Using Phase Coding Method for Audio Steganography with the Stream Cipher Encrypt Technique

Made Sutha Yadnya[1(✉)], Bulkis Kanata[2], and M. Khaerul Anwar[2]

[1] Electrical Engineering Department, University of Mataram, Mataram, Indonesia
msyadnya@unram.ac.id
[2] Engineering Department, University of Mataram, Mataram, Indonesia
uqinata@yahoo.co.id

Abstract. One of the most popular audio file formats suitable for hiding information is Windows Audio-Visual (WAV). The two main areas of modification in a WAV file for data embedding are the storage environment and digital representation of the signal that will be used. In order to conceal secret messages successfully, a variety of methods for embedding information. The proposed system consists of four steps of steganography techniques; encoding, embedding, extraction and decoding. The trend of the test results from the 10 audio tested, showing a decrease in the PSNR value along with the increase in the size of the embedded message. This is in accordance with the theory that the greater the embedded information, the smaller the PSNR generated the steganography system using the phase coding method with the addition of stream cipher encryption has been successfully implemented. The system works well, as indicated by the success of the embedding and extraction tests, as well as the test results from BER. The PSNR test itself gets good results when the size of the embedded message is not more than 45% of the cover size.

Keywords: Steganography · Encoding · Embedding · Extraction · Decoding

1 Introduction

Technological developments in the era of digitalization have led to a very rapid increase in internet use. The demands of the need and the ease of internet access are several factors that cause people to have a tendency to always use the internet for their daily needs, especially the ease of communicating and exchanging information, both private and public. However, there is also the possibility of various digital crimes such as theft of information or wiretapping of communications by irresponsible parties. It causes the need for continuous technology updates that can maintain the confidentiality of information on data before or after it is stolen, and steganography is one of the solutions for that.

Steganography is the art of hiding and transmitting data through apparently innocuous carriers in an effort to conceal the existence of the data. The word steganography, as

I G. P. Suta Wijaya et al. (Eds.): MIMSE-I-C 2022, ACSR 102, pp. 66–75, 2022.
https://doi.org/10.2991/978-94-6463-084-8_8

derived from Greek, literally means covered or hidden writing and includes a vast array of methods of secret communication that conceal the very existence of the message. Steganography can be viewed as akin to cryptography. Both have been used throughout recorded history as means to add elements of secrecy to communication. Cryptographic techniques "scramble" a message so that if it is intercepted, it cannot be understood. This process is known as encryption and the encrypted message is sometimes referred to as ciphertext. Steganography, in essence, "camouflages" a message to hide its existence and make it seem "invisible", thus concealing the fact that a message is being sent altogether. A ciphertext message may draw suspicion while an invisible message will not.

The onset of computer technology and the Internet has given new life to steganography and the creative methods with which it is employed. Computer-based steganographic techniques introduce changes to digital carriers to embed information foreign to the native carriers. Steganography encompasses methods of transmitting secret messages in such a manner that the existence of the embedded messages is undetectable. Carriers of such messages may resemble innocent sounding text, disks and storage devices, network traffic and protocols, the way software or circuits are arranged, audio, images, video, or any other digitally represented code or transmission. These provide excellent carriers for hidden information and many different techniques have been introduced [1].

Audio steganography is a concealment technique that uses an audio file as a cover object called carrier audio to carry a secret message transmitted by modifying the audio signal in such a way. Audio steganography takes advantage of the limitations found in hearing (human auditory system) that human hearing cannot distinguish between a low and loud sound, where a loud sound can drown out a quieter sound. In steganography, the audio format has advantages over the format of images and videos. Audio files are usually relatively larger in size compared to image format, so that they can accommodate a larger number of secret messages. As for the video format, the size is relatively very large, but it reduces its practicality and also lacks algorithms to support this format [2].

One of the most popular audio file formats suitable for hiding information is Windows Audio-Visual (WAV). The two main areas of modification in a WAV file for data embedding are the storage environment and digital representation of the signal that will be used. In order to conceal secret messages successfully, a variety of methods for embedding information in digital audio have been presented by different researchers [3]. Encoding methods that are usually used to hide data in this format, i.e.: low-bit encoding, phase coding, spread spectrum and echo data hiding [4].

