A Novel SVD and LWT Based Robust Blind Audio Watermarking Scheme

Wenliang Wu
(Corresponding author: Wenliang Wu)

Department of Basic Courses, Northern Beijing Vocational Education Institute
Beijing 101400, China
(Email: wuwenlig1982doc@sina.com)

(Received Mar. 19, 2018; Revised and Accepted Aug. 12, 2018; First Online Mar. 17, 2019)

Abstract

To efficiently protect copyright of digital audio products against illegal usage, in this paper, we present a robust, secure and Blind Audio Watermarking Algorithm based on Singular Value Decomposition (SVD) and Lifting Wavelet Transform (LWT) domain synchronization code, called BAWA-SL. More specifically, the synchronization code is embedded into the audio by leveraging the Quantization Index Modulation (QIM) to achieve blind extraction of watermarking. Furthermore, LWT instead of the traditional wavelet transform and discrete cosine transform is used. The synchronization code is embedded into low frequency coefficient of LWT domain and the low frequency coefficient is embedded into the maximum singular value obtained by SVD to improve the robustness of the proposed BAWA-SL. Moreover, the watermarking is encrypted by combing the improved cat transform and logistic transform to further improve the security of watermarking. Finally, the experimental results demonstrate that our proposed method obtains better performance than the chosen benchmarks in terms of security, signal to noise ratio and payload.

Keywords: Blind Audio Watermarking; QIM; SVD; Synchronization Code

1 Introduction

In recent years, copyright protection techniques for digital data have received a surge of attention due to its significant potential applications in a variety of aspects of people's daily lives. Illegal copy and unauthorized manipulation of the multimedia documents have been a crucial issue in term of their destructiveness for copyright protection. To avoid such information being invaded and destroyed, digital watermarking technology is proposed to deal with the problems [4, 21, 30]. In general, watermarking technologies can be classified into three categories, *i.e.*, audio watermarking, image watermarking and video watermarking according to the corresponding application. Among them, audio watermarking plays an important role in the watermarking technologies. International federation of the phonographic industry demands that any audio watermarking system should have the following properties and characteristics [10–12, 15]:

1) Robustness: The ability to extract a watermarking from a watermarked audio signal after various signal processing attacks;

2) Imperceptibility: Although the watermarking is embedded into the audio signal, the quality of the watermarked signal should not be degraded, and Signal to Noise Ratio (SNR) should exceed 20dB;

3) Payload: It should be also more than 20 bps (bit per second);

4) Security: We should ensure that watermarking encryption algorithm is safe and watermarking information will not be decrypted by attackers.

There exists a trade-off among the above four properties for each audio watermarking system. For example, to a certain extent, heavy payload causes the degradation of the imperceptibility, while robustness is in the inverse proportion to imperceptibility. Therefore, an ideal scheme should achieve the better trade-off according to the actual requirements. Generally, watermarking algorithms are divided into two main categories: time domain [2, 34] and frequency domain [13, 25]. Compared with frequency domain algorithms, the time domain algorithms are more effective and efficient. However, their robustness is much lower.

On the other hand, the most widely used method in the audio watermarking algorithms is to embed an image into the audio. It can enhance the security and robustness of the audio watermarking. The audio watermarking algorithms can be divided into two different categories, *i.e.*, non-blind audio watermarking algorithms and blind audio watermarking algorithms. In the non-blind audio watermarking algorithms, the watermarking can be extracted by the original image data and reservation information. Different form the general audio watermarking methods, the watermarking can be extracted without the original image data and reservation information. Therefore, the blind audio watermarking schemes are studied and developed. In recent years, the blind audio watermarking methods have been widely used in many fields [3, 14, 16–19, 23].

However, the current blind audio watermarking algorithms suffer from the following challenges:

1) A bulk of audio watermarking algorithms cannot resist cropping and shifting attack [18, 23];
2) The security of the watermarking cannot satisfy the customers' requirements, and the watermarking information may be decrypted by those clever pirates [16];
3) The robustness of some algorithms based on time domain synchronization code is not good enough [3];
4) Payload for a large amount of algorithms is far from users's requirements [14, 19];
5) The traditional wavelet transform is based on convolution and usually leads to a heavy computation load [17].

In order to address above challenges, in this paper, we propose an efficient audio watermarking algorithm based on Singular Value Decomposition (SVD) and Lifting Wavelet Transform (LWT) domain synchronization code to strengthen the confidentiality of information. The main contributions of this paper are summarized as follows.

1) We present a robust, secure and Blind Audio Watermarking Algorithm based on SVD and LWT domain synchronization code, called BAWA-SL to strengthen the copyright protection;
2) We use synchronization code to enhance the security of BAWA-SL. More specifically, we embed the synchronization code into the host audio by using the Quantization Index Modulation (QIM) to achieve the blind extraction of watermarking;
3) We apply LWT instead of traditional wavelet transform or discrete cosine transform to embed the synchronization code into low frequency coefficient of LWT domain and the low frequency coefficient into maximum singular value by leveraging SVD to improve the robustness of BAWA-SL;
4) We improve cat transform and logistic transform to encrypt the watermarking, further improving the security of watermarking;
5) We conduct extensive simulations in two different scale datasets for performance evaluation. The simulation results demonstrate that the proposed BAWA-SL outperforms the comparison algorithms in terms of security, robustness and payload.

