I.J. Image, Graphics and Signal Processing, **2016, 1, 36-42**
Published Online January 2016 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2016.01.04

A Hybrid Technique for Enhancing the Efficiency of Audio Steganography

Ali M. Meligy
Department of mathematics, Faculty of science, Menoufia University, Egypt
E-mail: meligyali@hotmail.com

Mohammed M. Nasef and Fatma T. Eid
Department of mathematics, Faculty of science, Menoufia University, Egypt
E-mail: {mnasef81@yahoo.com or Mohammed_nasef@science.menofia.edu.eg, fatma_taher111@yahoo.com or fatma.taher@fa-hists.edu.eg}

Abstract—Steganography is the art or science that is used in secret communication. It means that there is a secret message that is hidden within another cover media. The cover media may be image, video or audio and the secret message may be any type of digital message. The hidden message doesn't have any relationship with the cover media where the cover media is just to protect the secret message from hacking by unauthorized receiver. The audio cover is used in this paper because of the higher sensitivity of the human auditory system (HAS) than the human visual system (HVS). In this paper, we proposed a hybrid technique to audio steganography. This technique is based on a hybrid between two techniques of audio steganography. These techniques are Least Significant Bit (LSB) technique and modification of phase coding. The hybrid between them is for improving the performance of the phase coding where the performance of it is very low. Audio steganography performance is measured by several factors, the most important one of them is Signal to noise ratio (SNR) which is used to compare the performance of our technique with some known techniques.

Index Terms—Steganography, Cryptography, Audio Steganography, phase coding, Least Significant Bit, Signal to noise ratio.

I. INTRODUCTION

Security system plays an important role in our lives. It is used for impedance, or protecting the secret message from undesirable people. Therefore, an access to the best way to protect it and its confidentiality is a must. There are a lot of techniques in the security topic such as, cryptography, watermarking and steganography, [3, 6, 8].

Cryptography is used to encrypt or protect the secret message. The secret message was scrambled and became gibberish to everyone. So, everyone knows that there is a secret message but cannot read it without knowing the extraction key. So, the extraction key must be very difficult to be known or used, [1, 8].

The other technique of security is **watermarking**. It is used in copyrights' protection and authentication of digital files. Watermarking may be visible or invisible (It depends on the reason why we use it). The main purpose of using watermarking is to prevent the illegal copying or codification of ownership of digital media, [3, 8, 12].

On the other hand, **steganography** is used for hiding the secret message inside another message which is used as a cover for this secret message. Steganography is derived from the Greek word steganos which means *covered* or *secret* and graphy means *writing* or *drawing*. So, no one except the sender and the desirable receiver would be able to doubt or predict the existence of the secret message. The secret message may be text, audio, video, image, html file ….etc and the cover message may be image, audio or video, [3, 6, 8, 11].

In the following Fig .1, there is a comparison between the three techniques (cryptography, steganography and watermarking) to clarify the advantages and disadvantages of every one. The comparison is done according to the definition, secret message, objective and security of every technique.

In this paper, steganography is used because it provides much more security when compared with cryptography. In steganography, there is no chance for any unintended user to know that a message is being sent. Whereas, in cryptography there will always be a suspicion that a message is being sent, [4, 7], hence these are more prone to be hacked or suppressed.

To achieve more effective steganography technique, the following standards should be followed: [7, 11]

1. Imperceptibility:

No one can distinguish between the original file and the stego file.

2. Secrecy:

The secret message can't be extracted without the private key.

3. High capacity:

All the data in the secret message must be embedded in the cover file without any skip of any data in it.

4. Resistance:

The secret message should not be affected by the

manipulation which is made in the stego file.

5. Accurate extraction:

The secret message shouldn't be out with any change when the receiver extracts it from stego file.

