

An Audio Steganography Based on Run Length Encoding and Integer Wavelet Transform

Hanlin Liu, National University of Defense Technology, China

Jingju Liu, National University of Defense Technology, China

Xuehu Yan, National University of Defense Technology, China

 <https://orcid.org/0000-0001-6388-1720>

Pengfei Xue, National University of Defense Technology, China

Dingwei Tan, National University of Defense Technology, China

ABSTRACT

This paper proposes an audio steganography method based on run length encoding and integer wavelet transform which can be used to hide secret message in digital audio. The major contribution of the proposed scheme is to propose an audio steganography with high capacity, where the secret information is compressed by run length encoding. In the applicable scenario, the main purpose is to hide as more information as possible in the cover audio files. First, the secret information is chaotic scrambling, then the result of scrambling is run length encoded, and finally, the secret information is embedded into integer wavelet coefficients. The experimental results and comparison with existing technique show that by utilizing the lossless compression of run length encoding and anti-attack of wavelet domain, the proposed method has improved the capacity, good audio quality, and can achieve blind extraction while maintaining imperceptibility and strong robustness.

KEYWORDS

Audio Steganography, Integer Wavelet, Large Capacity, Run Length Encoding

1. INTRODUCTION

Steganography is an art of hiding secret information in another seemingly innocuous message, or carrier (Johnson, Duric&Jajodia, 2001). Steganography is different from cryptography. The primary goal of cryptography is to provide secure communication by changing data into a form that it cannot understand except for the sender and the receiver. Since when an attacker does not know the existence of secret message, he will not generate the idea of attacking it. Therefore, steganography is a powerful technique to enhance the security of data transmission. Throughout history, a multitude of methods have been used to hide information. With the development of the Internet and other new technologies, digital steganography technique which is used to embed the secret message into digital multimedia is

DOI: 10.4018/IJDCF.2021030102

This article, published as an Open Access article on February 15, 2021 in the gold Open Access journal, The International Journal of Digital Crime and Forensics (IJDCF) (converted to gold Open Access January 1, 2021), is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

gradually rising. It has developed a strong basis for the area of steganography with a growing number of applications for digital fields like covert communications, annotation etc. So far, various researches on steganography have been carried out on storage media, such as text, image, audio, and video.

Furthermore, auditory system is the largest source of information in addition to the Human Visual System(HVS). Thus the research of audio steganography is of great significance and has wide application scenarios.

Like steganography in other media, audio steganographic technique also has three evaluation metrics (Divya&Reddy, 2012):

1. Capacity means the amount of secret information that can be embedded into the host audio without affecting the perceptual quality of audio.
2. Imperceptibility evaluates how well a secret message is embedded into the cover audio. The difference between audio after hiding and audio before hiding should remain negligible.
3. Robustness indicates the ability of secret message to resist against attacks.

In audio steganography, Human Auditory System (HAS) is used to hide information in the audio. Because HAS has more precision than HVS, audio steganography is more challenging than image steganography (Shirali-Shahreza&Manzuri-Shalmani, (2007).

In this paper, the authors propose an audio steganography method based on run length encoding and integer wavelet transform. The ideal of this method is to design the embedding and extracting procedure of the secret information, and then the authors test the capacity, robustness and anti-detection and compare it with some existing methods. The advantages of the proposed method is large capacity.

The rest of the paper is organized as follows: the related work is introduced in Section II. In Section III the proposed methods are presented. Experimental results and analyses are presented in Section IV. Finally, Section V concludes this paper.

2. RELATED WORK

Steganography techniques can be classified into methods in time and transform domain. The common used methods of steganography in time domain include Least Significant Bits (LSB) (Cvejic&Seppänen, 2004; Roy, Parida, Singh &Sairam, 2012; Vimal& Alex, 2014), Echo Hiding (Chen& Wu, 2008; Yan, Sun & Lu, 2003), Spread Spectrum Method (Rupanshi, Preeti&Vandana, 2014) and etc. Besides, methods in transform domain contain hiding method based on DFT, DCT (Chen, Zhang, Liu, Niu, & Yang,2009; Premalatha, Narayanan,Vikash& Ramesh,2014; Chilhate, Patidar&Chandel, 2015) and DWT (Sheikhan, Asadollahi&Shahnazi, 2011; Bhattacharyya &Sanyal, 2012). It is hard to balance the capacity, imperceptibility and robustness. For instance, LSB has great hiding capacity, however, which is at the expense of more robust. Although methods in transform domain boost up robustness against some attacks, the capacity and imperceptibility are influenced a little.

Vimal and Alex (2014) discussed in their research that dual randomness LSB method was proposed which hides the secret message in randomly selected samples and bits. This method provided more confidentiality compared with conventional LSB method. But compared with the method of Cvejic and Seppanen (2003), the hiding capacity was smaller. Cvejic and Seppanen (2003) showed in their work that a novel modification to standard LSB algorithm was proposed, which was able to embed four bits per sample and improved the capacity of data hiding channel by 33%. They all belong to LSB method whose disadvantage is weak robust.

