Overview of Multifunctional Audio Watermarking based on Quantization

Haiyan Liu[1], Zenggang Xiong[2]

[1]*Department of Computer Engineering*, Beijing Information Technology College, Beijing,100018, China
[2]*School of Computer and Information Science*, Hubei Engineering University, Xiaogan 432000, Hubei, China

Abstract **— Multifunctional audio watermarking algorithms play an important role in the copyright protection and content authentication of digital media. Firstly, the paper summarized the method of generating the watermark information, focused on the multifunctional audio watermarking embedding algorithms based on quantization and analyzed their merits and demerits. And then, the tampering position detection schemes which are based on audio content and are based on tampering matrix were described in detail. Finally, the problems to be solved and the future research direction were discussed. The research on the problems will provide theoretical and technical supports for the copyright protection and content authentication.**

Keywords - Multifunctional audio watermarking; Quantization; Copyright protection; Content authentication

I. INTRODUCTION

In recent years, digital audio watermarking technologies have made great progress in terms of resisting conventional attacks and verifying audio content integrity. For example, in [1], Shijun Xiang et al proposed an audio watermarking algorithm that could resist stretching and local cutting attacks. In [2], Nakaya Shogo et al proposed an audio watermarking algorithm based on vector quantization index modulation. In [3], an adaptive robust audio blind watermarking algorithm using dithering quantization and SNR was proposed. In [4-6], fragile watermarking was embedded into the audio signal to verify the integrity of the audio content. However, the proposed algorithms can only solve copyright protection or the authentication of audio content. Some high-value audio signals require not only copyright protection but also content authentication. Theses actual requirements promote the generation of the multifunctional audio watermarking algorithms. The Multifunctional audio watermarking technologies can locate the tamper localization and judge the attack types besides robustness, imperceptibility, security. Therefore, multifunctional audio watermarking technologies have important value in the Internet age[7].

At present, the research of multifunctional audio watermarking embedding algorithms are mainly based on quantitative methods. Firstly, this paper outlines two methods that generate the watermark information, and then focuses on the multifunctional audio watermark embedding, detection algorithms based on quantization and tamper localization algorithms, and finally discusses the existing problems and makes recommendations for future research directions.

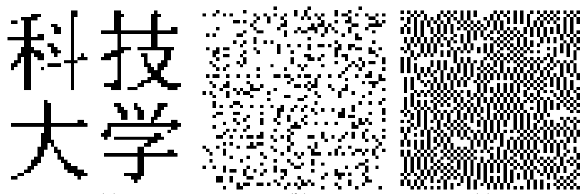
II. GENERATION OF WATERMARK INFORMATION

There are usually two ways to generate watermark information: (1) the meaningful binary images are adopted as the watermark; (2) the watermark is generated based on the audio content.

A. *Binary Image Watermark*

The Security of watermarking scheme is closely related with the watermark. To protect the security of scheme, the watermark is preprocessed with Arnold scrambling before it is embedded into the carriers. The scrambled watermark dispels the pixel space relationship of the binary watermark image, has high randomness, and can resist to a certain degree of decoding attack. When the attackers fail to obtain the watermark extraction algorithms and encryption keys, they can only extract the scrambled watermark. Therefore, the encryption method plays a protective effect to the embedded watermark and improves the security of watermarking schemes.

In [7], Haiyan liu et al used Arnold transformation and Logistic map to scramble the watermark. In [8-9], xiangyang wang, xiaodan lin et al used Arnold transformation to scramble the binary image. In [10], keqiang ren used Logistic map to encrypt watermark to dispel the pixel space relationship of the binary watermark image and improved the security and concealment. In figure 1, the figure(a)denotes the original binary image, the figure(b)is treated with Arnold scrambling, the figure(c) is processed with Logistic map.