The rest of this paper is organized as follows. Section 2 describes LWT and SVD. The embedding and extracting method of synchronization code is described in Section 3. Section 4 introduces the embedding and extracting method of watermarking. Section 5 shows the experimental results. Finally, this paper is concluded in Section 6.

2 Background and Related Work

After the synchronization code and watermarking information are embedded into the audio signal, the structure of the audio signal is described in Figure 1.

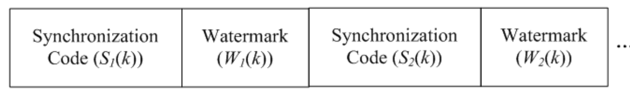

Figure 1: Structure of embedded audio

2.1 LWT

LWT has several advantages, summarized as follows [6, 9]:

1) It acquires biorthogonal wavelet construction completed in time domain without the participation of Fourier transform;
2) It has the time-frequency localization capability and is constructed by some simple wavelet functions;
3) LWT coefficients are integers without quantization errors compared to the first generation wavelet transform. Some researches [20, 31] have employed LWT to improve the robustness of audio watermarking system. By making full use of above advantages, this paper applies LWT into watermarking information embedding procedure.

Next, we give the detailed description of LWT implementation procedure. Note that the construction process of LWT is the inverse process of its decomposition process. In this section, we only introduce the specific flow of LWT decomposition. The specific LWT decomposition process consists of three steps:

Step 1. Split: It is defined as lazy wavelet implementation, which just divides audio $T(n)$ into even samples and odd samples, called $Te(n)$ and $To(n)$, respectively.

$$\begin{aligned} Te(n) &= T(2n), \\ To(n) &= T(2n+1), \end{aligned}$$

where n is the number of samples and is a non-negative integer.

Step 2. Predict: Let even samples predict odd samples as well as keep even samples unchanged. The difference between the prediction value of $P[Te(n)]$ and the real value of $To(n)$ is expressed as follows:

$$x(n) = To(n) - P[Te(n)],$$

where $P[\cdot]$ is the prediction operator, and $x(n)$ is high frequency component of $T(n)$, representing a high-pass filter.

Step 3. Update: Exploit $x(n)$ to update $Te(n)$:

$$c(n) = Te(n) + U[x(n)],$$

where $U[\cdot]$ is the update operator, and $c(n)$ is the low frequency component of $T(n)$, representing a low-pass filter.

2.2 SVD

Recently, some researches on applying SVD into the audio watermarking have been investigated [1, 26], since SVD has unique and special characteristic in the robust watermarking to withstand attacks. Specifically, for the biggest singular value $S(1,1)$, it is not obviously affected when going through some attacks. SVD is usually implemented with some frequency domain transform, for the reason that this combination can obtain a better robustness. SVD has become a frequently used method in the watermarking field and has been applied into the related researches [7] which connect SVD with frequency transform to obtain good robustness. Therefore, in this pa-per, we combine LWT with SVD to obtain a better robustness. The SVD of matrix B^{m*n} is described as follows:

$$B = USV^T,$$

$$B = \begin{pmatrix} U_{(1,1)} & \cdots & U_{(1,r)} \\ \vdots & \ddots & \vdots \\ U_{(m,1)} & \cdots & U_{(m,r)} \end{pmatrix} * \begin{pmatrix} S_{(1,1)} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & S_{(r,r)} \end{pmatrix} * \begin{pmatrix} V_{(1,1)} & \cdots & V_{(1,r)} \\ \vdots & \ddots & \vdots \\ V_{(n,1)} & \cdots & V_{(n,r)} \end{pmatrix}^T,$$

where U is an $m \times r$ matrix; V is an $n \times r$ matrix; S is an $r \times r$ diagonal matrix and its elements are non-negative. The diagonal elements of S are the singular values of B in descending order. The superscript T is the matrix transposition and r is the rank of matrix B.

3 The Proposed Synchronization Code

3.1 Synchronization Code Generating Process

Firstly, we generate the synchronization code according to the chaos theory [8] due to its characteristics of flexibility and safety. A binary shift Bernoulli is defined as follows:

$$x(k+1) = \begin{cases} 2x(k), & if\ 0 \leq x(k) \leq \dfrac{1}{2} \\ 2x(k) - 1, & if\ \dfrac{1}{2} \leq x(k) \leq 1 \end{cases}$$

where $x(0) \in (0,1)$ is the secret key, and its value is the specified at the stage of initialization.