As we all known, the classical wavelet transform has multi-resolution analysis and good reconstruction properties. But because of the dependence on the Fourier transform, there is also a large amount of calculation. In 1996, Sweldens proposed a lifting scheme that does not rely on Fourier

transform (Sweldens,1996). The lifting wavelet transform (LWT) is also called second-generation wavelet transform. Not only has it inherited the advantages of the first-generation wavelet transform, but also has the advantages of fast calculation speed, in-situ calculation, can easily perform integer wavelet transform and inverse wavelet transform(IWT). Due to all these advantages, it has attracted a lot of attention in the field of audio steganography. Lei, Soon and Li (2012) proposed an audio watermarking scheme based on LWT and SVD. They embedded the encrypted watermark into the low frequency coefficients of the IWT after SVD transform. This method has good robustness and anti-detection. Delforouzi and Pooyan (2008) proposed an adaptive digital audio information hiding technique based on IWT. They introduce the auditory threshold and use it as the threshold of embedding information. Then they embedded the secret information into the wavelet subband coefficients of the carrier audio. Compared with Lei's method, this method has a larger hidden capacity and more undetectable. Nehete, Sawarkar, andSohani (2011) showed in their paper that a novel method for digital audio steganography with security i.e., cryptography was presented where covert data was embedded into the coefficients of host audio in integer wavelet domain using quantization to reduce embedding errors. The disadvantage of the three methods above is that the capacity is small.

A transform domain method based on run length encoding and integer wavelet transform is proposed in this paper. Because of the use of run length encoding, the biggest advantage of this method is the large capacity, in addition, the secret information is embedded into wavelet domain, which leads to the result that it also has a good imperceptibility and the ability to resist the common attacks.

3. THE PROPOSED METHOD

The proposed method consists of two parts: secret information embedding method and secret information extraction method. First, integer wavelet transform is introduced.

3.1. Integer Wavelet Transform

Lifting scheme is presented by W.Sweldens in 1994 for creating wavelets entirely in the spatial domain. The main difference with classical constructions is that it is independent of the Fourier transform, wavelets are not necessary by translation and dilation of a function. In general, lifting scheme can be divided into two steps: split and upgrade. It starts with Lazy wavelet, a function which essentially does not do anything, but which has the formal properties of wavelet. Assuming S_{j-1} is the original signal, which is split into two separate parts, even and odd. That is:

$$\begin{aligned} Split_{(s_{j-1})} &= (s_j, d_j) \\ s_j &= s_{j-1,2k} \\ d_j &= s_{j-1,2k+1} \end{aligned} \tag{1}$$

Then, update every element in every part according to the expected features. The first is to lift the odd ones based on the even ones with the use of a predictor. This process is called dual lifting:

$$d_j = d_j - P(s_j) \tag{2}$$

If P is a good predictor, the difference or detail d_j will be regarded as the wavelet coefficients nearly without the correlation and called high-pass part. The second is to lift the even ones with the help of the high-pass part. The even ones will be replaced by smoothed values with the use of an update operator U applied to the details. This process is called primal lifting:

$$S_j = S_j + U(d_j) \quad (3)$$

Here S_j will be approximately approached to original signal and called low-pass part. The lifting scheme can go on using in the low-pass part S_j for multiresolution decomposition. The inverse lifting transform is the same as changing the signs of formulas described above and run them backwards, i.e. starts at the output:

$$\begin{aligned} S_j &= S_j - U(d_j) \\ d_j &= d_j + P(s_j) \\ \text{Merge}(S_{j-1}) &= (s_j, d_j) \end{aligned} \quad (4)$$

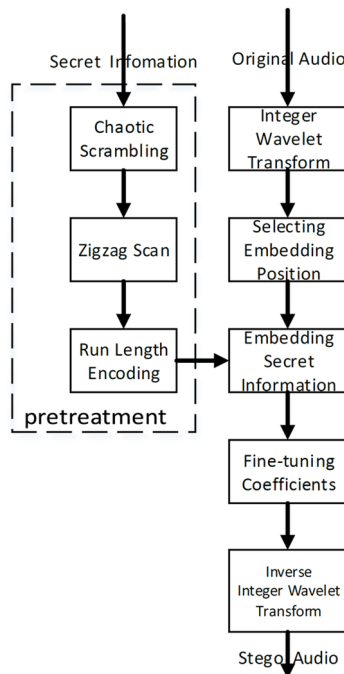
3.2. The Embedding Method

The embedding method consists of two steps. Step 1 needs to preprocess the secret information. In this step, there are three processes, including chaotic scrambling, Zigzag transform and run-length encoding. Step 2 is to embed the secret information into the integer wavelet coefficients of the original audio. Finally, the authors transform the integer wavelet coefficients into the time domain and obtain the stego audio in step 3. The embedding process is shown in Figure1.

3.2.1 Preliminary Process

Preliminary process is an important means to improve the security of secret information, and it has become an indispensable part of information hiding. It consists of three sub processes, that is, chaotic

Figure 1. The embedding process of the proposed method



scrambling, zigzag scan and run length encoding. The purpose of chaotic scrambling is to encrypt the secret information and increase the security and chaotic degree of secret information. The role of the zigzag scan is to convert two-dimensional secret information into one-dimensional secret information. And run length encoding is used to compress the secret information.