(a) (b) (c)

Fig.1 Original Watermarking Image and Scrambled Watermarking Image

III. EMBEDDING ALGORITHM BASED ON QUANTIZATIO

A. Embedding Algorithm Diagram Based On Quantization

Firstly, the original audio signal is preprocessed to embed watermark easily; Secondly, the preprocessed audio signal is transformed in the transform domain, the appropriate quantization step and coefficients are selected to embed watermark using the quantization method; Finally, the audio single coefficients embedded watermark is took inverse transform and obtains a watermarked audio signal. The following Fig.2 is the diagram of the watermarking embedding scheme.

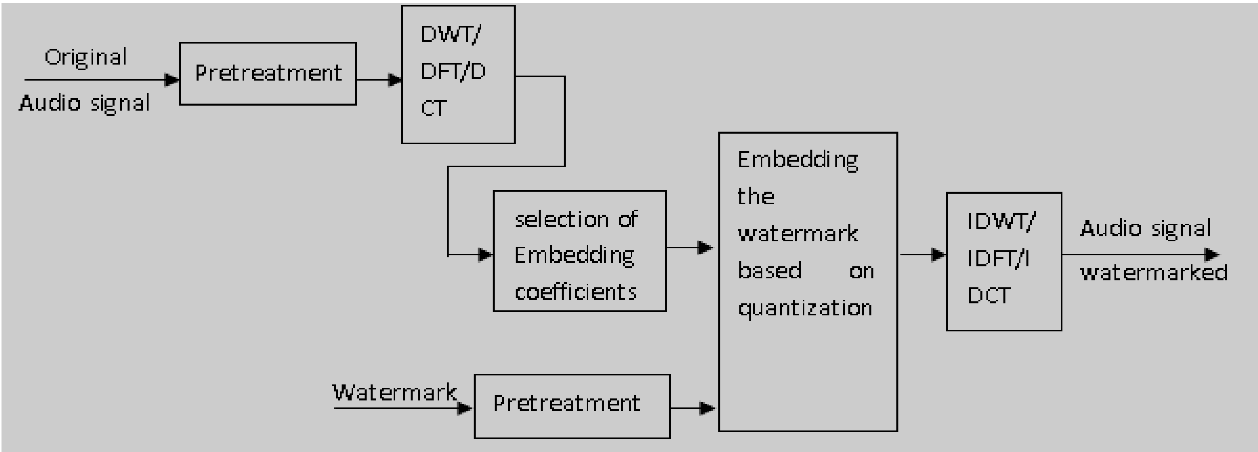


Fig.2 Embedding Watermark Algorithm Diagram

B. Embedding Algorithm Based On Quantization

(1) classic quantization watermarking algorithm

The principle of the classic quantization method is as follows.

The transformed audio coefficients maybe positive or negative, the principle that quantized coefficients hide the watermark is shown in Figure 3.

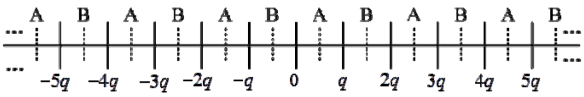


Fig.3 the Principle of the Classic Quantization Method Diagram

1)The axis of abscissa is divided into A interval set and B interval set according to the quantization step q . The coordinate values are expressed as"1" in A interval set, the coordinate values are expressed as "-1" in B interval set.

2)Using the following equation to evaluate the modulus and remainder of the coefficients $x(i,j)$:

$$m(i,j) = x(i,j) \bmod q \quad r(i,j) = x(i,j) - m(i,j) \cdot q$$

Where $m(i,j)$ expresses the modulus, $r(i,j)$ expresses the remainder.

3) Quantize $x(i,j)$ according to the interval sets of the watermark $w_f(i_1, j_1)$ and $x(i,j)$

If $w(i_1, j_1) = 1$, let the result of quantization $x'(i,j)$ take the middle coordinate value of a certain interval of proximate $x(i,j)$ in interval set A.

If $w(i_1, j_1) = -1$, let the result of quantization $x'(i,j)$ take the middle coordinate value of a certain interval of proximate $x(i,j)$ in interval set B.

In [12], random sequences have been used as the robust watermarking and embedded in the discrete cosine transform domain. Then the appropriate quantization has been done to produce the fragile watermarking and embedded in the discrete cosine transform domain. This method can not achieve blind extraction of robust watermarking and fragile watermark can not achieve the tamper localization.

