio signal, which will be embedded or detected watermark, is down-sampled into 11025 Hz and mono-channel. In the framing progress of audio segmentation approach based on BIC, the length of frame is 256 samples (23ms) with 50% overlapping. In the feature extraction phase, two types of features are calculated for each audio frame: (1) 12 order Mel Frequency Cepstrum Coefficient (MFCC),

(2) short-time average energy. In the way, an audio frame is represented by a 13-dimensional feature vector. The first 12 dimension features are represented by MFCC feature, the last one is short-time average energy [6].



(a) Segmentation positions (Δ) without cut-off



(b) Segmentation positions (Δ) after cropped 3s from the end of audio sequence

Figure 1. Cropping effect comparison with BIC audio segmentation

After splitting original audio into many segments, those audio segments with very little samples should be filtered. Not all of segments must be embedded or extracted watermarks from them. For example, if the length of a segment is less than half of audio frequency, it will be filtered out.

4. Zero-watermarking algorithm based on wavelet coefficients mapping

The watermark embedding and detection algorithm is based generally on this truth, that is, the excellent characters of exclusive-or operation: $A \oplus A = 0$, $A \oplus 0 = A$ and $A \oplus B = B \oplus A$. Suppose that the approximate coefficients of wavelet transform of an audio segment *X* are mapped into a binary matrix *M* by a kind of mapping method described in Part 4.1. Thus the key of the audio segment is $K = M \oplus W$. Meanwhile, the approximate coefficients of wavelet transformation of an unknown audio segment *X'*, which is supposed to be similar to *X* almost completely, is also mapped into a binary matrix *M'*. Because of the likeness of *X* and *X'*, *M'* is similar to *M* on bits, that is, the value of BER between *M'* and *M* is close to 0. According to the above characters of XOR, the following approximate equality can be obtained.

$$\begin{aligned} W' &= K \oplus M' \\ &= (M \oplus W) \oplus M' \\ &= (M \oplus M') \oplus W \\ &\xrightarrow{\because M \approx M'} W' \approx W \end{aligned} \tag{3}$$

Therefore, the owner of watermarking *W* and key *K* can claim the ownership of the unknown audio segment *X'*.

Let $A = \{a(i) \mid 0 \le i \le L\}$ presents an audio signal. $W = \{w(i, j) \in \{0,1\} \mid 1 \le i, j \le N\}$ is a binary watermark image. The special steps of our proposed algorithm are elaborated as follows.

4.1. Extract double-direction wavelet coefficients mapping matrixes

Let $X = \{x_1, x_2, ..., x_n\}$ presents an audio segment. Extracting the DWCM Matrix from a segment is that the approximate coefficients of wavelet decomposition of the audio segment and its reverse segment are transformed into two binary square matrixes. And DWCM Matrix extraction is a basis procedure before embedding and detecting watermark. The block diagram of extracting DWCM Matrix is shown in Figure 2.

The specific procedures of extracting DWCM Matrix from an audio segment are elaborated as follows:

Step 1: Perform m-level wavelet decomposition on the audio segment to get the approximate coefficients cA^m and detail coefficients $cD^m,...,cD^1$. Let L_{cA^m} denotes the length of cA^m.

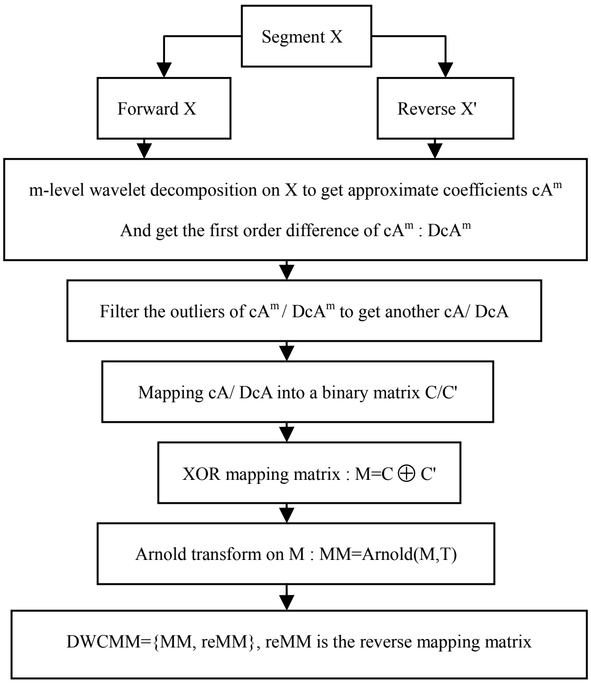


Figure 2. The block diagram of extracting DWCM Matrix

Step 2: Filter the outliers of cA^m which are lower than Q_α and upper than $Q_{1-\alpha}$ to get another coefficient sequence cA.

$$cA = \{cA^m(i) \mid Q_\alpha \le cA^m(i) \le Q_{1-\alpha},\ i = 1,2,...,L_{cA^m}\} \tag{4}$$

The Q_α and $Q_{1-\alpha}$ are α quantile and 1-α quantile of cA^m respectively. The value of α is between 0.0 and 0.1. In our experiment, the value of α is 0.05. Let L_{cA} denotes the length of cA.