The phase coding method works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is adjusted in order to preserve the relative phase between segments. Phase coding is one of the most effective coding methods in terms of the signal-to-perceived noise ratio. The procedure for phase coding is as follows [5, 6]:

1. The audio signal is broken down into segments with a signal length equal to the size of the bit to be encoded.
2. The Fast Fourier Transform (FFT) is applied to each segment to create a matrix of the phases and magnitudes.
3. S+qtore the phase difference between adjacent segments

4. Phase shifts between consecutive segments are easily detected. In other words, the absolute phases of the segments can be changed but the relative phase differences between adjacent segments must be preserved. Thus, the secret message is only inserted in the phase vector of the first signal segment as follows:

$$phase\_new = \begin{cases} \pi/2 \ if \ message \ bit = 0 \\ -\pi/2 \ if \ message \ bit = 1 \end{cases} \quad (1)$$

5. A new phase matrix is created using the new phase of the first segment and the original phase differences.
6. Using the new phase matrix and the original magnitude matrix, the audio signal is reconstructed by applying the IFFT and by re-concatenating the audio segments.

Symmetric cryptosystem or private-key cryptosystem, both sender and recipient share the same secret key. During the encryption process, the sender encrypts the secret message with the key. The recipient uses the same key to recover the secret message. Systems consist of two categories, which are stream ciphers and block ciphers.

Stream cipher is a symmetric cipher where binary messages combine with a pseudorandom key stream using a combination operation exclusive-or (XOR). A secret key is the initial state of a keystream. The key together with the initial vector are inserted into a pseudorandom byte generator that produces an arbitrary length of the keystream. The plaintext is also converted into a stream of bits. Combination operation XOR of key stream and plaintext stream output the ciphertext stream. Stream cipher requires a shorter time to encrypt a message than block cipher. Therefore, the stream cipher is suitable for operations that need fast encryption [7].

In this research, a combination of phase coding as a steganography system and stream cipher as an encryption system was carried out. This is intended to produce an audio steganography system that has good quality in terms of increasing various parameters. As is well known, phase coding is efficient in terms of PSNR, while stream ciphers can increase the robustness and security of the steganographic system itself.

2 Method

The proposed system consists of four steps of steganography techniques; encoding, embedding, extraction and decoding.

2.1 Encoding Algorithm

Encoding is purposed to increase the security of messages that will be embedded into audio. Besides that, there is also a process of spreading message bits which could increase the robustness of the steganography file. The algorithm for encoding is as follows:

1. Choose a message to be embedded in the form of txt format.
2. Converts each character in the text into a binary format, which will get output in the form of message bits.

3. Set the spreading coefficient.
4. Spreading the message bits according to the spreading coefficient factor.
5. Generating Pseudo-noise using Pseudorandom Number Generator (PRNG).
6. XOR operation of message bits with Pseudo-noise.
7. The final result of this encoding process is in the form of data bits which will later be embedded into an audio file (WAV).

2.2 Embedding Algorithm

The embedding process is done by inserting the data bit (which has previously been through the encoding process) into the first segment of the audio as a cover. The algorithm for embedding is as follows:

1. Choose audio as cover and data bits to be embedded in.
2. Get data bits length.
3. Separate the segment of audio based on data bits length.
4. Fast Fourier Transform (FFT) processing on each segment to form a phase and magnitude matrix.
5. Calculates the phase difference between adjacent segments.
6. Converts data bits into phase form.
7. The message phase is embedded in the phase of the first audio signal segment by doing a phase shift.
8. Create a phase matrix using the new phase from the first segment and the absolute phase difference.
9. Inverse Fast Fourier Transform (IFFT) reconstruction on a new phase and magnitude matrix to form new audio segments.
10. Re-concatenating audio segments to form new audio file.

2.3 Extracting Algorithm

The extraction process is carried out to retrieve the data bits that have previously been embedded in the audio. The algorithm for extracting is as follows:

1. Get the audio file from the previous embedding process.
2. Separate audio segments based on predefined length.
3. Fast Fourier Transform (FFT) processing on each segment to form a phase and magnitude matrix.
4. Reconstruct the first segment phase to get the embedded phase.
5. Convert phase to data bits.
6. The final result of this encoding process is in the form of data bits which will later go through the decoding process to retrieve the message that was previously embedded.

2.4 Decoding Algorithm

The decoding process converts data bits into characters so that messages or information previously hidden by the sender can be read by the recipient. The algorithm for decoding is as follows:

1. Read the previously resurrected Pseudo-noise data as the key.
2. XOR operation of data bits with Pseudo-noise to retrieve the message bits.
3. De-spreading message bits with the same spread coefficient factor as previously used.
4. Convert message bits to char format.
5. The final result of the decoding process is in the form of a txt file as previously embedded at first.