(2) mean quantization method and single coefficient quantization method

In [7],[13], the mean quantization method is used to embed the robust watermark, the signal quantization method is adopted to embed the fragile watermark. The reason of using mean quantization method is mainly that the sample mean has a smaller variance than a single sample. Every watermark bit is embedded into mean of a series of coefficients by quantization methods, so that the effect of the original audio signal distortion on single watermark bit will be reduced greatly. The mean quantization method and single quantization method are as follows:

1)Mean Quantization Algorithm

Calculate the mean of each segment of audio signal:

$$ave(i) = \frac{1}{N_1}\sum_{j=1}^{n} x(i,j)$$

Suppose $z(i) = \lfloor ave(i)/q + 1/2 \rfloor$,where, q is the pre-set quantization step。

If $z(i)\%2 = w_f(i_1, j_1)$， then $ave'(i) = z(i) \times q$；

If $z(i)\%2 \neq w_f(i_1, j_1), z(i) = \lfloor ave(i)/q \rfloor$， then $ave'(i) = (z(i)+1) \times q$；

If $z(i)\%2 \neq w_f(i_1, j_1), z(i) \neq \lfloor ave(i)/q \rfloor$， then $ave'(i) = (z(i)-1) \times q$

Where % denotes the modular arithmetic, $\lfloor \ \rfloor$ denotes the down rounding operation.

The error will produce in the process of using mean quantization method. Suppose the error is u, namely, $u = ave(i) - ave'(i)$ .Therefore, the audio coefficients need to be modified, $x'(i,j) = x(i,j) + u$.

3) single coefficient quantization method

The single coefficient quantization method is similar to mean quantization, only needs to use $x(i,j)$ instead of $ave(i)$, quantization step q_1 instead of q .

(3) Improved quantization method

1) selection of quantification subject

Suppose α_i is the difference of the sum of the odd and the even coefficients, $\alpha_i = |S_i\_X_{odd} - S_i\_X_{even}|$, $S_i\_X_{odd}$ is the sum of odd coefficients, $S_i\_X_{even}$ is the sum of even coefficients, α_i is the quantification subject, obtain α_i' after quantizing the, $\alpha_i'' = (\alpha_i - \alpha_i')/\lceil L/2 \rceil$, L is the length of audio.

If $S_i\_X_{even} \geq S_i\_X_{odd}$, then $X'_{even}(i) = \{x'_{even}(i,j) = x_{even}(i,j) - \alpha'', 0 \leq j < \lceil L/2 \rceil\}$

If $S_i\_X_{even} < S_i\_X_{odd}$, then $X'_{even}(i) = \{x'_{even}(i,j) = x_{even}(i,j) + \alpha'', 0 \leq j < \lceil L/2 \rceil\}$

2) Improved quantization method

The robust watermarking algorithm selects the classical quantization method. Divide into region A: $[2kq, (2k+1)q)$ and region B: $[(2k+1)q, (2k+2)q)$ respectively according to the figure 4, let $m_i = \alpha_i \bmod q$, $r_i = \alpha_i - m_i \cdot q$.

If $w_f(i_1, j_1) = 1$ then

$$\alpha_i' = \begin{cases} (2k+0.5)q & if \quad m_i = 2k \\ (2k+0.5)q & if\ m_i = 2k+1 \quad and \quad |r_i| \leq 0.5q \\ (2k+2.5)q & if\ m_i = 2k+1 \quad and \quad |r_i| > 0.5q \end{cases}$$

If $w_f(i_1, j_1) = 0$， $m(i,j) \neq 0$ then

$$\alpha_i' = \begin{cases} (2k+1.5)q & if \quad m_i = 2k+1 \\ (2k-0.5)q & if \quad m_i = 2k \quad and \quad |r_i| \leq 0.5q \\ (2k+1.5)q & if \quad m_i = 2k \quad and \quad |r_i| > 0.5q \end{cases}$$

If $w_f(i_1, j_1) = 0$， $m(i,j) = 0$, then $\alpha_i' = 1.5q$

The fragile watermarking scheme is as follows.

