Study on Protection of Genuine Digital Music Works by a Watermarking Algorithm

Juan Gao and Baike Zuo
(Corresponding author: Baike Zuo)

School of Music, Hebei Institute of Communications
Shijiazhuang 050073,China
Email: bk98493@163.com

(Received Apr. 9, 2022; Revised and Accepted Mar. 23, 2023; First Online Apr. 27, 2023)

Abstract

The protection of genuine digital music works is critical. In this paper, a watermarking algorithm was designed based on the technique of discrete wavelet transform and vector norms to achieve the protection of music works. The experiments on different types of digital music works showed that the watermarking algorithm designed in this paper had good security, a high mean opinion score and signal-to-noise ratio, and good imperceptibility, and it kept a high normalized coefficient and low bit error rate in the face of noise addition and resampling attacks and had an embedding capacity greater than 20 bps. The results demonstrate the effectiveness of the designed watermarking algorithm in protecting genuine digital music works.

Keywords: Digital Music; Genuine Protection; Robustness; Watermarking Algorithm

1 Introduction

With the continuous development of computer and Internet, traditional compact discs and tapes are gradually fading out of the market, music is getting closer and closer to digitalization, and the Internet is gradually becoming the way people enjoy music works, which, at the same time, brings the problem of rampant piracy [12]. In order to achieve protection of genuine digital music works, digital right management (DRM) has emerged [31], but the great inconvenience of DRM is not conducive to its practical promotion [8, 26]. Digital watermarking has a wide range of applications in digital media [6, 7, 13, 15] and plays a good role in the protection of genuine images and videos [23], and it can also be applied to audio signals.

Watermarking algorithms for audio signals has become a key issue for researchers to focus on [10]. Mosleh *et al.* [19] proposed a watermarking algorithm based on discrete cosine transformation (DCT) and LU decomposition and found through experiments that it had a good performance. Mohammed *et al.* [18] used discrete wavelet transform (DWT) to decompose the signal and then used DCT to encrypt the watermarked image. They found that the DCT method was not easy to be detected and the average signal-to-noise ratio (SNR) reached 61 dB.

Safitri *et al.* [22] proposed a watermark embedding method combining compressed sampling, DWT, and QR decomposition to embed the watermark into audio by quantization index modulation (QIM) and found through experiments that the SNR of the method was greater than 20 dB, indicating good robustness. Dronyuk *et al.* [9] designed a digital watermarking method based on generalized Fourier, Hartley transform, and Ateb function and verified the stability of the method in supporting the security of audio and image. This paper designed a watermarking algorithm based on DWT, verified the security and robustness of the algorithm through experimental analysis, and proved the reliability of the algorithm for genuine digital music protection. The designed watermarking algorithm can be applied in actual music works. This work provides a new method for the protection of audio signals.

2 Protection of Genuine Digital Music Works

DRM realizes the protection of genuine works through controlling the right to use the works. The technology strictly manages the playing and copying of works but brings greater inconvenience to the actual appreciation of music works [21]. Music works protected by DRM have a strong exclusivity, which restricts the use environment of the player and affects the experience of enjoying the works, which is not conducive to the promotion of the technology.

Digital watermarking is a method to protect works by embedding watermark information without changing the content of the original work [1]. Since the perception of the human ear is more sensitive than that of the human eye [20], modifications to the audio signal can be easily perceived by the human ear, so an excellent watermarking algorithm needs not only to have good resistance to attacks but also to be able to guarantee the quality of the musical work. The evaluation of watermarking algorithms includes the following three main aspects.

The subjective evaluation of perceptibility is based on the mean opinion score (MOS) [4], which refers to the listener's subjective perception of the music piece after embedding the watermark. The evaluation criteria of MOS are shown in Table 1.