3 Implementation and Result

Implementation is conducted using MATLAB software after confirming the steps of the steganography algorithm. Some audio is prepared to observe the system's performance, testing various parameters and comparisons against other systems.

3.1 Embedding and Extracting

Before testing the quality of the parameters generated by the system, it is necessary to first see whether the system is running successfully or not.

3.2 Peak Signal to Noise Ratio

This test is carried out to observe the audio quality after a message is embedded. PSNR testing is also one of the criteria that must be considered in research on steganography.

In this research, various PSNR tests were carried out on 10 audio samples under 2 different conditions. First, the condition of the size of the embedded message and the audio segment used is constant, with the purpose of observing the effect of segmentation on the audio cover. Second, by increasing the message size gradually, with the purpose of seeing the size limit of the embedded message on the resulting PSNR quality.

Table 1 shows the results of the message embedding and extracting process that has been carried out on the audio file. From the 10 audio covers with different sizes, it can be seen that the embedding and extracting process was successfully carried out on all the audio. This is because the size of the message is smaller than the audio file used as the cover medium. The message that is embedded into each audio file is text with the size 25 KB, which contains several characters such as letters, numbers and spaces.

Table 1. Results of the message embedding and extracting process

No.	Cover (.wav)	Size (KB)	Embedding	Extracting
1.	Sample1	103	Success	Success
2.	Sample2	207	Success	Success
3.	Sample3	405	Success	Success
4.	Sample4	603	Success	Success
5.	Sample5	810	Success	Success
6.	Sample6	1,034	Success	Success
7.	Sample7	1,507	Success	Success
8.	Sample8	2,024	Success	Success
9.	Sample9	5,168	Success	Success
10.	Sample10	6,115	Success	Success

3.3 Ratio Bit Error Rate

This test is conducted by looking at the bit comparison between the embedded message and the extracted message, whether there are different bits (error) or not. BER testing is also one of the most important criteria that must be considered in research on steganography.

Table 2 shows the results of the BER test between embedded messages and extracted messages. The BER value is said to be very good if there are no errors at all. From Table 2. it can be observed that all the samples used got a BER value of 0. Based on this, it was concluded that the BER test carried out had very good results.

Table 2. Results of the BER test

No.	Embedded Message	Size (KB)	Extracted Message	Size (KB)	BER
1.	Message.txt	25	Stego1.txt	25	0
2.	Message.txt	25	Stego2.txt	25	0
3.	Message.txt	25	Stego3.txt	25	0
4.	Message.txt	25	Stego4.txt	25	0
5.	Message.txt	25	Stego5.txt	25	0
6.	Message.txt	25	Stego6.txt	25	0
7.	Message.txt	25	Stego7.txt	25	0
8.	Message.txt	25	Stego8.txt	25	0
9.	Message.txt	25	Tego9.txt	25	0
10.	Message.txt	25	Stego10.txt	25	0

Table 3. Results of the PSNR testing

No.	Cover (.wav)	Cover Size		Message Size (KB)	PSNR (dB)
		Total (KB)	Segment (KB)		
1.	Sample1	103	80	25	61.79
2.	Sample2	207	80	25	58.56
3.	Sample3	405	80	25	61.08
4.	Sample4	603	80	25	59.41
5.	Sample5	810	80	25	56.50
6.	Sample6	1.034	80	25	57.72
7.	Sample7	1.507	80	25	60.08
8.	Sample8	2.024	80	25	58.74
9.	Sample9	5.168	80	25	51.18
10.	Sample10	6.115	80	25	42.56

Table 3 shows the results of the PSNR testing of various audio sizes with the condition that the size of the embedded message and the audio segment used is constant. Based on Table 3, it can be observed that from the 10 audio samples tested, the smallest PSNR value was 42.56 dB and the highest was 61.79 dB with an average value of 56.76 dB. Based on these results, it can be concluded that the PSNR testing carried out got very good results.

The next PSNR test is to gradually increase the size of the embedded message on the cover segment, which has a constant value. The purpose of this research is to observe the PSNR quality limit based on the comparison of message and cover sizes used.

Table 4 shows the message embedding 5–45% of the maximum size of the media cover, getting PSNR test results above 40 dB, which is a very good result.

Table 5 shows the message embedding 5–35% of the maximum size of the media cover, getting PSNR test results above 40 dB, which is a very good result.