Then, $x(k)$ is turned into the synchronization code sequence $A = \{a(k)|1 \leq k \leq Lsyn\}$ by Equation (1).

$$a(k) = \begin{cases} 1, & if\ x(k) > \tau \\ 0, & otherwise \end{cases} \tag{1}$$

Where $Lsyn$ is the number of segment and τ is a predefined threshold used for generating synchronization code. To acquire the safer scheme, we generate 10 bits synchronization code. Additionally, different from time domain synchronization algorithm, we embed synchronization code into low frequency coefficient of LWT domain to increase the robustness of the algorithm.

3.2 Embedding Process

Step 1. The synchronization code insertion part $S_m(k)$ is transformed into LWT domain, and m is the number of segment, expressed as follows:

$$[CA_m, CD_m] = LWT(S_m(k)),$$

where CA_m is the low frequency DC coefficient and CD_m is the high frequency AC coefficient.

Step 2. We select low frequency coefficient CA_m for synchronization code embedding where CA_m is divided into $Lsyn$ segments and each segment has p samples, and the process is described as follows:

$$LA_m(k) = CA_m(k \cdot p + u), 1 \leq k \leq Lsyn, 1 \leq u \leq p.$$

Step 3. Each bit of the synchronization code is embedded into $LA_m(k)$.

$$QA_m(k) = \begin{cases} round\left(\dfrac{LA_m(k)}{\Delta}\right) \cdot \Delta, & a(k) = 1 \\ floor\left(\dfrac{LA_m(k)}{\Delta}\right) \cdot \Delta + \dfrac{\Delta}{2}, & a(k) = 0 \end{cases}$$

Where Δ is the embedding strength; round() is the rounding to the nearest integer; floor() is the round to minus infinity. We can adjust it to achieve a better trade-off between robustness and transparency according to the actual requirement.

Step 4. We apply the inverse LWT to $QA_m(k)$ to get the completed synchronization code embedding process.

$$LWT^{-1}[QA_m(k), CD_m] = S_m^{'}(k).$$

Step 5. Stop the embedding process when n equals to 4096; otherwise, repeat Steps 1 to 4. All synchronization codes are embedded into the whole segments of the host audio signal.

3.3 Extracting Process

We employ the following rule to extract synchronization code.

Step 1. $QA_m^{'}(k)$ is obtained by applying LWT to $QS_m^{''}(k)$.

Step 2. We use the following formula to extract watermarking.

$$a^{'}(k) = \begin{cases} 0, & if \ \frac{1}{4}\Delta \leq \mathrm{mod}(QA_n^{'}(k), \Delta) \leq \frac{3}{4}\Delta \\ 1, & otherwise \end{cases}$$

Where $\mathrm{mod}(\cdot)$ is the modulus after the division.

4 The Proposed Watermarking Scheme

4.1 Marking Preprocessing

Arnold encryption is one of the most frequently used watermarking encryption methods [5]. However, traditional Arnold transform only adopts one pair of keys to encrypt watermarking image, which is easy to be decrypted by attackers. Thus, in this paper, we propose an improved Arnold transform to enhance the safety of watermarking information. We divide original picture into four individual parts and encrypt this four individual parts by traditional Arnold transform. Besides, we introduce five pairs of keys to encrypt a picture. Detailed steps are illustrated as follows.

Step 1. Transform an image into a binary image A.

Step 2. Cut A into four individual parts, named A_1, A_2, A_3 and A_4 respectively, and name the original picture A as A_5.

Step 3. Utilize the generalized Arnold transform to encrypt $A_i (i = 1, 2, 3, 4, 5)$, and the generalized Arnold transform is defined as follows:

$$\begin{aligned} \begin{pmatrix} x_{m+1} \\ y_{m+1} \end{pmatrix} &= \begin{pmatrix} 1 & p \\ q & pq+1 \end{pmatrix} \begin{pmatrix} x_m \\ y_m \end{pmatrix} \mathrm{mod}\ N \\ &= M \begin{pmatrix} x_m \\ y_m \end{pmatrix}, \end{aligned}$$

where $x_m, y_m \in \{0, 1, \ldots, N-1\}$, N is the size of original picture; (x_m, y_m) is the primitive matrix value; (x_{m+1}, y_{m+1}) is the matrix values after transformation; p and q are the control parameters; $M = \begin{pmatrix} 1 & p \\ q & pq+1 \end{pmatrix}$ is the key to encrypt the five individual parts. Firstly, we use M_1, M_2, M_3 and M_4 to the encrypted four parts $A_1^{'}$, $A_2^{'}$, $A_3^{'}$ and $A_4^{'}$. Then, these four parts are combined as a whole part, named $W_5^{'}$. The last matrix key M_5 is applied to $W_5^{'}$.

We can get the inverse Arnold transform according to the improved Arnold transform, described as follows:

Step 1. Read the encrypted binary image $W_5^{''}$.

Step 2. Get W_5 through decrypting $W_5^{''}$ by Equation (4.1).