Table 1: MOS evaluation criteria

Score	Audio indicators	Description
5	Excellent	Imperceptible
4	Good	Slightly perceptible
3	Medium	Perceptible, slightly unpleasant
2	Poor	Obviously perceptible, but tolerable
1	Very poor	Unbearable

The objective evaluation of perceptibility is based on the peak signal-to-noise ratio (PSNR) [27], which is a measure of the quality of the music piece after embedding the watermark. Let the audio signal before and after embedding the watermark be $x(n)$ and $x'(n)$ and the audio length be L. The PSNR is calculated by:

$$PSNR = 10\log_{10}(\frac{\max_{0\leq n\leq L}\{x^2(n)\}}{\sum_{n=1}^{L}[x'(n)-x(n)]^2}$$

The robustness is generally evaluated using the bit error rate (BER) [29] and the normalized coefficient (NC) [2]. The calculation formulas of BER and NC are:

$$BER = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}A(i,j)\oplus B(i,j)}{M\times N}$$

$$NC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}A(i,j)B(i,j)}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}A^2(i,j)}\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{N}B^2(i,j)}}$$

where A is the original watermark, B is the extracted watermark, $\oplus$ is the exclusive or operation, and M and N are the row and column of the watermark signal.

The unit of embedding capacity is bps. Let the watermark size be N and the length of audio signal be K. The embedding capacity is:

$$payload = \frac{N}{K}$$

According to the International Federation of the Phonographic Industry (IFPI) [16], this value needs to be greater than or equal to 20 bps.

3 Watermarking Algorithm Based on Vector Norms and Wavelet Transform

3.1 Watermark Pre-processing

A binary image is used as a watermark information for genuine digital music protection. It is preprocessed first in order to improve security. Arnold transform [24] is a method to achieve encryption by scrambling the coordinates of pixel points, defined as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \bmod N,$$

where x and y are the pixel coordinates before scrambling, x' and y' are the pixel coordinates after scrambling, and N is the row or column width of the image.

After scrambling, the watermarked signal is reduced dimensionally. For a $N\times M$ watermarked image W, its one-dimensional binary sequence is $w'(x,y)$. The dimensionality reduction process is written as:

$$V = \{v(k)=w'(x,y), \quad 0<i\leq N, 0<j\leq M, k=i\times M+j\}.$$

For a one-dimensional watermark, the encryption is performed using logistic mapping [28], and the relevant formula is:

$$x_(n+1) = \mu x_n(1-x_n),$$

where $x_n \in [0,1]$ and $\mu\in[0,4]$. When $3.569945<\mu\leq 4$, the system is chaotic. Let $\mu = 3.8$ and initial value $x_0 = 0.6$, the system is in a chaotic state. Sequence x_n is obtained according to the above formula, and chaotic sequence μ_k is obtained after quantification. Then, it is processed by exclusive or along with $v(k)$ to obtain the watermark sequence after secondary encryption:

$$W'' = v(k)\oplus\mu_k, 1\leq i\leq N\times M.$$

3.2 Watermark Embedding

The principle of the watermarking algorithm designed in this paper is to firstly DWT the original audio signal, write down the low frequency coefficients in it as a vector, and then combine the vector parametrization to achieve the embedding of the watermark. The two theories involved in the process are as follows.

1) Vector norm [14]: For vector $A=(a_1,a_2,\cdots,a_n)$, its P-norm is ρ, defined as $\rho = A_P = (\sum_{i=1}^{n}|a_i|^P)^{1/P}$.

2) Wavelet transform: It is a method of signal processing [11], which solves the shortage of Fourier transform in dealing with abrupt signals [25] and can extract more useful information. It has good performance in signal processing [30], image processing [5], etc. Suppose there is square productable

function $\psi(t)$, $\psi(t) \in L^2(R)$, whose Fourier transform $\psi(\omega)$ satisfies $\int \frac{|\psi(\omega)|^2}{\omega} d\omega < \infty$, then the equation is called the admissible condition. Let the scale factor of $\psi(t)$ be α and the translation factor be τ, then $\psi(t)$ after translational expansion is written as: $\psi_{\alpha,\tau}(t) = \alpha^{-1/2}\psi(\frac{t-\tau}{\alpha}$, $a > 0$, $\tau \in R$, and $\psi_{\alpha,\tau}(t)$ is the non-interrupted wavelet basis function. Since most of the signals in practice are discretized digital signals, DWT is more commonly used [3].