3.2.1.1 Chaotic Scrambling

Scrambling technology is widely used in the preprocess of secret information encryption. At present, there are many scrambling technologies. Most of them are used in the field of digital image processing. The most common scrambling methods are Arnold transform, magic square, Hilbert curve, Conway games, and Tangram algorithm et al. However, due to the large redundancy in the image or audio, it will consume a lot of time and space resource using the above methods for scrambling. In addition, some scrambling algorithms are periodic, which will greatly reduce the security of the scrambling. Compared with the above techniques, Logistic chaotic sequence is favored in data scrambling and encryption due to their noise-like, correlation and accurate reproduction.

Logistic chaotic sequence is defined as:

$$x_{i+1} = \mu x_i (1 - x_i) \quad (5)$$

Here, μ is called branch parameter, $0 \leq \mu \leq 4$. However, owing to the limitation of the data accuracy stored in the computer, the chaotic sequence will change periodically. In order to prevent this kind of periodical change from reducing security, the authors use the following Logistic chaotic modulated:

$$x_{i+1} = 0.999\mu x_i (1 - x_i) + 0.001x_i \quad (6)$$

or:

$$x_{i+1} = 1.001\mu x_i (1 - x_i) - 0.001x_i \quad (7)$$

The authors use the previous value x_i to fine-tune the current value x_{i+1} . By this way, random characteristic and uniform distribution characteristic of Logistic chaotic sequence will be enhanced.

The equation (1), (2), and (3) show the recursion relationship between value and its successor in the Logistic chaotic sequence. Therefore, the Logistic chaotic sequence has extremely sensitive dependence on the initial value x_0 . Even if the initial value has a minor change, the final chaotic sequence will be completely different. In the non-divergent dynamical system, the Lyapunov exponent is used as the index of system chaos discrimination. In 1983, Grebogi proved that as long as the maximum Lyapunov exponent greater than 0, it can be sure the existence of chaos in a system. The authors did some experiments and found that when $3.67 \leq \mu \leq 4$ and $0 < x_0 < 1$, Lyapunov exponent is greater than 0. Meanwhile, the equation (1), (2), and (3) show good chaotic characteristic.

In this paper, the authors generate the chaotic sequence with the same number of secret information by using the chaotic mapping equation (2) or (3) where $\mu = 3.9$ and $x_0 = 0.6$. All the values in the chaotic sequence are fractional numbers between 0 and 1. By expanding the same multiples and other processing, the authors transform each of these values into an integer between 0 and 255. Finally, after performing bitwise XOR operation on secret information and integer chaotic sequence, the authors will obtain the chaotic scrambled secret information.

3.2.1.2 Zigzag Scan

Zigzag scan is another scrambling method. It scans and gets the elements in a matrix from the upper left corner to the lower right corner one by one followed as the Z shape. Through this method, the authors can achieve the purpose of scrambling and dimensionality reduction. Zigzag scan is characterized by simple algorithm and low time complexity, which makes it widely used in the reading of discrete cosine transform coefficients, digital images compression and scrambling. Figure2 shows the zigzag scan.

Zigzag transform makes the chaotic sequence be further scrambled and improves the security. At the same time, it transforms the two-dimensional data to one-dimensional data, which will get prepare for the run length encoding.

3.2.1.3 Run Length Encoding

Run length refers to the repetition of each character in the data stream length. Run length encoding is a source encoding, which belongs to lossless compression encoding and can reduce data storage space and transmission capacity. The basic principle of run length encoding is to use a symbol value or string length instead of a continuous symbol with the same value, so that the symbol length is less than the length of the original data.

In run length encoding, if the binary data of the secret information is scanned in rows, it is always made up of several consecutive 0s and 1s. The length of the continuous 0 and 1 are called 0 run and 1 run, respectively, which always appear alternately. When the binary sequence starts at 0, then the first run is 0 run, the second must be 1 run and the third is 0 run. Thus any binary sequence can be converted into a run length sequence, so as to achieve the purpose of lossless compression. In this paper, the authors use this feature of run length encoding to improve the embedding capacity of audio information hiding and the extraction of secret information. The storage structure of run length encoding is shown in Table1.

In Table1, a and b in the first row represent the binary secret information value, which is 1 or 0. And Len (a) and Len (b) in the second row represent the run length of the corresponding a and b. The run length encoding of the secret information process as follows:

Figure 2. Zigzag scan

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

Table 1. Storage structure of the run length encoding

binary secret data	a	b	a	b	a	b	a
run length	len(a)	len(b)	len(a)	len(b)	len(a)	len(b)	len(a)

1. Transform the one-dimensional data which is after Zigzag scan into a one-dimensional binary sequence.
2. Scan the first bit of the binary sequence, record the secret information binary value a, and scan a continuous binary string with the same value, record its run length Len (a); and then scan the next successive data string, record the binary b and its run length Len (b), and continue the process until the last bit of the binary sequence is scanned.