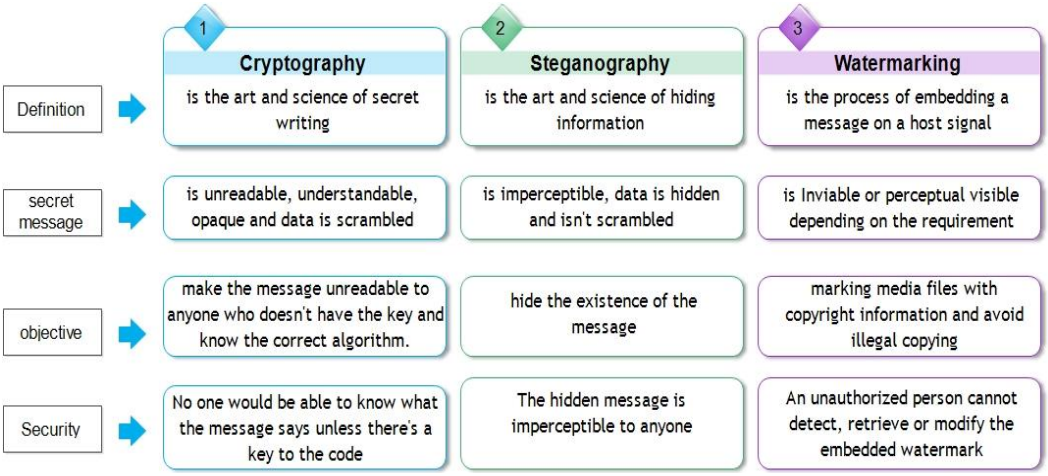

Fig.1. Comparison between Cryptography, Steganography and Watermarking

There are more types of steganography such as image steganography and audio steganography. In this paper, audio steganography is used where the secret message is hidden in digital audio file, [6]. The technique structure of audio steganography is clarified by using a general diagram in the following Fig .2. There are a lot of techniques that are used in audio steganography, such as, least Significant Bit (LSB), parity coding, echo coding and phase coding.

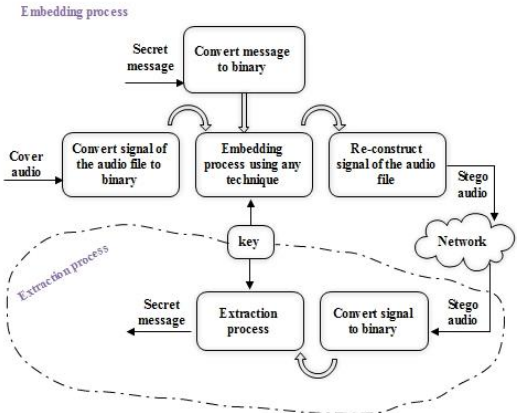

Fig.2. The General Diagram of the Audio Steganography Technique

Through this paper and after the introduction, the related work would be clarified and the idea of each one would be summarized. In section 2, the proposed technique would be explained and the embedding and extraction technique would be shown. In section 3, the results would be clarified and compared with some known algorithms. In section 4, the conclusion would combine a summary for this paper and finally the references that we depended on our work.

II. Related Work

F. Djebbar and B. Ayad, 2014, [5] proposed method that embeds high-capacity data in phase spectrum. The proposed method is based on the assumption that partial alteration of selected frequency bins in the phase spectrum leads to a smooth transition while preserving phase continuity. The cover-signal is divided into M frames of 4 ms, each contains N samples. The magnitude spectrum is isolated by transforming each frame to frequency domain using Fast Fourier Transform (FFT). The minimum and the maximum hiding band locations are specified. They only select high energy frequency components in an attempt to minimize the embedding distortion. The frequency bin is selected for data hiding only if its energy is higher or equal to the threshold value. Data is embedded along a chosen LSB layer (5th LSB layer at minimum). They benefit from the fact that noise that is 13 dB below the original signal spectrum for all frequencies is inaudible. Even though the frequency bins qualified for data hiding are selected in the magnitude spectrum. To embed in the phase spectrum, the frequency bins, in the phase, selected for data hiding are first defined in the magnitude through an election process and then mapped into the phase spectrum to embed data.

K. Kaur and D. Verma 2014, [10] proposed three different steganographic methods that have been used instead of using one steganographic method. Here, three secret messages rather than one, can be transmitted with a single cover file. Layering approach gives opportunity to do so. At the first level, cover file (C) can be embedded with the first secret message S1. Assuming the decoy file as C1 which is a cover file for next middle level where secret message can be denoted as S2. Assuming the decoy file as C12 which is a cover file for next lower level where secret message can be denoted as S3. Now the final stego file created as C123. So C123 holds three secret messages S1, S2 and S3. In this paper, three permutations of audio steganography methods are used for embedding the three secret message such as the first

secret message hided in level 1 under decoy object using LSB technique, second secret message hided in level 2 under decoy object using parity coding technique and third secret message hided in level 3 under decoy object using phase coding technique.