α and τ are discretized: $\alpha = 2^j$, $\tau = 2^k T_s$, and then the discrete wavelet function is obtained: $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$. For $f(t)$, its discrete wavelet function is: $WT_f(j,k) = (f, \psi_{j,k}) = \int_R f(t)\bar{\psi_{j,k}}(t)dt$. After the signal is processed by DWT, the energy is mainly concentrated in the low-frequency component, so the watermark information can be embedded into it.

The process of watermark embedding based on vector norms and DWT is as follows.

1) The watermark is divided into $N \times M$ frames. Two-stage DWT processing is performed on the audio signal to get a low-frequency component cA_2 and two high-frequency components cD_1 and cD_2.

2) The low-frequency coefficient is denoted as vector D, which is divided into vectors $V1$ and $V2$: $V1 = \{D(i), 1 \leq i \leq L_c/2\}$, $V2 = \{D(i), (L_c/2)+1 \leq i \leq L_c\}$, where L_c is the length of cA_2 and cD_2, $L_c = \frac{L}{N \times M \times 2^2}$ (L is the length of the audio signal).

3) The 2-norm of $V1$ and $V2$, i.e., $Norm_{V1}$ and $Norm_{V2}$, are calculated. Average value Norm is calculated. When the watermark bit to be embedded W'' is 1, the watermark is embedded according to the formula

$$\begin{aligned} Norm_{V1} &= Norm + q \\ Norm_{V2} &= Norm - q, \end{aligned}$$

when the watermark bit to be embedded W'' is 0, the watermark is embedded according to the formula

$$\begin{aligned} Norm_{V1} &= Norm - q \\ Norm_{V2} &= Norm + q, \end{aligned}$$

were q is the adjustable quantization strength, 0.03 here.

4) Vectors $V1'$ and $V2'$ are reconstructed using the revised norms, and the results are combined to obtain vector D'.

5) Inverse discrete wavelet transform (IDWT) is performed to obtain a frame of audio signal after embedding the watermark. The above steps are repeated until all watermark bits are embedded.

3.3 Watermark Extraction

The watermark extraction process is as follows.

1) The audio is divided into $N \times M$ frames for two-stage DWT processing. The low-frequency vector is denoted as D', which is divided into vectors $V1'$ and $V2'$.

2) The vector norms of $V1'$ and $V2'$, i.e., $Norm_{V1}$ and $Norm_{V2}$, are calculated. If $Norm_{V1} > Norm_{V2}$, then the watermark bit is 1; otherwise, it is 0.

3) Arnold transform and logistic mapping are used for decryption to get the watermark information.

4 Results and Analysis

Experiments were carried out in MATLAB2018 environment. Five different types of digital music (rock, pop, blues, classical, and jazz) were randomly selected from the network music library for experiments, and they were all wav format and monophonic. The sampling frequency was 44.1 kHz, 16 bit. A piece of 5 s was taken from every music as experimental audio signals. In the case of correct and incorrect keys, the extracted watermarked images are shown in Figures 1-3.

Figure 1: Original watermarked image

Figure 2: The extracted watermark image in the case of key error

Figure 3: The extracted watermark image when the key is correct

It was seen from the comparison between Figures 1-3 that when the watermark was embedded using the watermarking algorithm designed in this paper, the correct watermarked image was not obtained in the case of incorrect key, while the complete watermarked image was extracted when the key was correct.

Ten students with normal hearing functions were selected as listeners to make MOS evaluation on the musics in a quiet environment, and their SNRs were calculated. The results are shown in Figure 4.

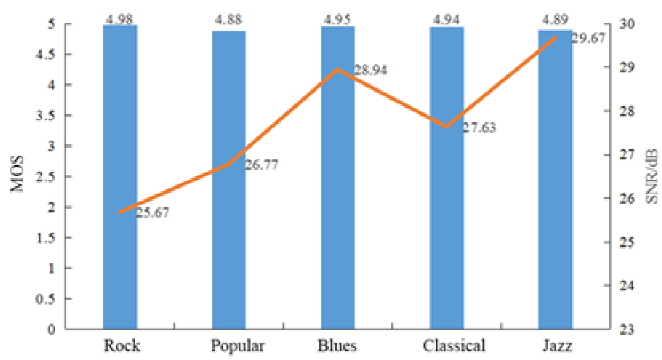

Figure 4: Imperceptibility evaluation results

It was seen from Figure 4 that, for these five types of music works, the MOS was always above 4.5 after embedding the watermark, with an average score of 4.93, indicating that the impact of embedding the watermark on the quality of music works was very small. Then, in terms of the objective perception, the SNR of different music genres was all above 20 dB after embedding the watermark, which was in line with the standard of greater than 20 dB as stipulated by IFPI.