3.2.2 Embedding

3.2.2.1 Integer Wavelet Transform

In the wavelet transform, since the filter has floating-point coefficients, even if the input data are composed of integers, the filtered data are no longer integers. In the Introduction, it is mentioned that the lifting wavelet can achieve integer wavelet transform. The proposed method is to embed the secret information into integer wavelet coefficients by using this feature of lifting wavelet.

The embedding algorithm in wavelet domain is mainly based on the coefficients of the approximate subband and the detail subband generated by the wavelet decomposition, which determines the embedding position of the secret information. The embedding position of the secret information is mainly the low frequency domain and the high frequency domain after wavelet decomposition, which correspond to the approximate subband and the detail subband respectively. Compared with the detail subband, the approximate subband is very similar to the original signal, and has a large energy, after the general signal processing, which can still retain its characteristics. So the method in which the secret information is embedded into the approximate subband has good robustness. Therefore, in the proposed method, the authors choose to embed the secret information into the low frequency part of the audio, that is, the approximate subband.

3.2.2.2 Embedding the Secret Information

In this paper, the main idea of the proposed method is to embed the secret information into the approximate subband of integer wavelet coefficients. There are two kinds of secret data need to be embedded, the binary value (1 or 0) of the secret information and the binary value of its run length, which can be seen in Table1. In original LSB method, only one bit is modified and can't

Represent the embedded two data. So the authors decide to embed secret information with two bits. According to the Zhang (2009), the auditory test shows that for the audio with 16-bit sampling precision, the authors can embed secret data into 4th bit utilizing LSB, and the change of the audio before and after the modification can't be perceived. In this case, the maximum change in original data is 8. However, Lu, Ye, and Wu, (2007) shows that the robustness of the algorithm is enhanced with the increase of the embedding position, but the imperceptibility of the carrier audio is reduced. If the authors choose the least significant bit and the second bit as the embedding position, the algorithm is less robust, and if choose the third and fourth bit as embedding position, imperceptibility of carrier audio will be greatly affected. Thus in order to balance both robustness and imperceptibility, the second and third bit are selected as the embedding position. The second bit represents the binary value of the run length, and the XOR value of the second bit and the third bit represents the secret data (1 or 0). The embedding rule are shown in Table 2.

Table 2. Embedding rules of the proposed method

B_3	B_2	binary secret data	binary run length
0	0	0	0
0	1	1	1
1	0	1	0
1	1	0	1

Where B_i represents the i th bit of the embedding coefficient. The run length of the secret data 0 is 3 (binary 11), which are embedded into coefficients 127(0111 1111) and 66 (0100 0010). After embedding, 127 (0111 1111) remains unchanged, 66 (0100 0010) is modified to 70 (0100 0110). The example of the embedding rule is shown in Table 3.

After embedding the secret data into carrier using the above rules, the integer wavelet coefficients of the carrier audio will change, as shown in Table 4.

The embedding position is not sequential or random. Because there is a large number of 0 in the integer wavelet coefficients and these 0 coefficients have a significant effect on audio quality after being modified, severely reducing imperceptibility. In addition, the change of the number of 0 is easy to be found in steganalysis. Therefore, the 0 coefficients should be avoided in the selection of embedding position. It can be seen from Table4 that the maximum change of the coefficient is 6, and the coefficients with absolute value less than 7 are likely to become an absolute value of 0 after modification, which leads to an error in the data extraction. So the coefficient with absolute value less than 7 that is called threshold value cannot be chosen as embedding position. However, in the coefficient with absolute value greater than 7, since the change value is negative, the absolute value of some coefficients will be less than 7 after embedding the secret information, which will result in selecting the wrong embedding position. Thus the coefficient must be fine-tuned after being into embedded data. The method is to change the fourth bit value to 1, so that the absolute value of the coefficient is greater than 7. The embedding process is shown in Figure 3.

3.3. The Extracting Method

The extraction process is similar to the embedding process, which is shown in Figure4.

In this paper, extracting secret message does not need original audio, so this method can achieve blind extraction. The step of extraction is as follows.

1. Perform integer wavelet transform on the stego audio, then get the integer wavelet coefficients.
2. Compare the coefficients with the threshold and find the extraction position of secret data.
3. According to the rule in Table1, extract the secret data and get the binary secret data as well as the corresponding run length.

Table 3. The example of the embedding rule

embedding position	the coefficient 127	the coefficient 66
Change of Coefficient		
Before embedding	0111 1111	0100 0010
After embedding	0111 1111	0100 0110

Table 4. The change of coefficient after embedding

binary secret data	binary run length	Before embedding B_3B_2	After embedding B_3B_2	Change of coefficient
0	0	00	00	0
		01	00	-2
		10	00	-4
		11	00	-6
	1	00	11	+6
		01	11	+4
		10	11	+2
		11	11	0
1	0	00	10	+4
		01	10	+2
		10	10	0
		11	10	-2
	1	00	01	+2
		01	01	0
		10	01	-2
		11	01	-4

4. Perform inverse run length encoding, get binary sequence then convert the binary sequence to one-dimensional decimal data.
5. Perform inverse Zigzag transform on the decimal data, get the two-dimensional data.
6. XOR the two-dimensional data and chaotic sequence, then obtain the secret messages.