K. Cho, S.H Bae, I.K Choi, N.S Kim, and M. Unoki, 2013 [9] proposed an audio data hiding method based on the phase information of the modulated complex lapped transform (MCLT) coefficients. A host audio signal is divided into consecutive MCLT frames and the data bits are embedded by modifying the phases of the MCLT coefficients. Since each MCLT frame overlaps by half with the adjacent ones, the MCLT-based approach reduces the blocking artifacts which degrade the quality of the data-embedded audio signal. The modified MCLT coefficients are then converted into a time-domain signal segment by applying inverse MCLT and overlapping with the previous frame. The main strategy of data embedding is to modify the phase of MCLT coefficients of the host audio signal in such a way that the phase can be detected as either 0 or π at the data extraction procedure.

M. Nutzinger and J. Wurzer, 2011 [13], proposed an algorithm is based on the phase coding technique which embeds data in the phase spectrum of the frequency domain signal. They retain the original phase values in order to best keep the quality of the cover audio signal. Secret bits are embedded by introducing a configurable phase difference between selected chunks of blocks from the cover medium instead of discarding the original phase values and introducing a random phase like other approaches. At the beginning of the operation the original audio cover medium is split into blocks of same length, corresponding to a configurable amount of milliseconds. For processing, each block is transformed into the frequency domain by the FFT algorithm. Two further variables depict the frequency interval which is used for embedding and extraction, giving the interval.

D. M. Ballesteros, J. M. Moreno, 2013 [4] proposed a new scheme of data hiding which takes advantage of the masking property of the Human Auditory System (HAS) to hide a secret (speech) signal into a host (speech) signal. The embedding process is carried out into the wavelet coefficients of the speech signals. The main point of the proposed scheme is that the embedding process is suitable for real-time processing, and the secret's coefficients are relocated by an adaptive key, instead of a pseudo-noise sequence of some approaches. The latency of the embedding module makes this approach useful for real-time speech communication because the total delay added by the proposed system is low compared to the highest delay allowed for a high quality speech transmission.

P. Chandrakar, M. Choudhary and C. Badgaiyan, 2013 [14] proposed a method that used Audio Steganography using LSB algorithm to hide the message into multiple audio files. The message hidden by this application is less vulnerable to be stolen than other similar applications. This is due to following reasons: Firstly, multiple files are taken to hide high amounts of messages which enhance information hiding capacity. Secondly, before being hidden, the message is broken into parts and shuffled randomly based on permutation generated at runtime so even if the LSB gets encountered the message is still unarranged and meaningless which enhances its security.

III. Proposed Technique

Our proposed technique is based on a hybrid between LSB and the modification of phase coding. Most researches worked on LSB only or on phase coding only and every technique has advantages and disadvantages. The maximum SNR value is accessed using the phase coding was 32.31 and it's very bad result. So, we proposed a hybrid technique that intends to take the advantages of the two techniques to enhance the efficiency of audio steganography. We, firstly, clarified these techniques, the advantages and disadvantages of them, then, clarified our proposed technique and how a hybrid is done between them.

A. LSB and Phase coding techniques:

- ***LSB:***

This technique is based on embedding the secret message bits in the least significant bit of the cover audio file, [6, 11]. This technique is clarified in the following Fig .3. using a general diagram, [17]. Then, the major advantages and the disadvantages of this technique are clarified.

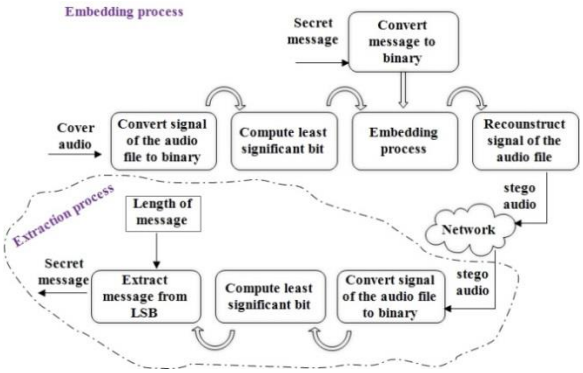

Fig.3. The General Diagram of the LSB Technique

The major Advantages of this technique:

- High rate of embedding bits.
- Low computational complexity of the algorithm compared with others techniques
- Low noise in the cover audio and may be not sensible.

The major disadvantages of this technique:

- Low robustness against the attacks and signal processing.
- Easy to extract and destroy.