Step 3. Divide $W_5^{'}$ into four individual parts, named $A_1^{'}$, $A_2^{'}$, $A_3^{'}$ and $A_4^{'}$ respectively.

Step 4. Apply inverse Arnold transform to $A_1^{'}$, $A_2^{'}$, $A_3^{'}$ and $A_4^{'}$ respectively. Arnold transform is defined as follows.

$$\begin{pmatrix} x_m \\ y_m \end{pmatrix} = \begin{pmatrix} 1 & p \\ q & pq+1 \end{pmatrix}^{-1} \begin{pmatrix} x_{m+1} \\ y_{m+1} \end{pmatrix} \mathrm{mod}\ N.$$

Step 5. Combine A_1, A_2, A_3 and A_4 to obtain A_5.

4.2 Watermarking Embedding and Extracting Procedure

4.2.1 Embedding Procedure

In this section, we employ QIM to achieve the embedding and extracting processes, and the embedding algorithm is described as follows:

Step 1. Perform LWT on the watermarking insert segment $W_n(k)$.

$$[KA_m, KD_m] = LWT(W_m(k)).$$

Step 2. Select the low frequency coefficient KA_m for watermarking information embedding. KA_m is recombined into Matrix MA_m.

Step 3. Apply SVD into Matrix to obtain $S_m(1,1)$, expressed as follows:

$$KA_m = U_m S_m V_m^T.$$

Step 4. Insert the watermarking into $S_m(1,1)$ with QIM. Specifically, the encrypted watermarking after the improved Arnold transform $w(k)$ is inserted to $S_m(1,1)$ of each matrix, defined as follows.

$$Q_m = round\left(\frac{S_m(1,1)}{\beta}\right), D_m = \mathrm{mod}(Q_m, 2).$$

Here, we can adjust it to a lower value to increase the imperceptibility of the watermarking algorithm. Conversely, the robustness of the algorithm will decrease. This situation requires us to adjust β to obtain the reasonable balance between imperceptibility and robustness.

Step 5. Apply $Q_m = Q_m + 1$ when $D_m = 0$ and $w(k) = 1$ or $D_m = 1$ and $w(k) = 0$.

$$Q_m = \begin{cases} Q_m + 1, & if \ w(k) = 1 \ and \ D_m = 0 \\ Q_m + 1, & if \ w(k) = 0 \ and \ D_m = 1 \end{cases}$$

Where $w(k)$ is the encrypted watermarking information by the improved Arnold transform.

Step 6. $S_m(1,1)$ is further modified by the updated Q_m, as follows:

$$S_m^{'}(1,1) = \beta \times round(Q_m).$$

Step 7. Apply the inverse SVD transform as follow:

$$KA_m^{'} = U_m S_m^{'} V_m^T.$$

Step 8. Exploit the inverse LWT to obtain the watermarked audio, described as follows:

$$LWT^{-1}[KA_m^{'}, KD] = W_m^{'}(k).$$

Step 9. Stop the embedding process when all watermarking information is embedded into the whole host audio signal; Otherwise, repeat Steps 1-8.

4.2.2 Extracting Procedure

Extracting process is implemented as follows.

Step 1. Implement LWT on watermarked audio $W_m^{''}(k)$ which suffers from some attacks.

$$LWT(W_m^{''}(k)) = [KA_m^{''}, KD_m^{'}].$$

Step 2. Obtain the low frequency coefficient $KA_n^{''}$.

Step 3. Reconstruct $KA_m^{''}$ into Matrix $KA_m^{'''}$, and perform SVD of $KA_m^{'''}$.

$$KA_m^{'''} = U_m^{'} S_m^{''} V_m^{'T}.$$

Step 4. Let $Q_m^{'} = round(S_m^{''}(1,1)/\beta)$ and $D_m^{'} = $ mod $(Q_m^{'}, 2)$, and extract water-marking information according to Equation (2):

$$w^{'}(k) = \begin{cases} 1, & D_m^{'} = 1 \\ 0, & D_m^{'} = 0 \end{cases} \qquad (2)$$

Step 5. Stop extracting process when all encrypted watermarking information is obtained; otherwise, repeat Steps 1, 2, 3 and 4, and combine these information as a whole part, named $h(k)$.

Step 6. Apply the improved inverse Arnold transform to $h(k)$, and obtain the watermarking information.

We introduce a segment where the scanned size is L_1 and synchronization code is bit by bit. When the synchronization code is extracted successfully, the watermarking information is founded accurately. However, when the synchronization code is not extracted, the segment is moved to the next bit.

5 Performance Evaluation

5.1 Experiment Settings

The simulation is implemented on Matlab programming platform, and the test environment is set up on a personal computer with Intel(R) Core(TM) i5-4590M CPU processor and 4.00 G RAM over Windows 7.

In the simulation, two datasets are used for performance evaluation. The first dataset is 16bit mono WAV audio whose sample rate is 24000 Hz, and watermarking is 64*64 bit binary image. The second dataset is twelve host signals which are from the RWC music-genre database [8]. For the simplicity, we use DS1 and DS2 to represent the two datasets respectively.