4. EXPERIMENTAL RESULT AND ANALYSIS

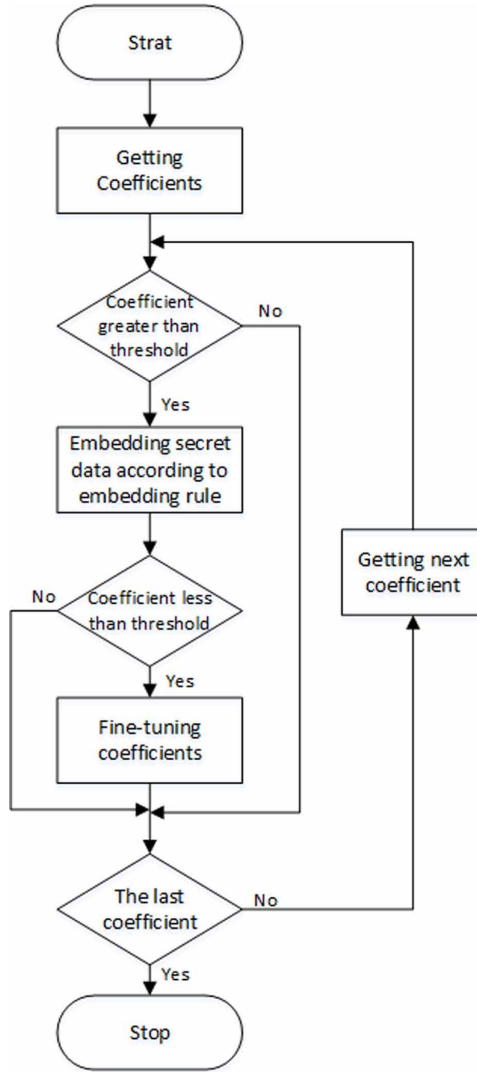
In this paper, the authors use the Matlab platform to complete the proposed test, and the imperceptibility, robustness and capacity are analyzed. The concrete experimental environment as follows:

Experimental platform: Matlab R2014a;
Secret information: 255×255 grayscale image, 'lena.bmp';
Original carrier: 6 minutes and 38 seconds of a two-channel stereo music, sampling frequency 44.1KHz, 16bit quantification, 'music.wav';
Parameter setting: The threshold value is set to 7.

In this section, the authors utilize MSE (Mean Square Error), SNR (Signal to Noise Ratio), BER (Bit Error Rate) and NC (Normalization Coefficient) to evaluate the performance of the proposed method.

Four formulas for calculating MSE, SNR, BER and NC are given as follows, respectively.

Figure 3. Embedding process of the proposed method

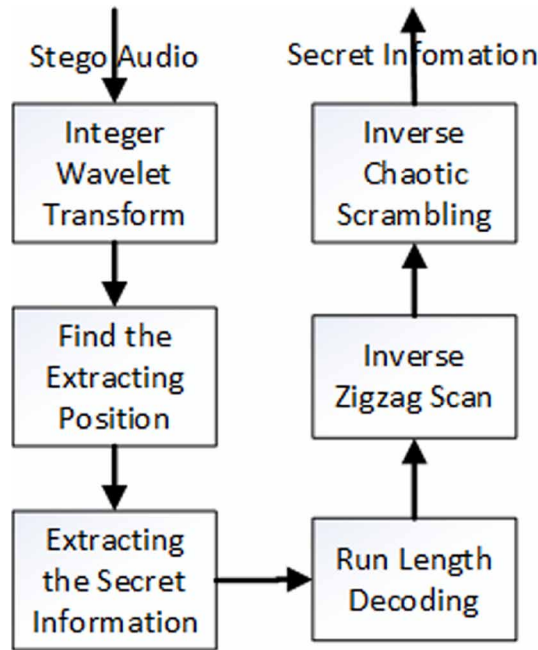


$$MSE = \frac{1}{N} \sum_{i=1}^N (c_i - c'_i)^2 \quad (8)$$

$$SNR = 10 \log \left(\frac{\sum_{i=1}^M c_i^2}{\sum_{i=1}^M (c_i - c'_i)^2} \right) \quad (9)$$

$$BER = (l/L) \cdot 100\% \quad (10)$$

Figure 4. The extracting process of the proposed method



$$NC = \sum_{i=1}^N c_i c_i' / \left(\sqrt{\sum_{i=1}^M c_i^2} \sqrt{\sum_{i=1}^M c_i'^2} \right) \quad (11)$$

In the equations given above, L represents the total number of original confidential information bits and l represents the number of extracted error bits. c_i is the original host audio and c_i' is the steganography audio obtained after the embedding process. Original confidential information taken as s_i and s_i' represents the extracted confidential information.

4.1. Imperceptibility Experiment and Analysis

The authors will analysis and evaluate the imperceptibility of stegoaudio from the objective and subjective aspects. First making comparison between original audio and stego audio from imperceptibility.

