Information Hiding in Audio Signals

H.B.Kekre
Sr. Professor,
Ph.D Research Scholar,
MPSTME, SVKM's NMIMS, Mumbai-56

Archana Athawale
Asst.Prof., Thadomal Shahani Engineering College,
Mumbai-50

Swarnalata Rao
MPSTME,
SVKM's NMIMS University,
Mumbai-56

Uttara Athawale
MCA Dept.,
Bharti Vidyapeeth's Institute of Mnmt & Info.Tech.,

ABSTRACT

Steganography is the practice of encoding secret information in indiscernible way. Audio Steganography is a technique used to transmit hidden information by modifying an audio signal in an imperceptible manner. It is mainly required for increasing security in transferring and archiving of audio files. Steganography complements rather than replaces encryption by adding another layer of security- it is much more difficult to decrypt a message if it is not known that there is a message. The basic idea of the paper is to present methods that hides information (audio, image and text) in cover audio using Least Significant Bit (LSB) coding method along with encryption so as to increase the security. Two novel methods have been proposed in this paper, one is considering parity of the digitized samples of cover audio and the other is considering the XOR operation. A novel method which is an extension to the XOR method that uses multiple LSB's for data embedding is also proposed. Experimental results are presented in this paper to demonstrate the effectiveness of the proposed methods. In addition, subjective listening tests are performed and the perceptual quality of the stego audio signal is found to be high.

Keywords

Information Hiding, Audio Steganography, Cryptography, Least Significant Bit (LSB) coding, Human Auditory System (HAS)

1. INTRODUCTION

By development of computer and the expansion of its use in different areas of life and work, the issue of security of information has gained special significance. One of the concerns in the area of Information security is the concept of hidden exchange of information [1]. Steganography is a sub-discipline of Information hiding that focuses on concealing the existence of messages [2]. The term hiding refers to the process of making the information imperceptible or keeping the existence of the information secret. Steganography is a word derived from the ancient Greek words *steganos*, which means *covered* and *graphia*, which in turn means *writing* [3]. The Eq. (1) provides a very generic description of the pieces of the steganographic process [4]:

cover_medium +hidden_data + stego_key = stego_medium (1)

In this context, the cover_medium is the file in which we will hide the hidden_data, which may also be encrypted using the stego_key. The resultant file is the stego_medium. Any steganography technique has to satisfy two basic requirements. The first requirement is perceptual transparency, i.e. cover object (object not containing any additional data) and stego object (object containing secret message) must be perceptually indiscernible. The second constraint is high data rate of the embedded data [5]. Unlike cryptography, which simply conceals the content or meaning of the message, Steganography conceals the very existence of a message [6].

Modern advances in computer, communication and signal processing have enabled the discovery of sophisticated techniques of steganography. These advances have broadened steganography's use to include various types of medium and various forms of information. The developed techniques allow text, audio, video, graphics, or codes to be concealed in electronic documents containing text, graphics, images and even in electronic audio or video files. Steganography has numerous applications like digital rights management, access control, covert communication, annotation etc.

Among the methods of Steganography, the most common one is to use images for applying steganography. Image steganography has been explored extensively with various steganographic schemes. Since nowadays, audio files are available everywhere and moreover; today's technology allows the copying and redistribution of audio files over the Internet at a very low or almost no cost. So it is necessary to have methods that confines access to these audio files and also for its security. Audio Steganography is one of the solutions. In Audio Steganography, the weakness of the Human Auditory System (HAS) is used to hide information in the audio [1]. That is, while using digital images as cover files the difficulty of the human eye to distinguish colors is taken advantage of, while using digital audio one can count on the different sensitivity of the human ear when it comes to sounds of low and high intensity; usually, higher sounds are perceived better than lower ones and it is thus easier to hide data among low sounds without the human ear noticing the alteration [7]. Audio Steganography is more challenging than Image Steganography because the human Auditory System (HAS) has more precision than Human Visual System (HVS) [1]. Audio Steganography has wide range of applications such as covert communication, digital watermarking, access control, digital rights management, etc. An effective audio steganographic scheme should possess the following three characteristics: Inaudibility of distortion (Perceptual Transparency), Data Rate (Capacity) and Robustness. These characteristics (requirements) are so called the magic triangle for data hiding and are contradictory [1].