Furthermore, the related parameter settings are as follows: The number of bits embedded into the host audio Nw=4096 bits, the duration of the host audio T=19s and the data embedding payload P=215bps.

In order to comprehensively evaluate the performance of the proposed watermarking scheme, four performance indexes including Normalized Correlation (NC) [27], Peak Signal to Noise Ratio (PSNR) [24], SNR and Payload are adopted, described as follows.

5.2 Experiment Results

5.2.1 Security Analysis

We employ Mean Opinion Score (MOS) [27] and SNR to evaluate imperceptibility of the watermarked audio signal. In particular, SNR is the objective way, while MOS is the subjective way. The score sheet of MOS is depicted in Table 1. In our experiment, 5 students are required to classify the difference between the original and the watermarked audio according to a 5-point MOS, which is described as follows:

1) Very annoying;
2) Annoying;
3) Slightly annoying;
4) Perceptible but not annoying;
5) Imperceptible.

MOS way: Figures 2, 3 depict the original audio signal and embedded audio signal in the dataset DS1 respectively. We can observe that the difference between the original signal and the embedded signal is not very obvious. However, as shown in Table 2, MOS can be used to prove indistinguishable. In Table 2, the average of MOS for the tested audio excerpts in the dataset DS1 is 4.77. It means that the watermarked audio and the original audio in the dataset DS1 are perceptually indistinguishable. Similarly, as illustrated in Table 3, the average of MOS in the dataset DS2 is higher than that in DS1. It means that it is more difficult for the large-scale dataset to perceptually distinguish the original audio and watermarking audio. The comparison results demonstrate that the proposed algorithm can enhance the security efficiently.

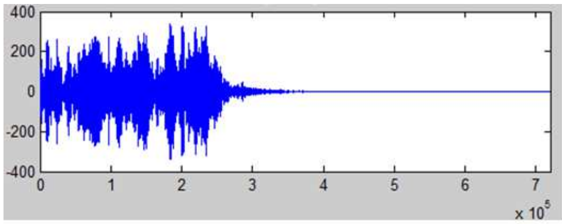

Figure 2: Original audio signal (DS1)

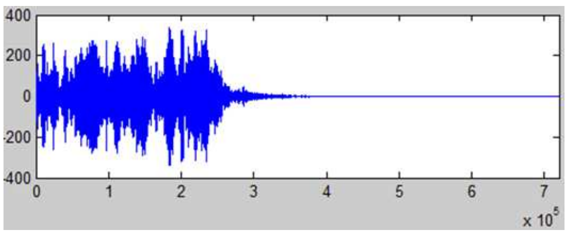

Figure 3: Watermarked audio signal (DS1)

Table 2: MOS of the watermarked audio (DS2)

Student	1	2	3	4	5
MOS	4.83	4.87	4.85	4.79	4.85
Average	4.838				

Table 3: MOS of the watermarked audio (DS1)

Student	1	2	3	4	5
MOS	4.7	4.78	4.82	4.69	4.88
Average	4.77				

SNR way: Figures 4, 5 show the relationship between SNR and quantization step in the dataset DS1 and DS2 respectively. We can observe that the minimum SNR in DS1 is higher than 20 dB, and the minimum SNR in DS2 is higher than 32 dB.

In summary, SNR and MOS results demonstrate that the imperceptibility of the proposed algorithm is up to the regulated standard. This is because we embed the synchronization code into the host audio by using the Quantization Index Modulation (QIM) to achieve the blind extraction of watermarking.