Through the experiment we find that there is no significant difference between the waveform of the original audio and stego audio. By observing the difference between stego audio and original audio, it can be found that the main difference between their sampling points are concentrated in ± 0.02 , $-0.05 \sim -0.03$ and $0.05 \sim 0.03$, which is because of the modification and fine-tuning of coefficients. According to the analysis in Subsection Embedding the Secret Information, after embedding the data, the maximum change of the coefficient is ± 6 , and through dividing by 255, this change is quantized to ± 1 interval, which is exactly 0.02. The change caused by fine-tuning the fourth bit of the coefficient is ± 8 , and plus the change ± 6 , the total change is ± 14 . And through dividing by 255, the total change is quantized to ± 1 interval, which is exactly ± 0.03 and ± 0.05 . And the last part of the difference is clearly zero, which is because the last part of coefficients has not been used to embedded and not be modified, so the difference is 0.

In addition to observing the waveforms, the authors can also use objective quantitative criteria to evaluate imperceptibility, such as MES and SNR. MSE is a simple method to measure the average error. The smaller the MSE is, the smaller the change of the original audio after embedding the secret information is, and the higher the imperceptibility of the algorithm is. According to the requirement of the International Recording Industry Association (IFPI), if the audio SNR after embedding the information is greater than 20dB, the change of audio will not be perceived by the human auditory system. The SNR and MES of the original audio and stego audio are shown in Table 5.

Through the experiment and analysis, the authors can see that the method proposed in this paper has good performance in term of imperceptibility, and after embedding the secret information, the change of the original audio is slight, and the human auditory system cannot be perceived.

4.2. Robust Experiment and Analysis

Robustness is an important criterion to evaluate the performance of an algorithm. In the experiment, common audio processing includes noise addition, re-quantization, and resampling.

1. No attack state: do not deal with stego audio, directly using the extraction method to extract the secret information. The extracted secret information and the parameters of the test results are shown in Figure 5 and Table 6.

Table 5. The test result of stego audio

Evaluation criterion	SNR	MER
Value	25.911	0.065

Figure 5. Original secret information and extracted information in condition of no attack(a) Original secret information (b)Extracted information in condition of no attack



(a)



(b)

Table 6. The test result in condition of no attack

Attacks	BER (%)	NC
No attack	0	1

It can be concluded from Figure5 and Table6 that in the case of no attack, the secret information can be extracted at 0 bit error rate, and the extracted secret information has no difference in visuality from the original secret information.

2. Stego audio with original sampling rate 44.1 kHz have been down-sampled to 22.05kHz and up-sampled back to 44.1 kHz. Then up-sampled to 88.2kHz and down-sampled back to 44.1kHz. The extracted secret information and the parameters of the test results are shown in Figure 6 and Table 7.

It can be seen from Figure 6 and Table 7 that the ability of the method to resist the down-sampling attack is obviously superior to that of the upper sampling attack, but the contents also can be properly identified.

Figure 6. Original secret information and extracted information after resampling (a) Original secret information (b) Extracted information after up-resampling (c) Extracted information after down-resampling



Table 7. The test result after resampling

Attacks	The Change of Frequency	BER (%)	NC
up-sampling	44.1KHz ~ 88.2KHz ~ 44.1KHz	0.015	0.992
down-sampling	44.1KHz ~ 22.05KHz ~ 44.1KHz	0.012	0.9965

3. Re-quantization: the authors tested the process of re-quantization of a 16-bit stego audio to 8-bit and back to 16-bit. Then the authors tested the process of re-quantization of a 16-bit stegoaudio to 32-bit and back to 16-bit. The extracted secret information and the parameters of the test results are shown in Figure 7 and Table 8.

It can be seen from Figure 7 and Table 8 that the proposed method has good performance in term of re-quantitative attack. The effect of re-quantitative attack on this method is very small, the extracted secret information is not distorted and the contents of the secret information can be normal identified.

4. Noise Addition: white noise with 50 db, 40 db, and 30 db SNR is added to stego audio. The extracted secret information and the parameters of the test results are shown in Figure 8 and Table 9.

It can be seen from Figure 8 and Table 9 that the ability of this method to resist noise attack varies from noise to noise, and when the SNR is 30 db, the extracted secret information is seriously distorted. This method has strong resistance to white noise with greater than 40 db SNR, and in this case, the extracted secret information without distortion can be normal identification.

4.3. Capacity Analysis

The existing audio steganography method based on wavelet coefficients mostly embeds the secret message bits into the wavelet coefficients in a one-to-one relationship. The number of wavelet

Figure 7. Original secret information and extracted information after re-quantization(a) Original secret information (b) Extracted information after lift quantization (c)Extracted information after reduced quantization



Table 8. The test result after re-quantization

Attacks	The Change of Quantization Bits	BER (%)	NC
lift quantization	16bit ~ 32bit ~16bit	0.001	0.998
reduced quantization	16bit~ 8bit~16bit	0.001	0.998

Figure 8. Original secret information and extracted information after Noise Addition(a) Original secret information; (b)Extracted information after adding noise with 50db SNR; (c) Extracted information after adding noise with 40 db SNR (d) Extracted information after adding noise with 30 db SNR



Table 9. The test result after noise addition

Attacks	SNR (db)	BER(%)	NC
Noise Addition	50	0.011	0.996
	40	0.018	0.98
	30	0.0373	0.869

coefficients determines the embedding capacity. Because the proposed method utilizes compression of run length encoding, compared with other method, it just consumes less space to hide same secret message, and accordingly, same space can be embedded into more secret messages.