Step 3: Map the coefficient sequence cA into a $N \times N$ matrix C. Initialize C with 0 firstly. Afterwards, update the elements of C through following approach.

$$T(i) = \frac{N \times N - 1}{\max(cA) - \min(cA)}[cA(i) - \min(cA)], \tag{5}$$

Then, $$C\left(\left\lfloor \frac{T(i)}{N} \right\rfloor + 1, \lfloor \mathrm{mod}(T(i), N) \rfloor + 1\right) = 1 \tag{6}$$

Where $i = 1,...,L_{cA}$. In Equation (5), cA is mapped into interval $[0, N \times N - 1]$ through a linear function, which is shown in Figure 3.

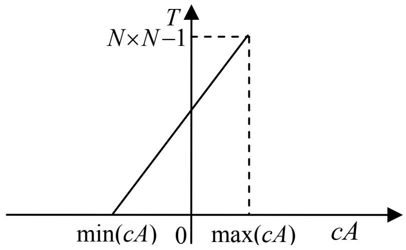


Figure 3. Linear mapping function

Step 4: In the steps 2 and 3, we can discover that the different permutations of cA^m will produce same mapping matrix C. Therefore, we introduce first-order difference of cA^m to get another binary matrix C'. Let DcA^m represents the first order difference of cA^m. Perform the above same Step 2 and 3 to map the DcA^m into another $N \times N$ matrix C'.

Step 5: Perform XOR operation on the C and C' to get a matrix M: $M = C \oplus C'$. Then, we can avoid the drawback of single mapping matrix.

Step 6: Implement Arnold transform on M to get a secure mapping matrix MM:

$$MM = Arnold\,(M, T) \tag{7}$$

Where T is number of iterations of Arnold transform. The purposes of Arnold transformation are scrambling the matrix M and enforcing security of the algorithm.

Step 7: Reverse the audio segment X to get another segment $X' = \{x_n, x_{n-1}, ..., x_1\}$. Transform the approximate coefficients of wavelet decomposition of X' into another mapping matrix $reMM_{N \times N}$ based the above Step 1~6. Here the time-frequency characteristics of wavelet transform are make full use to eliminate the deviation of audio segmentation under various attacks.

Step 8: Then, we get two mapping matrixes MM and $reMM$, which belong to the mapping matrix of forward X and reverse X' respectively. We name them as Double-direction (forward and reverse direction) Wavelet Coefficients Mapping Matrixes of an audio segment simply.

$$DWCMM = \{MM, reMM\} \tag{8}$$

4.2. Watermark embedding

Step 1: For an audio signal A, Divide A into q segments with based on BIC segmentation algorithm: $X_1, X_2, ..., X_q$.

Step 2: Extract the DWCM Matrix from each audio segment $X_i (1 \le i \le q)$ with the approach described in above part. Supposed the number of iterations of Arnold transform is T.

$$DWCMM_i = \{MM_i, reMM_i\} \tag{9}$$

Step 3: The key pairs $K_i = \{K_i^1, K_i^2\}$ of an audio segment are obtained as follows:

$$\begin{cases} K^1{}_i = MM_i \oplus W \\ K^2{}_i = reMM_i \oplus W \end{cases}, 1 \le i \le q. \tag{10}$$

The set of key pairs ($K = \{K_1, K_2, ..., K_q\}$) of entire audio is kept as the secret for detection. Finally, the original audio signal and the secret registered information $R = \{K, W, T\}$ are registered with an authentication center for copyright protection [17, 18].

4.3. Watermark detection

Without the help of original audio, the unknown audio A' will be detected blindly watermark to determine its affiliation. The steps are described as follows:

Step 1: Divide A' into non-overlapping p segments through based on BIC segmentation algorithm:

$X_1^{'}, X_2^{'}, ..., X_p^{'}$.