If $w_f(i_1, j_1) = 1$,Then

$$\alpha_i' = \begin{cases} \alpha & if \quad m_i = 2k \\ (2k+1)q - \theta & if\ m_i = 2k+1 \quad and \quad |r_i| \leq 0.5q \\ (2k+2)q + \theta & if\ m_i = 2k+1 \quad and \quad |r_i| > 0.5q \end{cases}$$

If $w_f(i_1, j_1) = 0$, $m_i \neq 0$,Then

$$\alpha_i' = \begin{cases} \alpha & if \quad m_i = 2k+1 \\ 2kq - \theta & if \quad m_i = 2k \quad and \quad |r_i| \leq 0.5q \\ (2k+1)q + \theta & if \quad m_i = 2k \quad and \quad |r_i| > 0.5q \end{cases}$$

If $w_f(i_1, j_1) = 0$， $m_i = 0, \alpha_i' = q + \theta$

Where, θ is the embedding parameter, $0 < \theta < q$.

In [15], a zero-watermarking scheme system based on RBF neural network was constructed as robust watermarking .Since the scheme did not change the original audio data, it had a good transparency .The fragile watermarking embedded watermark image block by double bipolar, it had a good sensitivity .The experiment shows that the watermarks are robust and sensitive to many signal operations .Moreover, it has a good accuracy of tamper location.

In addition to the several typical quantization algorithms, N. Chen and J. Zhu[16]proposed a multifunctional speech watermarking algorithm based on the multistage vector quantization (MSVQ) of linear prediction coefficients (LPCs). The robust watermarking algorithm took use of the feature that the vector quantization (VQ) indices of the LPCs amongst neighboring frames tended to be very similar to embed the robust watermark in the indices of the first-stage VQ(VQ1). Then, the semi-fragile watermark was embedded in the indices of the second-stage VQ(VQ2) with index constrained VQ encoding scheme. Both the robust watermark and the semi-fragile watermark could be extracted blindly. Experiment results proved that the proposed algorithm in terms of robustness and semi-fragility was effective. However, the algorithm exists two problems: one is the product codebook size is too large to search for a long time, it is difficult to implementation of the algorithm; the other is that the LPC coefficient of speech is made vector quantization directly without introducing psychological acoustic model ,so the quality of speech signal watermarked is poor. In [17], the masking threshold was adopted as the quantization step and embedded the watermark in FFT(fast fourier transform) domain by the positive and negative modulation. The different detection process is used to copyright protection or tamper detection. The disadvantage of this method is that the blind detection can not be achieved and can not locate the tamper regions.

IV. WATERMARK EXTRATION

The extraction algorithm of multifunctional audio watermark is the inverse process of embedding algorithm. The audio signal tested is preprocessed and then is transformed in domain transform. The robust watermark and fragile watermark are extracted finally. The robust watermark is used to verify the capacity of algorithm resisting conventional attacks and malicious attacks, commonly used evaluation criteria have error rates. The authentication watermark is used to verify the integrity of the audio content and to locate tamper localization.

V. TAMPER DETECTION ALGORITHM

The fragile watermarking scheme can detect if the audio signal tested is tampered and may locate the tamper regions.

A. Tamper Matrix-based Authentication Method

Firstly, the watermark is extracted from the audio signal embedded watermark, and then the watermark extracted is compared with the original watermark to determine whether the audio signal embedded watermark is tampered. The information of the audio signal tampered is given including tamper type, tamper location if the signal is tampered when tamper matrix-based authentication algorithm is used to authenticate the integrity of digital audio. The steps of authentication algorithm is as follows[8]:

Assuming that the original audio signal is divided into i segments.

Step1: define the tamper matrix D_H : $\mathrm{D}_H = W_f' \oplus W_f$

Where, W_f expresses the original binary watermark image, $W_f{}'$ expresses the watermark image extracted, $\oplus$ expresses XOR operation. The segments of the pixels whose value are 1 express that the segments can be tampered in tamper matrix.

Step2: Eliminate noise D_H to eliminate the false alarm of malicious attacks.