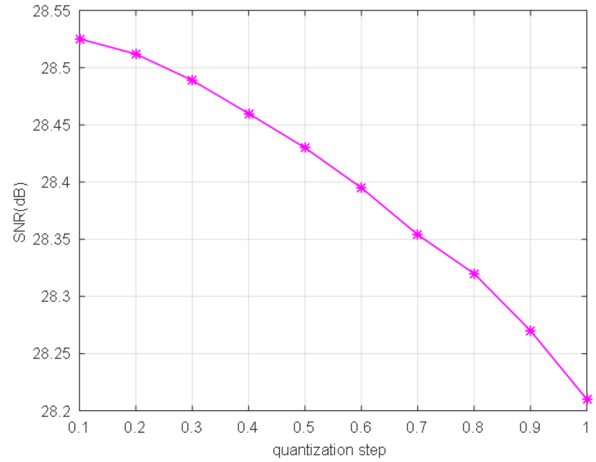

Figure 4: SNR of the proposed method in DS1

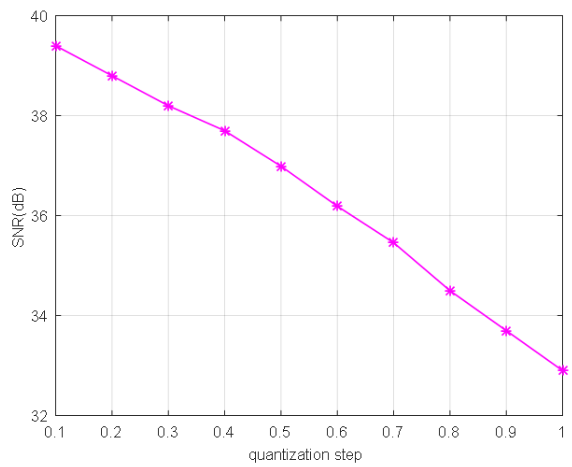

Figure 5: SNR of the proposed method in DS2

5.2.2 Robustness Analysis

Several attacks are used to investigate the performance of the proposed algorithm, including cropping, resampling, filtering attack, white noise and MP3 compression. In the simulations, we use [29, 35] as the comparison algorithms. Figures 6, 7, 8 and 9 represent the robustness test results of the proposed algorithm in $DS1$ and $DS2$ respectively. Here, $C1$, $C2$, $R1$, $R2$, $L1$, $L2$, $G1$, $G2$, $M1$ and $M2$ represent Cropping (500 bits), Copping (1000 bits), Resampling (2kHz-C3kHz-C2kHz), Resampling (10kHz-4kHz-10kHz), Low Pass (19.2 KHz), Low Pass (20.4 KHz), Gauss Noise (SNR 11 dB), Gauss Noise (SNR 13 dB), MP3 Compression (64 kbps) and MP3 Compression(80 kbps) respectively. We can observe that in different datasets, the proposed BAWA-SL obtains higher NC and PSNR than the comparison algorithms. It indicates explicitly that BAWA-SL improves the robustness more dramatically and effectively. This is because we apply LWT instead of traditional wavelet transform or discrete cosine transform to embed the synchronization code into low frequency coefficient of LWT domain and the low frequency coefficient into maximum singular value by leveraging SVD to improve the robustness of BAWA-SL.

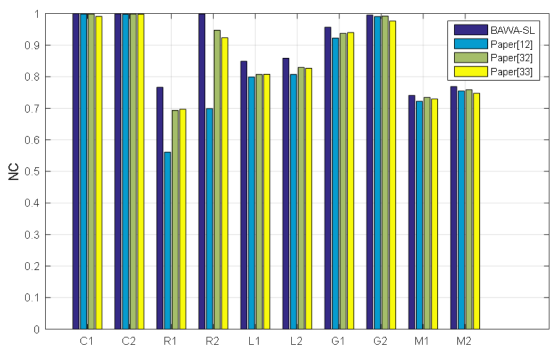

Figure 6: NC comparisons in DS1

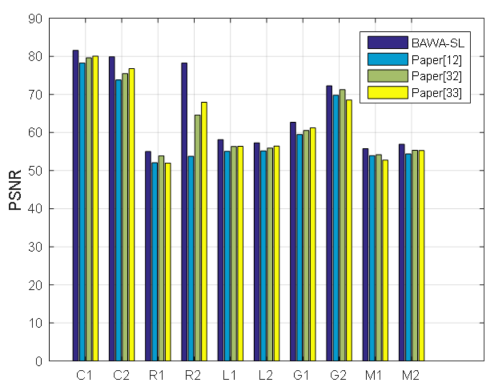

Figure 7: PSNR comparisons in DS1

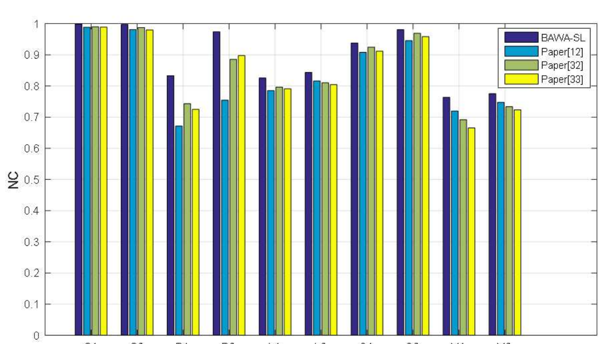

Figure 8: NC comparisons in DS2

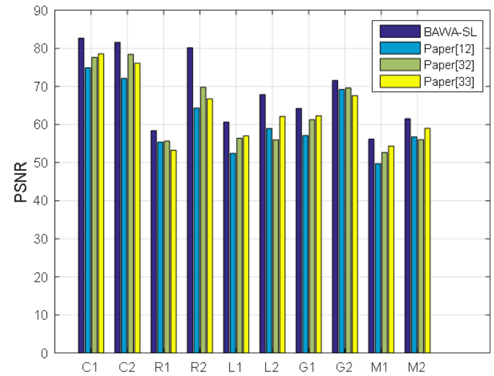

Figure 9: PSNR comparisons in DS2

5.2.3 Payload Analysis

Tables 4, 5 show the payload comparison results of different algorithms in two different datasets respectively. We can observe that compared with the other algorithms, the proposed BAWA-SL algorithm can obtain the biggest payloads in both datasets. The proposed BAWA-SL can play a significant role in finding the promising solutions. From the above comparison experiments, we conclude that BAWA-SL can protect copyright information against being compromised and invaded effectively and efficiently.