Step 2: Extract the DWCM Matrixes from each audio segment $X_i^{'}, 1 \le i \le p$.

$$DWCMM_i = \{MM_i, reMM_i\} \tag{11}$$

Step 3: With the key K and original image watermark W, the watermark $I_i^B (1 \le i \le p)$ of each segment can be gained as follows:

$$\begin{cases} I_{i,j}^1 = MM_i \oplus K_j^1 \\ I_{i,j}^2 = reMM_i \oplus K_j^2 \end{cases}, 1 \le i \le p, 1 \le j \le q \tag{12}$$

$$I_{i,j} = \begin{cases} I_{i,j}^1 & if\ BER(I_{i,j}^1, W) \le BER(I_{i,j}^2, W) \\ I_{i,j}^2 & if\ BER(I_{i,j}^1, W) > BER(I_{i,j}^2, W) \end{cases} \tag{13}$$

$$I_i^B = I_{i,j}^B, if\ BER(I_{i,j}^B, W) = \min_{j=1}^{q}(BER(I_{i,j}, W)), 1 \le i \le p \tag{14}$$

Step 4: Select the best quality watermark I of pending detection audio. I corresponds to the minimum of $BER(I_i^B, W), 1 \le i \le p$, which is obtained with

$$I = I_i^B, if\ BER(I_i^B, W) = \min_{i=1}^{p}(BER(I_i^B, W)) \tag{15}$$

Then, the binary watermarking image I is the watermark detected from the unknown audio signal.

5. Experimental results

In order to evaluate the performance of our watermarking scheme, various robustness experiments were performed. We evaluated the robustness of the watermark to common signal manipulations, such as low-pass filtering, re-quantization, re-sampling, white Gaussian noise addition, and MP3 compression, and geometrical attack (de-synchronization) processing including time scale modification, jitter attack, and specially large-area random cropping. Table 1 and 2 summarize the obtained robustness results against the attacks mentioned above, and Table 3 gives the comparative results with the other watermarking schemes.

An audio signal with length of 1949655 (177s) samples (16-bit signed mono audio file sampled at 11025 Hz with Wave format) and a 64×64 ($N = 64$) binary watermark image with Chinese Characters of Chongqing University are used to report the experimental results, shown in Figure 4 and Figure 5 respectively. In the experiment, the audio is performed 3-level wavelet decomposition with Daubechies-4 wavelet, and the value of number of iterations of Arnold transform is 15(T=15).

BER mentioned above is used to evaluate the similarity between original watermark and the detected watermark.

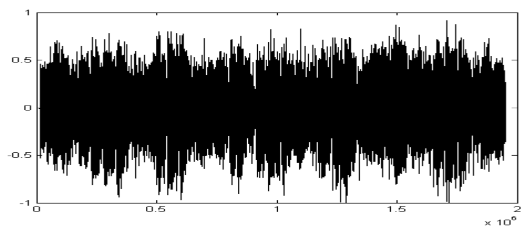


Figure 4. Original audio signal

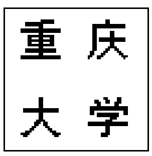


Figure 5. Binary watermark image

Table 1. Attacks and results

Attack Types	BER (%)	Detected Watermark
Without attack	0	重庆 大学
Resample: from 11025 Hz to 44.1 kHz, and back to 11025 Hz	0	重庆 大学
Re-quantization: from 16-bit to 8-bit, and then back to 16-bit	0	重庆 大学
Addition white Gaussian noise with 20 dB	6.1	重庆 大学
MP3 compression: from Wave format to MP3 with 64kbps, and back to Wave format	9.6	重庆 大学
Low-pass filter: Butterworth filter with cut-off frequency 1000Hz	6.3	重庆 大学
Time scale modification: time scale modification up to 150% (265s)	7.5	重庆 大学
Jitter attack: crop 10 samples interval 100 samples	6.6	重庆 大学

Table 2. Random cropping attacks

Cropping Types	BER (%)	Detected Watermark
Cut off 50% from the beginning of audio	0	重庆 大学
Cut off 90% from the beginning of audio	0	重庆 大学
Cut off 95% from the beginning of audio	0	重庆 大学
Cut off 50% from the end of audio	0	重庆 大学
Cut off 90% from the end of audio	0	重庆 大学
Cut off 95% from the end of audio	0	重庆 大学

Table 3. Comparison with other schemes ([13] and [16])
in terms of BRE(%) under various attacks

Cropping Types	Ours	[13]	[16]
Resample: from 11025 Hz to 44.1 kHz, and back to 11025 Hz	0	5.6	10.5
Addition white Gaussian noise with 20 dB	6.1	6.0	7.5
Cut off 10% from the beginning of audio	0	8.7	8.6
Cut off 50% from the end of audio	0	42.2	69.1
Cut off 90% from the end of audio	0	74.9	84.9
MP3 compression: from Wave format to MP3 with 64kbps, and the back to Wave format	9.6	6.8	7.1

From the experimental results, we can see, even the audio signal has a certain degree of damage and distortion under various attacks, the watermark still can be detected clearly. In particular, the proposed scheme has excellent performance in term of resisting attack of random large-area cropping. Therefore, we consider that the robustness of the audio zero-watermarking algorithm is better than other watermarking algorithm.