The same test was also carried out on 8 other audio, so that a total of 10 audio samples were tested from the smallest to the largest size. The resulting PSNR can be observed in Fig. 1.

Figure 1 show the trend of the test results from the 10 audio tested, showing a decrease in the PSNR value along with the increase in the size of the embedded message. This is in accordance with the theory that the greater the embedded information, the smaller the PSNR generated.

Table 4. Message embedding 5–45%

No.	Cover (.wav)	Cover Size (Bytes)	Message Size		PSNR (dB)
			(%)	(Bytes)	
1.	Sample1	26,460	5	1,323	73.25
2.	Sample1	26,460	10	2,646	70.13
3.	Sample1	26,460	15	3,969	64.37
4.	Sample1	26,460	20	5,292	62.08
5.	Sample1	26,460	25	6,615	61.25
6.	Sample1	26,460	30	7,938	60.92
7.	Sample1	26,460	35	9,261	58.53
8.	Sample1	26,460	40	10,584	52.04
9.	Sample1	26,460	45	11,907	48.18
10.	Sample1	26,460	49	12,965	38.01

Table 5. Message embedding 5–35%

No.	Cover (.wav)	Cover Size (Bytes)	Message Size		PSNR (dB)
			(%)	(Bytes)	
1.	Sample10	1,565,550	5	78,278	114.82
2.	Sample10	1,565,550	10	156,555	107.43
3.	Sample10	1,565,550	15	234,833	62.57
4.	Sample10	1,565,550	20	313,110	54.51
5.	Sample10	1,565,550	25	391,388	47.43
6.	Sample10	1,565,550	30	469,665	43.27
7.	Sample10	1,565,550	35	547,943	41.84
8.	Sample10	1,565,550	40	626,220	39.51
9.	Sample10	1,565,550	45	704,498	37.58
10.	Sample10	1,565,550	49	767,120	26.38

4 Conclusion

The steganography system using the phase coding method with the addition of stream cipher encryption has been successfully implemented. The system works well, as indicated by the success of the embedding and extraction tests, as well as the test results from

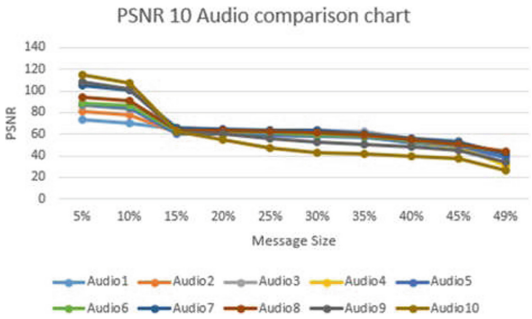

Fig. 1. Trend of the test results

BER. The PSNR test itself gets good results when the size of the embedded message is not more than 45% of the cover size.

Bibliography

1. Jhonson, N., Duric, Z. and Jajodia, S.: Information Hiding: Steganography and Watermarking - Attacks and Countermeasures. 1st ed. Virginia: SPRINGER-SCIENCE+BUSINESS MEDIA, LLC. pp. 1–6 (2001)
2. Mustakmal, M. E.: Audio Steganography with LSB Algorithm for Digital Data Security. Yogyakarta, pp.11–12 (2018)
3. S. Bhattacharyya, A. Kundu, and G. Sanyal.: A Novel Audio Steganography Technique by M16MA. Int. J. Comput. Appl., vol. 30, no. 8, pp. 26–34 (2011). https://doi.org/10.5120/3671-5113
4. B. Kuniadi, D. Puspitaningrum, and F. F. Coastera.: Design and Development of Text Message Steganography Applications on Digital Audio Using the Least Significant Bit Method. J. Rekursif, vol. 5, no. 3, pp. 285–297 (2017)
5. Rolasris.: Audio Watermarking Analysis Based on Discrete Wavelet Transform and Phase Coding Methods in Ambient Mode. vol. 3, no. 2, p. 52 (2016)
6. W. Bender, D. Gruhl, N. Morimoto, and A. Lu.: Techniques for data hiding. IBM Syst. J., vol. 35, no. 3–4, pp. 313–335 (1996). https://doi.org/10.1147/sj.353.0313
7. A. A. Alsabhany, F. Ridzuan, and A. H. Azni.: The Adaptive Multi- Level Phase Coding Method in Audio Steganography. IEEE Access, vol. 7, pp. 129291–129306 (2019)