Step3: Calculate the tamper radio of audio segments.

$$r(i) = \frac{\sum_{j_1=1}^{n_1} d_H(i, j_1)}{n_1}$$

Step4：Suppose detection threshold δ to judge whether audio segments tested is attacked maliciously.

TABLE I. TAMPER AUTHENTICATION RULES

Condition	Result
$r(i) = 0$	Audio segments are not attacked.
$r(i) \leq \delta$	Audio segments have been attacked by conventional attacks
$r(i) > \delta$	Audio segments have been attacked by conventional attacks

Xiangyang wang[8] et al uses such method, the extracted watermark image is compared with the original watermark image to determine the tampering locations. The attack types are determined according to the ratio of black dots and white dots of dense point number in difference map generated by the tamper matrix D_H . However, the distribution of black dots and white dots is each more concentrative and the distribution ways is different for different watermarking image, so it is difficult to find a suitable sparse point definition, which affects the accuracy of the tamper localization. Meanwhile, in the process of judgment, the detector needs to set a non-negative threshold value δ and the size of δ determines the stringency and accuracy of the type of attack.

(2)audio content-based tamper authentication method

The audio content-based tamper authentication method is that the content or features of the audio signal as watermark are embedded into the original audio signal. The watermark extracted is compared with the audio content or features tested without additionally providing original watermark when authenticating.

In [11] and [18], the second method is adopted. In [11], the authentication sequences is defined $T = W' \oplus W^*$ (W' is the watermark extracted from the tested audio, W^* is the reconstruction watermark) and then the T is divided into M groups, let $T_A(p) = \begin{cases} 1 & S(p) \neq 0 \\ 0 & S(p) = 0 \end{cases}$ ($S(p)$ is the sum of elements in each group), if $T_A(p) = 1$,then audio content in the pth frame is tampered ,else, the audio content is not tampered.

VI. CONCLUSION AND PROSPECT

In summary, the multifunction digital audio watermarking algorithms based on quantization have solved the following issues: (1) most of algorithms have taken into account the human auditory system, so the imperceptibility of algorithms have been improved; (2) The malicious tampering can be detected effectively and the tamper localization can be located; (3) The schemes have the capacity of copyright protection and audio content integrity authentication. However, the multifunctional audio watermarking algorithms need to hide a large amount of information generally, which needs a better utilization of characteristic of the human visual system (HVS) and human auditory system (HAS). However the human visual system is more sensitive than the human auditory system, the research of multifunctional audio watermarking algorithm faces more challenges. there are many practical problems to solve. According to the characteristics of audio signal and the actual needs of the audio watermarking technology, the following issues should be further gone into:

(1) The ability of resistance synchronization attacks such as TSM, Random Cropping, up and down sampling is poor in general, therefore, how to further improve the robustness of multifunctional audio watermark schemes is an interesting question;

(2) How to avoid mutual interference between robust watermarking and fragile watermarking needs to be solved;

(3) The embedding amount of the original audio is increased because of embedding multiple watermarks, how to balance the embedding amount, robustness and imperceptibility needs to be solved;

(4) The audio watermarking algorithms which have repair function are the research emphasis. How to use the differences between extracted watermark and the original watermark to recover and reconstruct the original audio content will be the research emphasis of the multifunctional audio watermarking algorithms.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflicts of interest.

ACKNOWLEDGMENT

The National Natural Science Foundation of China(No. 61370092), Science and Technology Project of Beijing Municipal Commission of Education Natural Science(No.KM201310857001), Foundation of Hubei Province(No.2013CFC005, 2014CFB188), Hubei Provincial Department of Education Outstanding Yo uth Scientific Innovation Team Support Foundation.

REFERENCES

[1] Shijun Xiang, Yongjin Huo, Shangyi Liu, Xinrong Luo, " Audio Watermarking against Synchronization Attacks Using Statistical Features," Journal of Applied Sciences—Electronics and Information Engineering, Shanghai, China, vol. 4, pp.936-940.2014

[2] Nakaya Shogo, Wada Shigeo, "Audio Watermarking based on Vector Quantization Index Modulation using Audio Fingerprint," Electronics and Communications in Japan, Japan, vol.3, pp. 47-56,2015.