In the past few years, several algorithms for the embedding and extraction of messages in audio sequences have been proposed. All of the developed algorithms exploit the characteristics of the human auditory system (HAS) in order to hide data into the host signal in a perceptually transparent manner [8].

The methods proposed in this paper combine the techniques of audio Steganography and cryptography, in order to make the message more secure. In these methods, the cover medium can be audio, image or video and the secret medium can be audio, image, video or text. In this paper, cover medium is an audio and the secret messages used are audio, image and text.

The rest of the paper is organized as follows: Section 2 explains in brief the common methods of audio Steganography, Section 3: explains the proposed methods, Section 4: gives experimental results and its discussion and Section 5: concludes the paper.

2. METHODS OF AUDIO STEGANOGRAPHY

Some commonly used methods of audio steganography are listed and discussed below in brief.

- Least Significant Bit (LSB) Coding
- Parity Coding
- Phase Encoding
- Spread Spectrum
- Echo Data Hiding

Least Significant Bit (LSB) Coding : One of the earliest techniques studied in the information hiding of digital audio (as well as other media types) is LSB coding. In this technique LSB of binary sequence of each sample of digitized audio file is replaced with binary equivalent of secret message [9]**.**

For example if we want to hide the letter 'A' (binary equivalent s **1000001**) into a digitized audio file where each sample is represented with 16 bits, then LSB of 7 consecutive samples (each of 16 bit size) is replaced with each bit of binary equivalent of the letter 'A' [10] .

Advantages: It is the simplest way to embed information in a digital audio file. It allows large amount of data to be concealed within an audio file, use of only one LSB of the host audio sample gives a capacity equivalent to the sampling rate which could vary from 8 kbps to 44.1 kbps (all samples used) [11]. This method is more widely used as modifications to LSBs usually not create audible changes to the sounds. Disadvantage: It has considerably low robustness against attacks.

Parity Coding [12]: Instead of breaking a signal down into individual samples, the parity coding method breaks a signal down into separate regions of samples and encodes each bit from the secret message in a sample region's parity bit. If the parity bit of a selected region does not match the secret bit to be encoded, the process flips the LSB of one of the samples in the region. Advantage: The sender has more of a choice in encoding the secret bit, and the signal can be changed in a more unobtrusive manner. Disadvantage: This method like LSB coding is not robust in nature.

Phase Coding [12]**:** Phase coding relies on the fact that the phase components of sound are not as perceptible to the human ear as noise is. It "works by substituting the phase of an initial audio segment with a reference phase that represents the data. The phase of subsequent segments is then adjusted in order to preserve the relative phase between segments". Disadvantage: It is a complex method and has low data transmission rate.

Spread Spectrum (SS) [9]: It attempts to spread out the encoded data across the available frequencies as much as possible. This is analogous to a system using an implementation of the LSB coding that randomly spreads the message bits over the entire sound file. However, unlike LSB coding, the SS method spreads the secret message over the sound file's frequency spectrum, using a code that is independent of the actual signal. As a result, the final signal occupies a bandwidth in excess of what is actually required for transmission. Advantage: It offers moderate data transmission rate while maintaining a high level of robustness. Disadvantage: It can introduce noise into a sound file.

Echo data hiding [9]: Text can be embedded in audio data by introducing an echo to the original signal. The data is then hidden by varying three parameters of the echo: initial amplitude, decay rate, and offset. If only one echo is produced from the original signal, then only one bit of information could be encoded.