The original audio used in this experiment is sampled at 44.1 KHz with length of 398 seconds and quantized by 16 bits. After one level integer wavelet transform, there are 7760000 coefficients

Table 10. Test results comparing with existing methods

Attacks	BER(%)			
	Method in this paper	Method in (Wu, Wu, & E, 2016)	Method in (Tan, Wu, Liu, & Zhou, 2010)	Method in (Tewari, Saxena & Gupta, 2014)
Up-sampling	0.015	0.013	5.40	0.00
Down-sampling	0.012	2.810	0.10	0.00
lift quantization	0.01	0.013	0.20	-
reduced quantization	0.01	7.220	0.20	-
30db white noise	0.037	0.018	-	-
40db white noise	0.018	0.015	0.48	7.00

that can be embedded into secret messages. Suppose each coefficient can be embed into 1 bit, the capacity of the proposed method is $7760000/398=19500$ bps. After using run length encoding, for image message the lowest compression rate can reach 85%, and for text message, can reach 65%. So the capacity of embedding image and text is $19500/0.85=22940$ bps and $19500/0.65=30000$ bps, respectively. If the two-level integer wavelet coefficient is chosen as the embedding position, the minimum capacity will decrease slightly, which is 17460 bps and 24900 bps, respectively.

4.4. Comparison with Related Method

In this subsection, the authors compare the proposed method with existing methods (Wu, Wu, &E, 2016; Tan, Wu, Liu, & Zhou, 2010; Tewari, Saxena&Gupta, 2014), since these three methods are also performed in transform domain, and the first two method also have the characteristic of large capacity. And the results of the comparison are shown in Table 10.

By comparing BER of this method with existing methods', the authors can see that the proposed method has a greater advantage in resisting resampling and re-quantization attack, which is because of the use of fine-tuning. In term of noise attack, it has a poor resistance to noise with SNR less than 40 db.

As for capacity, compared with the highest capacity of 22050 bps and 100bps in (Wu, Wu, &E, 2016; Tan, Wu,Liu,&Zhou, 2010), the method proposed in this paper has higher capacity, 22940bps. And the authors should attribute large capacity of the proposed method to run length encoding, since it compresses the secret information. However, in practice, it is necessary to measure imperceptibility, robustness and capacity according to the practical needs. Even so the advantage of this method in capacity is still evident.

5. CONCLUSION

5.1. Discussion

This paper presents an audio steganography method based on run length encoding and integer wavelet transform. By setting embedding threshold and selecting the embedding position reasonably, stego audio distortion is effectively controlled. And the embedding capacity is improved via utilizing run

length encoding. Experiments show that the proposed method has great imperceptibility. Moreover, compared with existing method, it is evident that the method not only has large capacity, but also has strong robustness for resisting the common attacking behaviors, such as re-quantization and down-sampling. Although this method also has the disadvantage of poor resistance to noise with low SNR, in summary, it is still a practical audio steganography method.

5.2. Future Work

The authors embedding data into fixed 2nd and 3rd coefficients. By reference to existing methods (Huang, Nobutaka & Akira Nishimura, 2014), in future work, we will adaptive hiding positions to achieve complexity of extraction.