6 Conclusion

In this paper, a novel robust blind audio watermarking algorithm based on the improved SVD and LWT domain synchronization code, called BAWA-SL, is proposed to protect the copyright of products. We embed the synchronization code into LWT domain to resist several synchronization attacks and SVD is used to acquire singular values to improve the robustness of BAWA-SL more apparently. We further develop an improved watermarking encryption based on cat map and Arnold transform to improve the security of audio information. Experimental results demonstrate that the security of watermarking information gets promoted drastically more than five times than the chosen benchmarks. Moreover, our algorithm's payload and SNR are much higher than several mainstream algorithms. In summary, the BAWA-SL is feasible and promising for addressing copyright protection for digital audio data.

Acknowledgements

We would like to thank the editors and all anonymous reviewers for helpful comments and suggestions, which have considerably improved and enhanced the quality of this paper. The authors declare that there is no conflict of interests regarding the publication of this paper.

Table 4: Payload comparison in DS1

Reference	Method	Payload(bps)	Synchronization code
Lie [22]	Amplitude modification	43.1	Adopted
Bhat [3]	DWT– SVD	45.9	Adopted
Lei [9]	DCT–SVD	43	Adopted
Wu [33]	QIM–DWT	172.41	Adopted
Ozer [26]	STFT–SVD	32	Not adopted
Wang [32]	FFT–RSVD	187	Not adopted
BAWA-SL	**LWT–SVD**	**215**	**Adopted**

Table 5: Payload comparison in DS2

Reference	Method	Payload(bps)	Synchronization code
Lie [22]	Amplitude modification	62.1	Adopted
Bhat [3]	DWT– SVD	76.8	Adopted
Lei [9]	DCT–SVD	59	Adopted
Wu [33]	QIM–DWT	216.38	Adopted
Ozer [26]	STFT–SVD	48	Not adopted
Wang [32]	FFT–RSVD	231	Not adopted
BAWA-SL	**LWT–SVD**	**297**	**Adopted**

References

[1] I. A. Ansari, M. Pant, C. W. Ahn, "Robust and false positive free watermarking in IWT domain using SVD and ABC," *Engineering Applications of Artificial Intelligence*, vol. 49, pp. 114-125, 2016.

[2] P. Bassia, I. Pitas, N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 232-241, 2001.

[3] V. Bhat, I. Sengupta, A. Das, "An adaptive audio watermarking based on the singular value decomposition in the wavelet domain," *Digital Signal Processing*, vol. 20, no. 6, pp. 1547-1558, 2010.

[4] 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.

[5] W. Chen, C. Quan, C. J. Tay, "Optical color image encryption based on Arnold transform and interference method," *Optics Communications*, vol. 282, no. 18, pp. 3680-3685, 2009.

[6] I. Daubechies, W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier Analysis & Applications*, vol. 4, no. 3, pp. 247-269, 1998.

[7] P. K. Dhar, T. Shimamura, "Blind SVD-based audio watermarking using entropy and log-polar transformation," *Journal of Information Security and Applications*, vol. 20, pp. 74-83, 2015.

[8] Q. He, X. Wang, M. Huang , J. Lv, L. Ma, "Heuristics-based influence maximization for opinion formation in social networks," *Applied Soft Computing*, vol. 66, pp. 360–369, 2018.

[9] Q. He, X. Wang, Z. Lei, M. Huang, Y. Cai, and L. Ma, "TIM: A two-stage iterative framework for influence maximization in social networks," *Applied Mathematics and Computation*, vol. 354, pp. 338–352, 2019.

[10] M. S. Hwang, C. C. Chang, K. F. Hwang, "A watermarking technique based on one-way hash functions", *IEEE Transactions on Consumer Electronics*, vol. 45, no. 2, pp.286–294, May 1999.

[11] 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. 548–556, Jan. 2000.

[12] M. S. Hwang, K. F. Hwang, C. C. Chang, "A timestamping protocol for digital watermarking", *Applied Mathematics and Computation*, vol. 169, pp. 1276–1284, 2005.

[13] M. S. Hwang, J. S. Lee, M. S. Lee, H. G. Kang, "SVD based adaptive QIM watermarking on stereo audio signals," *IEEE Transactions on Multimedia*, vol. 20, no. 1, pp. 45-54, 2017.

[14] A. Iacovazzi, Y. Elovici, "Network Flow Watermarking: A Survey," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 1, pp. 512-530, 2017.

[15] R. Jain, M. C. Trivedi, S. Tiwari, "Digital audio watermarking: A survey," in *Advances in Computer and Computational Sciences*, pp. 433-443, 2018.