3. PROPOSED METHODS

3.1 Considering Parity [6]:

This method uses LSB coding technique for data hiding in audio. However, instead of directly replacing LSBs of digitized samples with the message bits, it first checks the parity of the samples and then carries out data embedding. The process of data embedding and data retrieval are explained as follows:

Steps for Data embedding:

1. Read the cover audio signal.
2. Read the audio signal to be embedded, its size less than the size of the cover audio signal and convert it into binary sequence of message bits.
3. Depending upon the value of the message bit to be embedded (0/1), the LSB of the sample of the cover audio signal is modified or unchanged.
4. If the message bit to be embedded is 0, then the LSB of the sample of the cover audio signal is modified or unchanged such that the parity of the sample after embedding of this message bit is even.
5. If the message bit to be embedded is 1, then the LSB of the sample of the cover audio signal is modified or unchanged such that the parity of the sample after embedding of this message bit is odd.
6. The modified cover audio samples are then written to the file forming the stego audio signal.

Steps for Data Extraction/ Retrieval:

1. The Stego audio file is read.
2. The parity of every sample of the stego is checked.
3. If the parity is even, then the message bit retrieved is 0.
4. If the parity is odd, then the message bit retrieved is 1.
5. After every such 16 message bits are retrieved, they are converted to decimal equivalents.
6. Finally the secret message is reconstructed.

3.2 Using XORing of LSB's [6]:

This method performs XOR operation on the LSBs and then depending on the result of XOR operation and the message bit to be embedded, the LSB of the sample is modified or kept unchanged. The method described below performs XOR operation on first 2 LSBs. The XORing can be further expanded to 3 LSBs, 4 LSBs upto 16 LSBs so as to increase the level of encryption. The primary merit of the XOR operation is that it is simple to implement, and it is computationally inexpensive. The steps for data embedding and data retrieval are explained below and a tabular representation of the embedding procedure is as given in Table 1.

Steps for Data Embedding:

1. Read the cover audio signal.
2. Read the audio signal to be embedded, its size less than the size of the cover audio signal and convert it into binary sequence of bits.
3. Every message bit is embedded into the LSBs of the cover audio after processing.
4. Processing is done as follows:
   - If the message bit to be embedded is 0, then adjust or flip the LSB such that the XORing of LSB and next to LSB is 0.
   - If the message bit to be embedded is 1, then adjust or flip the LSB such that the XORing of LSB and next to LSB is 1.
5. The modified cover audio samples are then written to the file forming the stego audio signal.

Steps for Data Extraction/ Retrieval:

1. The Stego audio file is read.
2. Retrieval of the message bit is done by XORing the LSB and the bit next to LSB. If the result of XORing is 0, then the message bit is 0. If the result of XORing is 1, then the message bit is 1.
3. After every such 16 message bits are retrieved, they are converted to their decimal equivalents.
4. Finally the secret signal is reconstructed.

Table 1. Procedure for data embedding

LSB	Bit next to LSB	XOR	Action if message bit is 0	Action if message bit is 1
0	0	0	No Change	Flip LSB
0	1	1	Flip LSB	No Change
1	0	1	Flip LSB	No Change
1	1	0	No Change	Flip LSB

4. RESULTS AND DISCUSSIONS

Proposed methods were tested on music clips and speech samples. Clips were with varying sampling frequencies, mono audio files, represented by 16 bits per sample. Duration of the clips ranged from 2 to 8 seconds. The secret messages used for embedding were audio clips, image and text. The image used as secret message is of Lena of dimensions 64*64 and 128*128. The text files used for embedding purpose were taken from 2 famous personalities. The first text is the letter written by Abraham Lincoln to his son's teacher and the other text is the inspiring speech of Dr. Abdul Kalam Azad. The performance of the proposed methods is analyzed in terms of MSE (Mean Squared Error), PSNR (Peak Signal-to-Noise Ratio) and SNR (Signal-to-Noise Ratio). Subjective quality evaluation of these methods has also been carried out by performing listening tests involving ten persons. From the results of subjective tests, no difference has been found in the perceptual quality of the original audio files and their corresponding stego audio files.