REFERENCES

- Bhattacharyya, S., & Sanyal, G. (2012). A robust image steganography using dwt difference modulation (dwtDM). *International Journal of Computer Network & Information Security*, 4(7), 27–40. doi:10.5815/ijcnis.2012.07.04
- Chen, M., Zhang, R., Liu, F. F., Niu, X. X., & Yang, Y. X. (2009). Audio steganography by quantization index modulation in the dct domain. *Journal of Communication*, 30(8), 105–111.
- Chen, T. C., & Wu, W. C. (2008). Highly robust, secure, and perceptual-quality echo hiding scheme. *Audio Speech & Language Processing IEEE Transactions on*, 16(3), 629–638.
- Chilhate, K., Patidar, K., & Chandel, G. S. (2015). Advanced audio steganography technique based on coefficient comparison in dct domain. *International Journal of Advanced Research in Computer Science and Electronics Engineering*, 4(12), 162–164.
- Cvejic, N., & Seppanen, T. (2003). Increasing the capacity of LSB-based audio steganography. *Multimedia Signal Processing, 2002 IEEE Workshop on*, 336–338.
- Cvejic, N., & Seppänen, T. (2004). Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method. *International Conference on Information Technology: Coding and Computing*, 2, 533–537. doi:10.1109/ITCC.2004.1286709
- Delforouzi, A., & Pooyan, M. (2008). Adaptive digital audio steganography based on integer wavelet transform. *Circuits, Systems, and Signal Processing*, 27(2), 247–259. doi:10.1007/s00034-008-9019-x
- Divya, S. S., & Reddy, M. R. M. (2012). Hiding text in audio using multiple lsb steganography and provide security using cryptography. *International Journal of Scientific & Technology Research*, 1(6), 68–70.
- Huang, Ono, Echizen, & Nishimura. (2014). Reversible Audio Information Hiding Based on Integer DCT Coefficients with Adaptive Hiding Locations. Springer Lecture Notes in Computer Science, Digital-Forensics and Watermarking, 8389, 376–389.
- Johnson, N. F., Duric, Z., Jajodia, S., & Memon, N. (2001). Information hiding: Steganography and watermarking-attacks and countermeasures. *Journal of Electronic Imaging*, 10(3), 825. doi:10.1117/1.1388610
- Lei, B., Soon, I. Y., & Li, Z. (2012). A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition. *Signal Processing*, 92(9), 1985–2001. doi:10.1016/j.sigpro.2011.12.021
- Lu, X., Ye, C., & Wu, X. (2007). Design of lsb method based on energy feature. *Market Modernization*, (11), 37–38.
- Nehete, S., Sawarkar, S. D., & Sohani, M. (2011). Digital audio steganography using DWT with reduced embedding error and better extraction compared to DCT. *ICwet'11 International Conference & Workshop on Emerging Trends in Technology*, 167–168. doi:10.1145/1980022.1980059
- Premalatha, P., Narayanan, K. V., & Ramesh. (2014). Steganography in audio signals using variable bit replacement method in dct domain. Eersa Publications.
- Roy, S., Parida, J., Singh, A. K., & Sairam, A. S. (2012). Audio steganography using LSB encoding technique with increased capacity and bit error rate optimization. *International Conference on Computational Science, Engineering and Information Technology*, 372–376. doi:10.1145/2393216.2393279
- Rupanshi, Preeti, & Vandana. (2014). Audio steganography by direct sequence spread spectrum. *International Journal of Computer Trends and Technology*, 13(2).
- Sheikhan, M., Asadollahi, K., & Shahnazi, R. (2011). Improvement of embedding capacity and quality of dwt-based audio steganography systems. *World Applied Sciences Journal*, 13(3), 507–516.
- Shirali-Shahreza, S., & Manzuri-Shalmani, M. T. (2007). Adaptive Wavelet Domain Audio Steganography with High Capacity and Low Error Rate. *International Conference on Information and Emerging Technologies*, 1–5. doi:10.1109/ICIET.2007.4381305
- Sweldens, W. (1996). The lifting scheme: A custom-design construction of biorthogonal wavelets. *Applied and Computational Harmonic Analysis*, 3(2), 186–200. doi:10.1006/acha.1996.0015

Tan, L., Wu, B., Liu, Z., & Zhou, M. T. (2010). An audio information hiding algorithm with high capacity which based on chaotic and wavelet transform. *Tien Tzu Hsueh Pao*, 38(8), 1812–1811.

Tewari, T. K., Saxena, V., & Gupta, J. P. (2014). A digital audio watermarking scheme using selective mid band dct coefficients and energy threshold. *International Journal of Speech Technology*, 17(4), 365–371. doi:10.1007/s10772-014-9234-8

Vimal, J., & Alex, A. M. (2014). Audio steganography using dual randomness LSB method. *International Conference on Control, Instrumentation, Communication and Computational Technologies*, 941-944. doi:10.1109/ICCICCT.2014.6993093

Wu, Q., & Wu, M., & E.I.T. (2016). A new method of voice information hiding based on wavelet transform. *Dianzi Yu Xinxu Xuebao*, 38(4), 834–840.

Yan, B., Sun, S. H., & Lu, Z. M. (2003). *Improved echo hiding using power cepstrum and simulated annealing based synchronization technique*. Academic Press.

Zhang, M. (2009). *Research on digital audio watermarking technology* (PhD thesis). Yan shan University.

Hanlin Liu was born in China, in Aug 1993, received the B.Sc. degree with honor rank in Computer Application, China in 2015. He is now a postgraduate student at National University of Defense Technology, Hefei, P. R. China. His areas of interests are multimedia security, information hiding.

Jingju Liu was born in China, in 1974, received the B.Sc. degree with honor rank in Computer Application, China in 1996 and M.Sc. degree in Computer Application in 2001 from Hefei electronic engineering institute. She now is a professor at National University of Defense Technology, Hefei, P. R. China. Her areas of interest is information security.

Xuehu Yan was born in China, in Feb 1984, received the B.Sc. degree with honor rank in Science in Information & Calculate Science, China in 2006, M.Sc. degree in Computational Mathematics in 2008, and doctoral degree in Computer Science and Technology in 2015 from Harbin Institute of Technology. He now is an Associate Professor at National University of Defense Technology, Hefei, P. R. China. His areas of interests are visual cryptography, secret image sharing, information hiding, cryptography and multimedia security.

Pengfei Xue was born in China, in Aug 1989, received the B.Sc. degree with honor rank in Computer Application, China in 2012, M.Sc. degree in Information Security in 2015 from Hefei electronic engineering institute. He now is a PhD candidate at National University of Defense Technology, Hefei, P. R. China. His areas of interests are multimedia security, information hiding, steganography.

Dingwei Tan was born in China, in Mar 1991, received the B.Sc. degree with honor rank in Information Engineering, China in 2013. He now is a postgraduate at national university of defense technology, Hefei, P. R. China. His areas of interests are cryptography, audio steganography and steganalysis.