The following attacks were performed on the watermarked audios.

1) Noise addition: Gaussian white noise with an expectation value of 0.01 and a variance of 0.05 was added.

2) Low-pass filter: a low-pass filter with six orders and a cut-off frequency of 10 kHz was used.

3) Requantization: the audio resolution was quantized from 16 bit to 8 bit and from 8 bit to 16 bit.

4) MP3 compression: the audios were converted from wav format to mp3 format and from map format to wav format.

5) Resampling: the audios were resampled using 22.05 kHz and then 44.1kHz.

6) Random cropping: ten locations were randomly selected to cut and remove 200 sampling points. The BER and NC of different types of music works under different attacks are listed in Table 2.

It was seen from Table 2 that the NC and BER of the audios was 1 and 0, respectively, under no attack. Different types of music works always maintained low BER (below 0.1) and high NC (above 0.9) under different attacks, indicating that the audios were less affected by attacks. The BER and NC values of different types of musical works were averaged, and the results are shown in Figure 5.

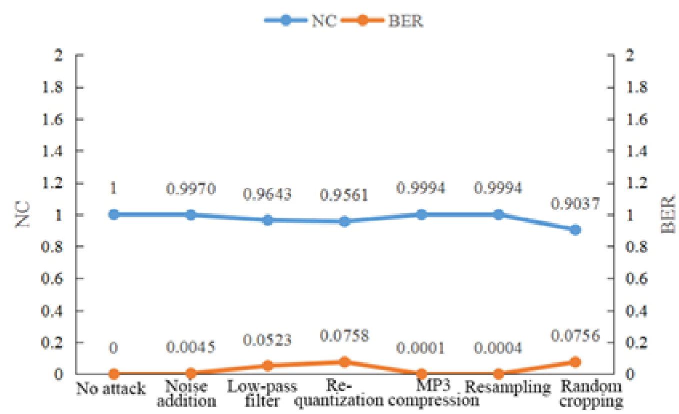

Figure 5: Results of robustness analysis of the watermarking algorithm

It was seen from Figure 5 that the BER of the audios was small under different attacks, the BER of the audios under random cropping was the largest, 0.0756, and the BER under MP3 compression was the smallest, 0.0001, which indicated that the watermarking algorithm designed in this paper maintained a low BER under different attacks. Then, in terms of NC, it was found that the NC of the audios under different attacks was very close to 1, with a maximum of 0.9994 and a minimum of 0.9037, both above 0.9, indicating that the designed watermarking algorithm maintained a high NC under different attacks. Finally, the embedding capacity was analyzed. The embedding capacity of the watermark was 87.64 bps after calculation, which met the IFPI standard–at least 20 bps.

Table 2: BER and NC of different music works

	NC/ BER	No attack	Noise addition	Low-pass filter	Requantization	MP3 compression	Resampling	Random cropping
Rock	NC	1	0.9964	0.9652	0.9568	0.9998	0.9995	0.9021
	BER	0	0.0044	0.0524	0.0763	0.0001	0.0004	0.0754
Popular	NC	1	0.9956	0.9646	0.9525	0.9998	0.9996	0.9056
	BER	0	0.0043	0.0526	0.0756	0.0001	0.0003	0.0764
Blues	NC	1	0.9974	0.9626	0.9556	0.9989	0.9989	0.9025
	BER	0	0.0051	0.0512	0.0758	0.0001	0.0004	0.0755
Classical	NC	1	0.9986	0.9646	0.9578	0.9997	0.9996	0.9056
	BER	0	0.0043	0.0525	0.0765	0.0001	0.0005	0.0752
Jazz	NC	1	0.9969	0.9646	0.9578	0.9989	0.9996	0.9025
	BER	0	0.0042	0.0526	0.0749	0.0001	0.0004	0.0754

5 Conclusion

This paper mainly studied the protection of genuine digital music works, designed a watermarking algorithm based on norms and DWT, and conducted experiments on actual music works. It was found that the designed watermarking algorithm had good security, high MOS, SNR above 20 dB, and good imperceptibility, and it maintained high NC and low BER even under different attacks, suggesting good robustness, and the embedding capacity also met the demand of digital music genuine protection. The designed watermarking algorithm can be promoted and applied in practice.