- ***Phase coding:***

This technique is based on embedding the secret message bits in the initial phase of the cover audio file [6, 11]. This technique is clarified in the following Fig .4. using a general diagram, [17]. Then, the major advantages and the disadvantages of this technique are clarified.

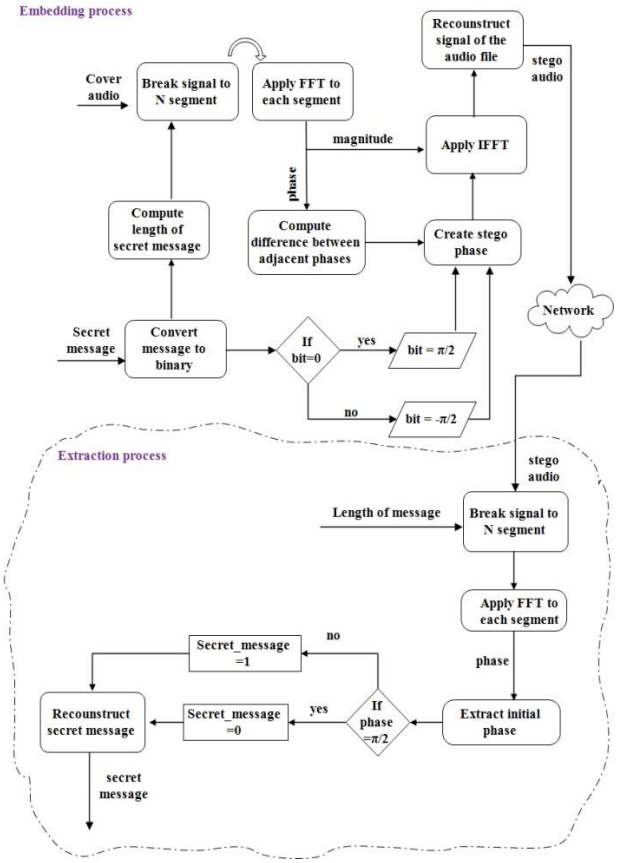

Fig.4. The General Diagram of the Phase Technique

The major Advantages of this technique:

- Robustness against signal modification

The major disadvantages of this technique:

- Low data transmission rate. It is used when only a small amount of data, such as a watermark, needs to be concealed.

B. Proposed Technique (hybrid between LSB and phase coding):

Our proposed technique depended on hybrid between LSB and phase coding techniques. This technique is meant to be different from the previous techniques by making a modification in the phase coding then using the LSB.

In the embedding process, the modification of phase coding depended on using all the audio signal and not breaking it into smaller segments. The steps of this process are explained in the following Fig .5.

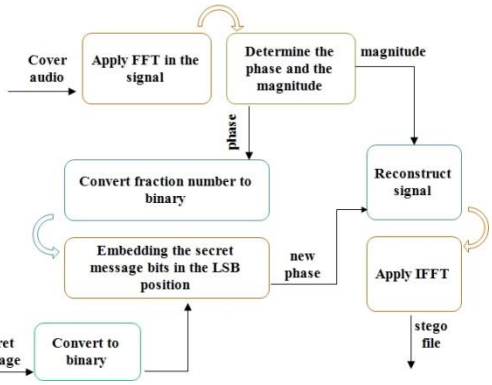

Fig.5. The General Diagram of the Embedding Process of the Proposed Technique

In the extraction process, the receiver must enter the length of message to extract the secret message from the stego file. The steps of this process are explained in the following Fig .6.

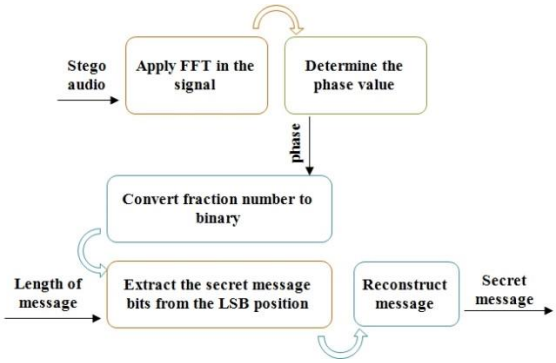

Fig.6. The General Diagram of the Extraction Process of the Proposed Technique

IV. RESULTS

The proposed technique is implemented using Matlab (2013a) program. We used several audio signals for embedding different types of secret message. The secret message is text, audio or image (color or gray). The efficiency of the stego file is compared with the original cover audio by determine the SNR (signal to noise ratio) value and the PSNR (peak signal to ratio) value.