6. Conclusion

We have presented a novel approach for the audio watermark embedding and extraction based on audio segmentation and approximate wavelet coefficient mapping. Simulation results demonstrate that the proposed zero-watermark scheme has excellent performance in terms of security, robustness, anti-cropping attack. Furthermore, we make sure that the imbedded watermark can be detected in a tiny segment. In addition, we have to point out that we do not give a detail description to how to register watermark information in security and make a complete set of scheme of audio copyright protection. This is the work we will study later.

7. Acknowledgment

This work described in the paper was supported by National Natural Science Foundation of China (61103114), and the Fundamental Research Funds for the Central Universities under Grant NO. CDJXS11180009.

8. References

[1] M. Swanson, M. Kobayashi, A. Tewfik, "Multimedia Data Embedding and Watermarking Techniques", Proc. IEEE, vol. 86, no. 6, pp. 1064-1087, Jun. 1998.
[2] M.D. Swanson, B. Zhu, A.H. Tewfik, L. Boney, "Robust Audio Watermarking Using Perceptual Masking", Signal Process. vol. 66, no. 3, pp. 337–355, May 1998.
[3] M. Cettolo, M. Vescovi, "Efficient Audio Segmentation Algorithms Based on the BIC", ICASSP, vol. 6, pp. 537 - 540, 2003.
[4] Wen Quan, Sun Tan-Feng, Wang Shu-Xun, "Concept and Application of Zero-watermark", Tien Tzu Hsueh Pao/Acta Electronica Sinica, vol. 31, no. 2, pp. 214-216, 2003.
[5] Tanfeng Sun, Wen Quan, Shuxun Wang, "Zero-watermark Watermarking for Image Authentication", Proceedings of the Fourth IASTED International Conference Signal and Image Processing, pp. 503-508, 2002.
[6] Yuxin Peng, Chong-Wah Ngo, Cuihua Fang, st al, "Audio Similarity Measure by Graph Modeling and Matching", Proceedings of the 14th Annual ACM International Conference on Multimedia, pp. 603-606,2006.
[7] Shaoliang Meng, Wei Cao, Yixin Yan, Shengming Li, "Bit Replacement Audio Watermarking Using Stereo Signals", AISS, Vol. 2, No. 3, pp. 18 - 24, 2010.
[8] Jungyeop Kim , "Robust Vector Digital Watermarking Using Angles and a Random Table", AISS, Vol. 2, No. 4, pp. 79 - 90, 2010.
[9] X.-Y. Wang, H. Zhao, "A Novel Synchronization Invariant Audio Watermarking Scheme Based on DWT and DCT", IEEE Trans. Signal Processing, vol.54, no.12, pp. 4835–4840, Dec. 2006.
[10] S. Wu, J. Huang, D. Huang, Y.Q. Shi, "Efficiently Self-synchronized Audio Watermarking for Assured Audio Data Transmission", IEEE Trans. Broadcasting, vol.51, no.1, pp.69–76, Mar. 2005.
[11] I.-K. Yeo, H.J. Kim, "Modified Patchwork Algorithm: A Novel Audio Watermarking Scheme", IEEE Trans. Speech Audio Process, vol.11, no.4, pp. 381–386, July 2003.
[12] X. Li, H.H. Yu, "Transparent and Robust Audio Data Hiding in Cepstrum Domain", IEEE International Conference on Multimedia and Expo, vol. 1, 2000, pp. 397–400.
[13] H. Wang, Y. Sun, L. Lu; W. Shu, "Anti-cropping Synchronization Audio Digital Watermark Algorithm Based on Watermark Sequence Number", ICSP, vol. 4, 2006.
[14] Ning Chen, Jie Zhu, "A Robust Zero-Watermarking Algorithm for Audio", EURASIP Journal on Advances in Signal Processing, vol. 2008, Article ID 453580, 7 pages, 2008.
[15] Ali Hamadou, Xingming Sun, Lingyun Gao, Saeed A. Shah, "A Fragile Zero-Watermarking

Technique for Authentication of Relational Databases", JDCTA, Vol. 5, No. 5, pp. 189 ~ 200, 2011.
[16] Gang Chen, Qiping Hu, "Zero-Watermark of Audio Based on Multi-Scale Entropy Vectors", Computer Science and Information Technology, vol. 3, pp. 131-135, 2009.
[17] Chaoju Hu, Xuning Wang, "Zero-watermark Protocol Based on Time-stamp and Digital Signature", International Forum on Information Technology and Applications, vol.3, pp.193-196, 2009.
[18] D Augot, J- F Delaigle , C Fontaine. "DHWM: A Scheme for Managing Watermarking Keys in the Aquarelle Multimedia Distributed System", In: Proceedings of the 5th European Symposium on Research in Computer Security-(ESORICS98), Louvain -la –Neuve,Belgium, Lectures Notes in Computer Security, Springer-Verlag, pp.241-256, 1998.