[3] Xionghua Huang, Hongxia Wang, Weizhen Jiang, Gengshen Cui, " An Adaptive Robust Audio Blind Watermarking Algorithm Using Dithering Quantization and SNR, " Geomatics and Information Science of Wuhan University, Wuhan, China, vol.38, pp. 121-125,2013.

[4] Yun Zhou, Rangding Wang, "Fragile watermarking scheme based on index constrained vector quantization," Application Research of Computers , Chengdu, China, vol. 28, pp. 1940-1942, 2011.

[5] Haiyan Liu, Xuefeng Zheng, Ying Wang, "A Fragile Digital Audio Watermarking Algorithm based on SMM," 7th World Congress on Intelligent Control and Automation, Chongqing, China, pp.8409 – 8413, 2008.

[6] Jinquan Zhang, Bin Han, "Fragile Audio Watermarking Scheme based on Sample Mean Sequence," 2011 International Conference on Multimedia Technology (ICMT), Hangzhou, China, pp.333–336, 2011.

[7] Haiyan Liu, Zenggang Xiong, Yi Liu, "A multifunctional Audio Watermarking Algorithm for Copyright Protection and Content Authentication", Journal of Beijing University of Technology, Beijing, China, vol.41, pp. 68 – 75, 2014.

[8] Xiangyang Wang, Qi Wei,"A Semi-fragile Audio Watermarking for Copyright Protection and Content Authentication", Acta Automatica Sinica, Beijing, China, vol. 33, pp.936 – 940, 2007.

[9] Xiaodan Lin, "Audio Watermarking for Copyright Verification and Content Authentication", Journal of Computer Applications, Chengdu, China, vol. 30, pp. 2269 – 2271, 2010.

[10] Keqiang Ren, Hui Li, Bin Xie, "Adaptive dual Audio Watermarking based on DWT AND DCT," Application Research of Computer, Chengdu, China, vol. 30, pp. 2120 – 2123,2013.

[11] Hongxia Wang, Mingquan Fan, "Centroid–based Sem-fragile Audio Watermarking Algorithm Inhybrid Domain", Sciece China Information Sciences, Beijing, China, vol. 40, pp.313 – 326, 2010

[12] Yang Yang, Xiaoping Chen, "Digital Audio Multi-Watermarking Algorithm based on Discrete Cosine Transform", Journal of Data Acquisition & Processing, Nanjing, China, vol. 19, pp.215-219, 2006.

[13] Rangding Wang, Dawen Xu, "Multiple audio watermarks based on lifting wavelet transform," Journal of Electronics & Information Technology, Beijing, China, vol.28,pp. 1820 – 1826,2006.

[14] Wanming Liao, Yuxian Zhang, Dongxiao Li, et al, "Audio fragile-robust dual watermarking scheme based on digital wavelet transform", Journal o f Zhejiang University (Engineering Science), Zhengjiang, China, vol. 43, pp. 721 – 726.

[15] Chaoyang Ma, Xueying Zhang, Wenlu Li, "A Digital Audio Dual Watermarking Scheme Used for Tamper Location ", Journal of Taiyuan University of Technology, taiyan,China, vol. 42, pp. 107 – 110, 2011.

[16] N. Chen, J. Zhu, "Multipurpose Speech Watermarking based on Multistage Vector Quantization of Linear Prediction Coefficients", The Journal of China Universities of Posts and Telecommunications, Beijing, China, vol. 14, pp. 64 – 69, 2007.

[17] CHEN Ning, ZHU Jie, "Multipurpose audio watermarking algorithm", Journal of Zhejiang University Science A, Zhejiang, China, vol. 2, pp. 517– 523, 2008.

[18] Chaokui Ning, Hongjie He, fan Chen et al, "Semi-fragile Audio Watermarking Scheme Based on the Approximate Components Energy", Journal of the China Railway Society, Beijing,China,vol.35, pp. 46 – 50, 2013.