The SNR is calculated by formula:

$$SNR = 10 * log_{10} \frac{\sum_{i=1}^{n} X^{2}(n)}{\sum_{i=1}^{n} [X(n) - Y(n)]^{2}} \quad (1)$$

The PSNR is calculated by formula:

$$PSNR = 10 * \log_{10} \frac{R^{2}}{MSE} \quad (2)$$

where,

$$MSE = \frac{\sum_{1}^{n}[X - Y]^2}{M*N} \quad (3)$$

In the previous equations, X is the original signal, Y is the stego signal, M and N are the numbers of rows and columns in the input signals and R is the maximum value of the signal, [4, 10].

In the proposed technique, the phase value is a floating number. We tried to re-quantize the values of the phase to convert it to integer but this occurred loss of phase information and got bad SNR values. The phase value is two parts (integer and fraction number). By testing, we found that embedding in the integer number using LSB led to create high noise in the audio signal. This because that the integer number of the phase is very small (i.e. is smaller than 4). So, we used the fraction number where it is more than 5 digits.

When we used from 5 to n digits of the phase fraction number, it got infinity value for SNR. This means that there is no different between the original and the stego audio and there was no noise affect in the original audio. However, the extraction message was damaged and this mustn't be occur. So, these values doesn't take in consideration as an evaluation.

In this section, we firstly clarified some examples for the extraction secret message (image, text and audio) from using our proposed technique. Then, we clarified the SNR and PSNR values of using 3 and 4 digits. Then, we clarified the comparison between our technique result and some knows techniques.

A. Some examples of the extraction message using our proposed technique:

In Fig .7, Fig .8 and Fig .9, there are examples for the extraction secret message (image, text and audio) from using 4 and 3 digits of the phase fraction number. To clarify the difference of extraction messages, In Fig .7. "Lena.bmp" is used as a secret message, In Fig .8. A small text from this paper is used as a secret message and In Fig .9. The secret message is an audio file and it is represented visually by time-domain plot.

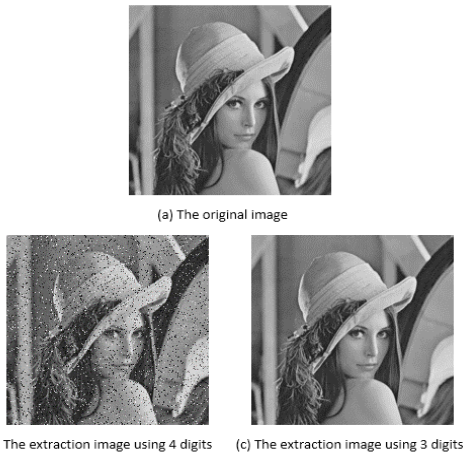

Fig.7. Comparison between the Extraction Gray Image Message Using 4 And 3 Digits of the Fraction Number of the Stego Phase

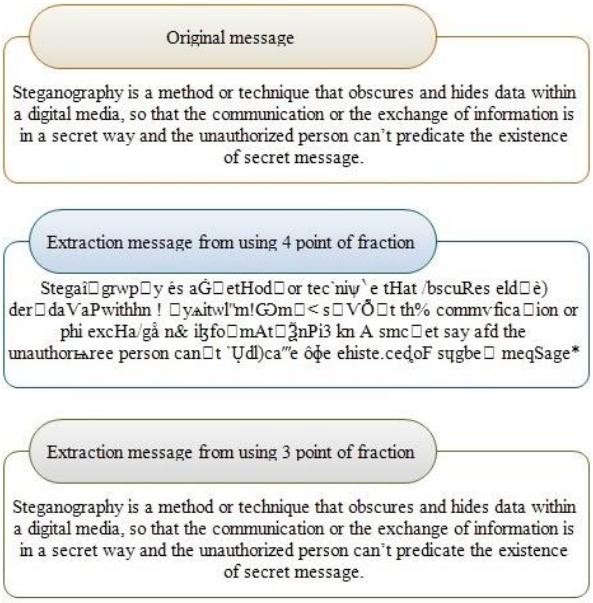

Fig.8. Comparison between the Extraction Text Message using 4 and 3 Digits of the Fraction Number of the Stego Phase