[16] A. M. Joshi, S. Gupta, M. Girdhar, R. Sarker, "Combined DWT-DCT-based video watermarking algorithm using Arnold transform technique," in *Proceedings of the International Conference on Data Engineering and Communication Technology*, pp. 455-463, 2017.

[17] Y. Kang, K. Yang, J. Wang, Y. Liu, "Multiple delayed position of echo hiding algorithm research and

development," in *IEEE International Conference on Signal and Image Processing*, pp. 514-518, 2016.

[18] B. S. Ko, R. Nishimura, Y. Suzuki, "Time-spread echo method for digital audio watermarking using PN sequences," *IEEE Transactions on Multimedia*, vol. 7, no. 2, pp. 212-221, 2002.

[19] L. Laouamer, "Towards a robust and fully reversible image watermarking framework based on number theoretic transform," *International Journal of Signal and Imaging Systems Engineering*, vol. 10, no. 4, pp. 169-177, 2017.

[20] B. Lei, I. Y. Soon, E. L. Tan, "Robust SVD-based audio watermarking scheme with differential evolution optimization," *IEEE Transactions on Audio Speech & Language Processing*, vol. 21, no. 11, pp. 2368-2378, 2013.

[21] J. Li, T. Zhong, X. Dai, C. Yang, R. Li, Z. Tang, "Compressive optical image watermarking using joint Fresnel transform correlator architecture," *Optics and Lasers in Engineering*, vol. 89, pp. 29-33, 2017.

[22] W. N. Lie, L. C. Chang, "Robust and high-quality time-domain audio watermarking based on low-frequency amplitude modification," *IEEE Transactions on Multimedia*, vol. 8, no. 1, pp. 46-59, 2006.

[23] H. Liu, A. Kadir, X. Sun, "Chaos-based fast colour image encryption scheme with true random number keys from environmental noise," *IET Image Processing*, vol. 11, no. 5, pp. 324-332, 2017.

[24] G. Mihajlovic, J. C. Read, N. Smith, P. V. D. Heijden, C. H. Tsang, "Improved sig-nal-to-noise ratio in current perpendicular-to-plane giant magnetoresistance sensors using strong exchange-biased reference layers," *IEEE Magnetics Letters*, vol. 8, pp. 1-4, 2017.

[25] S. P. Mohanty, A. Sengupta, P. Guturu, E. Kougianos, "Everything you want to know about watermarking: From paper marks to hardware protection: From paper marks to hardware protection," *IEEE Consumer Electronics Magazine*, vol. 6, no. 3, pp. 83-91, 2017.

[26] H. Ozer, B. Sankur, N. Memon, "An SVD-based audio watermarking technique," in *Signal Processing and Communications Applications Conference*, pp. 51-56, 2005.

[27] F. Ribeiro, D. Florencio, C. Zhang, M. Seltzer, "Crowdmos: An approach for crowdsourcing mean opinion score studies," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'11)*, vol. 7882, pp. 2416-2419, 2011.

[28] RWC database, 2018. (`https://staff.aist.go.jp/m.goto/RWC-MDB/`)

[29] Z. Su, G. Zhang, F. Yue, *et al.*, "SNR-constrained heuristics for optimizing the scaling parameter of robust audio watermarking," *IEEE Transactions on Multimedia*, vol. 99, pp. 1, 2018.

[30] K. Uehira, K. Suzuki, H. Ikeda, "Does optoelectronic watermark technology migrate into business and industry in the near future? -applications of optoelectronic watermarking technology to new business and industry systems utilizing flat panel displays and smart devices," *IEEE Transactions on Industry Applications*, vol. 52, no. 1, pp. 511-520, 2016.

[31] X. Y. Wang, 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.

[32] J. Wang, R. Healy, J. Timoney, "A robust audio watermarking scheme based on reduced singular value decomposition and distortion removal," *Signal Processing*, vol. 91, no. 8, pp. 1693-1708, 2011.

[33] J. Wu, D. Huang, Y. Q. Huang, "Efficiently self-synchronized audio watermarking for assured audio data transmission," *IEEE Transactions on Broadcasting*, vol. 51, no. 1, pp. 69-76, 2005.

[34] S. Xiang, L. Yang, Y. Wang, "Robust and reversible audio watermarking by modifying statistical features in time domain," in *Advances in Multimedia*, pp. 1-10, 2017.

[35] X. Zhu, J. Ding, H. Dong, *et al.*, "Normalized correlation-based quantization modulation for robust watermarking," *IEEE Transactions on Multimedia*, vol. 16, no. 7, pp. 1888-1904, 2014.

Biography

Wu Wenliang is currently a lecturer in the basic Department of Northern Beijing Vocational Education Institute. He received a Master of Science degree from the Faculty of Science of Central University for Nationalities in 2008 and a Bachelor's degree from Tangshan Normal University in 2005. At present, his research mainly focuses on mathematical applications and image algorithms. He has published about 10 papers, conferences and periodicals.