References

[1] Z. Al-Husseiny, "Using discrete wavelet transformation algorithm for authentication digital image watermark," *Journal of Engineering and Applied Sciences*, vol. 13, no. 11, pp. 4001-4008, 2018.

[2] M. H. Annaby, S. H. Basha, Y. M. Fouda, "Defect detection methods using boolean functions and the φ-coefficient between bit-plane slices," *Optics and Lasers in Engineering*, vol. 139, no. 6, pp. 106474, 2020.

[3] R. Arvanaghi, S. Danishvar, M. Danishvar, "Classification cardiac beats using arterial blood pressure signal based on discrete wavelet transform and deep convolutional neural network," *Biomedical Signal Processing and Control*, vol. 71, pp. 1-8, 2022.

[4] M. Barazzetta, M. Colombo, L. Bastianelli, F. Moglie, V. Mariani Primiani, R. Diamanti, D. Micheli, "Testing of VoLTE mean opinion score in reverberation chambers," *IET Science Measurement & Technology*, vol. 14, no. 8, pp. 949-954, 2020.

[5] H. K. Bhagya, N. Keshaveni, "Contrast enhancement technique using discrete wavelet transform with just noticeable difference model for 3D stereoscopic degraded video," *International Journal of Innovative Technology and Exploring Engineering*, vol. 10, no. 3, pp. 7-13, 2021.

[6] C. C. Chang, K. F. Hwang, M. S. Hwang, "A digital watermarking scheme using human visual effects", *Informatics*, vol. 24, no. 4, pp. 505–511, Dec. 2000.

[7] C. C. Chang, K. F. Hwang, M. S. Hwang, "A block based digital watermarks for copy protection of images," in *Asia-Pacific Conference on Communications*, vol. 2, pp. 977-980, 1999.

[8] Y. Chou, K. Anggriani, N. Wu, M. S. Hwang, "Research on e-book text copyright protection and anti-tampering technology", *International Journal of Network Security*, vol. 23, no. 5, pp. 739-749, 2021.

[9] I. Dronyuk, O. Fedevych, N. Kryvinska, "Constructing of digital watermark based on generalized fourier transform," *Electronics*, vol. 9, no. 7, pp. 1-13, 2020.

[10] R. R. Ginanjar, D. S. Kim, C. B. Moon, "High-capacity and transparent digital audio watermarking using rounding reduced-arc MPSK and a genetic algorithm," *IEIE Transactions on Smart Processing and Computing*, vol. 8, no. 1, pp. 49-57, 2019.

[11] M. Gogowski, S. Osowski, "Classical versus deep learning methods for anomaly detection in ECG using wavelet transformation," *Przeglad Elektrotechniczny*, vol. 97, no. 6, pp. 72-76, 2021.

[12] L. Guo, A. L. Wang, S. K. Li, J. P. Pang, K. Xin, S. X. Fan, F. P. Liu, "Authentication algorithm of secure digital halftone watermark based on SM4 algorithm," *IOP Conference Series: Materials Science and Engineering*, vol. 563, no. 5, pp. 1-5, 2019.

[13] M. S. Hwang, C. C. Chang, K. F. Hwang, "Digital watermarking of images using neural networks", *Journal of Electronic Imaging*, vol. 9, no. 4, pp. 548–555, Jan. 2000.

[14] T. Ke, L. Zhang, X. Ge, H. Lv, M. Li, "Construct a robust least squares support vector machine based on Lp-norm and L∞-norm," *Engineering Applications of Artificial Intelligence*, vol. 99, no. 3, pp. 104134, 2021.

[15] V. Korzhik, D. Flaksman, "Digital watermark system with an ability of its extraction from hard copies

of data," *Proceedings of Telecommunication Universities*, vol. 5, no. 3, pp. 75-85, 2019.