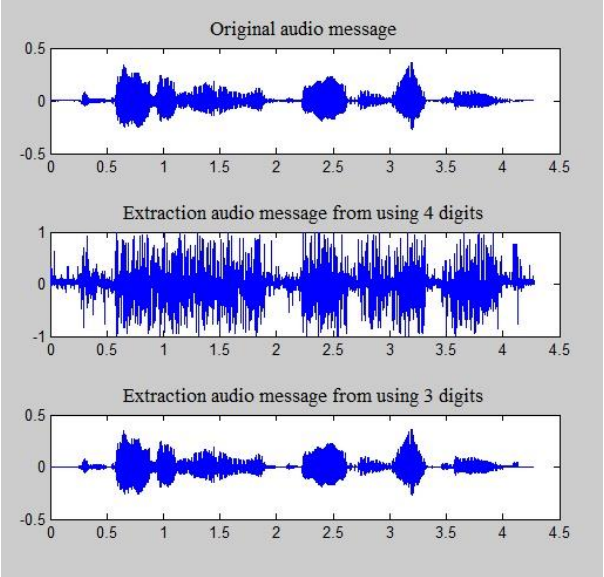

Fig.9. Comparison between the Extraction Audio Message using 4 and 3 Digits of The Fraction Number of the Stego Phase by Representing the Audio Wave of Secret Message in Time-Domain

B. Numerical Results:

We calculated the SNR and PSNR values of set of audio signals. We do this for four types of secret message which also have different length. These secret messages are text, audio, gray image and color image. SNR is calculated by applying equation (1) and PSNR is calculated by applying equation (2). Then, we calculated the average value of SNR and PSNR for each type of the secret message.

In the following Table 1. we presented the type of the secret message, these messages are a random samples, which embedded in the audio and the capacity of each one. Then, we presented the SNR and PSNR values of each type where these values are the average values from calculating the SNR and PSNR of set of audio signals.

In Table 2. we compared between our proposed technique and some known algorithms using SNR values. These known algorithms is only depend on one technique may be Phase coding or LSB. Although, we concluded from this comparison that our technique enhanced the efficiency of audio steganography and our technique is better than using phase coding only.

Table 1. The Results of SNR and PSNR Calculation for Embedding Text, Audio, Gray Image or Color Image using Our Proposed Technique

Type of the secret message	Embedded bytes	Embedding in the 3 digit of the phase fraction		Embedding in the 4 digit of the phase fraction	
		SNR	PSNR	SNR	PSNR
Text	36337	67.593975	187.88659	94.76241	215.595
Audio	38632	66.877725	187.17034	93.03729	213.8699
Gray image	80000	66.42215	186.71476	94.39743	215.23
Color image	227448	67.26055	187.5532	94.59339	215.4259

Table 2. Comparison between SNR Values of Proposed Technique with Some Known Algorithms

Researchers/algorithm	Technique	SNR
F. Djebbar and B. Ayad,[5]	Phase coding	32.31
S.S. Divya and M. R. Mohan Reddy [15]	LSB	55.37
M. Nutzinger and J. Wurzer [13]	Phase coding	30
K.P.Adhiya ,S.A. Patil,[11]	LSB	68
S.K. Bandyopadhyay, B. Datta, [16]	LSB	54.7
D.Pal, N.Ghashol,[2]	LSB	64.4402
Proposed technique	**Phase & LSB**	**67.0386**

V. Conclusion

In the proposed technique, we hybrid between two techniques (LSB and modification of phase coding). The major disadvantages of the LSB have been dealt with by using the phase coding to increase the robustness of the LSB. Also, the major disadvantages of the phase coding were addressed by using the whole signal and were not based only on the initial phase to increase the capacity of the embedding area of the phase coding. The proposed technique have been applied in 3, 4, 5 to n digits of the phase fraction number and the SNR and PSNR values have been calculated. From these values, we found that using the 3 digits of the fraction number is better than others. This is because that using 3 digits did not damage the secret message in the extraction process like 4, 5, ……, n digits. Also, it did not make a perceptual effect in the cover file. From The SNR values, we found that the results of this technique is better than the results of using these two techniques individually.

References

[1] A. Joseph Raphael Dr. V. Sundaram, Head & Director, "Cryptography and Steganography – A Survey", International Journal of Computer Technology and Applications, vol. 2 (3), May-June 2011.

[2] D. Pal and N. Ghashol, "A robust audio steganography scheme in time domain," International Journal of computer Applications, vol.80 (15), October 2013.