[16] Laing D. , "International federation of the phonographic industry

[17] . Moneyscience, 2009: 585-594.

[18] R. Mohammed, M. E. Abdulmunim, "Using Rubik's cube in fragile audio watermark encryption," *Diyala Journal of Engineering Sciences*, vol. 15, no. 3, pp. 103-124, 2020.

[19] M. Mosleh, S. Setayeshi, B. Barekatain, M. Mosleh, "High-capacity, transparent and robust audio watermarking based on synergy between DCT transform and LU decomposition using genetic algorithm," *Analog Integrated Circuits and Signal Processing*, vol. 100, no. 3, pp. 513-525, 2019.

[20] W. Mynarski, J. H. Mcdermott, "Ecological origins of perceptual grouping principles in the auditory system," *Proceedings of the National Academy of Sciences*, vol. 116, no. 50, pp. 25355-25364, 2019.

[21] S. Oliver, T. Winarta, "The lack of enforcement of the DRM policy," *Journal of Applied Information Communication and Technology*, vol. 7, no. 1, pp. 15-25, 2021.

[22] I. Safitri, G. Budiman, A. Putri, "Audio watermarking combined with compressive sampling based on QIM and DST-QR techniques," *Jurnal Elektronika dan Telekomunikasi*, vol. 19, no. 1, pp. 20, 2019.

[23] R. Sailaja, C. Rupa, A. Chakravarthy, "EPNN based high secure intensive hidden digital watermark application in telemedicine," *Advances in Modelling and Analysis*, vol. 56, no. 1, pp. 21-25, 2019.

[24] A. Shrivastava, J. Sharma, S. Purohit, "Image encryption based on fractional wavelet transform, arnold transform with double random phases in the HSV color domain," *Recent Advances in Computer Science and Communications*, vol. 15, no. 1, pp. 5-13, 2022.

[25] A. K. Singh, A. Saxena, N. Roy, U. Choudhury, "Inter-turn fault stability enrichment and diagnostic analysis of power system network using wavelet transformation-based sample data control and fuzzy logic controller," *Transactions of the Institute of Measurement and Control*, vol. 43, no. 12, pp. 2788-2798, 2021.

[26] C. Y. Tsai, C. Y. Yang, I. C. Lin, M. S. Hwang, "A survey of e-book digital right management", *International Journal of Network Security*, vol. 20, no. 5, pp. 998-1004, 2018.

[27] B. J. Tyler, R. Kassenböhmer, R. E. Peterson, D. T. Nguyen, M. Freitag, F. Glorius, B. J. Ravoo, "Denoising of mass spectrometry images via inverse maximum signal factors analysis," *Analytical Chemistry*, vol. 94, no. 6, pp. 2835-2843, 2022.

[28] J. Wang, K. Han, S. Fan, Y. Zhang, H. Tan, G. Joen, Y. Pang, J. Lin, "A logistic mapping-based encryption scheme for wireless body area networks," *Future Generation Computer Systems*, vol. 110, pp. 57-67, 2020.

[29] A. K. Yadav, P. K. Sahoo, Y. K. Prajapati, "Peak-to-average power ratio reduction and improved bit error rate performance of orthogonal frequency division multiplexing system using discrete Fourier transform precoder and root-based nonlinear companding," *Optical Engineering*, vol. 58, no. 7, pp. 076106.1-076106.9, 2019.

[30] L. Zamikhovsky, O. Zamikhovska, V. Pavlyk, "Research of the characteristics of acoustic processes using wavelet transformation for detecting a diagnostic sign of the technical state of gas pumping units," *Technology Audit and Production Reserves*, vol. 1, no. 2(57), pp. 32-36, 2021.

[31] X. Zhu, M. Cho, "The end of ownership?: An investigation of users' preferences and perceptions of ownership configurations," *Proceedings of the Association for Information Science & Technology*, vol. 55, no. 1, pp. 618-627, 2018.

Biography

Juan Gao, lecturer, master of arts, graduated from Guizhou University in 2010 and currently works at Hebei Institute of Communications. Her main research areas are music education and music aesthetics.

Baike Zuo, lecturer, master of education, graduated from Hebei University in 2017 and currently works at Hebei Institute of Communications. His main research areas are education management and music performance.