[3] D. M. Ballesteros and J. M. Moreno, "Highly transparent steganography model of speech signals using Efficient Wavelet Masking", Expert Systems with Applications, vol. 39 (10), August 2012.

[4] D. M. Ballesteros, J. M. Moreno, "Real-time, speech-in-speech hiding scheme based on least significant bit substitution and adaptive key", Computers and Electrical Engineering, vol. 39 (4), May 2013.

[5] F. Djebbar and B. Ayad, "Audio Steganograpgy by Phase Modification" The Eighth International Conference on Emerging Security Information, Systems and Technologies, 2014.

[6] F. Djebbar, B. Ayad, K. A Meraim and H. Hamam," Comparative study of digital audio steganography techniques", Journal on Audio, Speech, and Music Processing, 2012.

[7] H.I. Shahadi, R. Jidin and W.H. Way, "Lossless Audio Steganography based on Lifting Wavelet Transform and Dynamic Stego Key," Indian Journal of Science and Technology, Vol 7-No. 3, March 2014.

[8] H. V. Desai, "Steganography, Cryptography, Watermarking: A Comparative Study", Journal of Global Research in Computer Science, vol. 3 (12), December 2012.

[9] K. Cho, S.H Bae, I.K Choi, N.S Kim, and M. Unoki, "Robust Audio Data Hiding Method Based on Phase of Modulated Complex Lapped Transform", Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing IEEE, 2013.

[10] K. Kaur and D. Verma, "Multi-Level Steganographic Algorithm for Audio Steganography using LSB, Parity Coding and Phase Coding Technique", International Journal of Advanced Research in Computer Science and Software Engineering, vol. 4 (1), January 2014.

[11] K. P.Adhiya and S.A. Patil,"Hiding Text in Audio Using LSB Based Steganography" Information and Knowledge Management, vol. 2 (3), 2012.

[12] L. K. Saini1, V. Shrivastava, "A Survey of Digital Watermarking Techniques and its Applications", International Journal of Computer Science Trends and Technology (IJCST), vol. 2 (3), May-Jun 2014.

[13] M. Nutzinger and J. Wurzer, "A Novel Phase Coding Technique for Steganography in Auditive Media", Sixth International Conference on Availability, Reliability and Security, 2011.

[14] P. Chandrakar, M. Choudhary and C. Badgaiyan,

"Enhancement in Security of LSB based Audio Steganography using Multiple Files", International Journal of Computer Applications, vol. 73 (7), July 2013.
[15] S.S. Divya and M. R. Mohan Reddy, "Hiding Text In Audio Using Multiple LSB Steganography And Provide Security Using Cryptography", International Journal Of Scientific & Technology Research, vol. 1 (6), July 2012.
[16] S.K. Bandyopadhyay and B.Datta, "Higher LSB Layer Based Audio Steganography Technique" The International Journal on Electronics & Communication Technology IJECT, Vol. 2 (4), 2011.
[17] W. Bender, D. Gruhl, N. Morimoto, A. Lu, "Techniques for data hiding", IBM Systems Journal, vol. 35 (3&4), 1996.

Authors' Profiles

Ali M. Meligy is a professor of computer science at the Menoufia University in Egypt. Previously, he was the head of computer science and information technology departments at Al-Hussein Bin Talal University in Jordan. His research interests include parallel processing and applications, distributed systems, Petri nets, and reuse-based software engineering.

Mohammed M. Nasef was born in Egypt March 10th 1981. He received the M.Sc and Phd. Degree in computer science at the faculty of science, Menoufia University, Egypt in 2007 and 2011, respectively. His research interests include artificial intelligence, audio steganography, audio classification, and audio retrieval. Currently he is a lecturer of computer science in faculty of science, Menoufia University, Egypt. Member of the faculty projects for education development as DSAP and CIQAP. He is the manager of It-Unit, Faculty of science, Menoufia University since 2013 until now.

Fatma T. Eid is a demonstrator of computer science at Higher Future Institute for Specialized Technological Studies (Future Academy) in Egypt. She was born in Menoufia, Egypt, in 1991. She received the BSc degree in pure Mathematics and Computer Sciences in 2012, from the Faculty of Science, Menoufia University, Egypt.

How to cite this paper: Ali M. Meligy, Mohammed M. Nasef, Fatma T. Eid,"A Hybrid Technique for Enhancing the Efficiency of Audio Steganography", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.8, No.1, pp.36-42, 2016.DOI: 10.5815/ijigsp.2016.01.04