Table 2 gives the results of the proposed method based on parity. The first three entries in the secret column are the secret audio clips followed by 2 text messages and the Lena image. Fig.1 shows one of the secret signal used in the parity method. Fig.2 shows the secret signal retrieved from stego. It is seen from both these figures that there is no difference between the original and the retrieved message, thereby assuring that the recovery is 100%. Fig. 3 shows the message retrieved when the LSBs of the stego signal are extracted directly. This indicates that the direct extraction of LSBs will only result in noise if embedding is done using parity method, thereby increasing security. Fig.4 shows the original secret image of Lena used for hiding. Fig.5 shows the retrieved image and Fig.6 is the plot of LSBs being extracted directly from stego which looks like noise.

Table 2. Results of Proposed method (considering parity)

Cover	Secret	MSE	PSNR	SNR
Guitar	Pingpong	4.17E-10	190.12	69.85
	Chimes	4.65E-10	189.65	69.38
	Newmail	4.65E-10	189.65	69.37
	Text1	7.44e-011	197.61	77.34
	Text2	2.45e-010	192.43	72.15
	Image	1.69E-10	194.03	73.76
Triangle	Pingpong	2.07E-10	193.16	69.32
	Chimes	4.65E-10	189.65	65.79
	Newmail	4.69E-10	189.61	65.76
	Text1	3.62e-011	200.73	76.88
	Text2	1.22e-010	195.47	71.62
	Image	3.35E-10	191.07	67.22
Speech	Pingpong	2.11E-10	193.08	62.72
	Chimes	4.65E-10	189.65	59.28
	Newmail	4.65E-10	189.65	59.28
	Text1	3.72e-011	200.62	70.25
	Text2	1.25e-010	195.35	64.99
	Image	3.41E-10	191	60.65
Average		**2.84E-10**	**192.92**	**68.08**

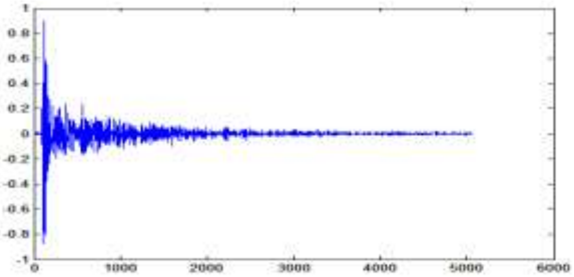

Fig. 1 Plot of secret signal used in parity method

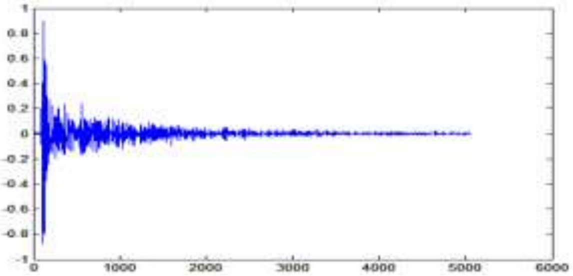

Fig. 2 Plot of secret signal retrieved using the parity method

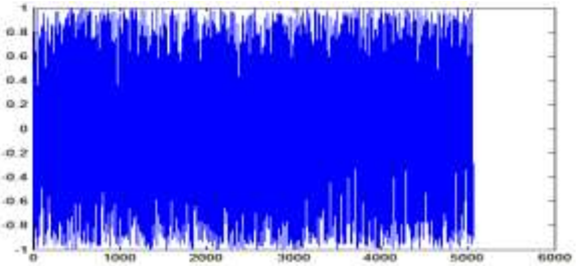

Fig. 3 Plot of signal retrieved by extracting LSBs directly which looks like noise

Fig.4 Original Secret Image

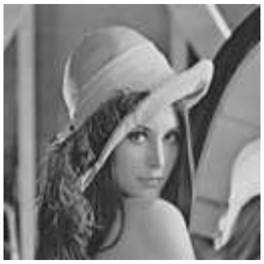

Fig.5 Retrieved Secret Image using Parity method

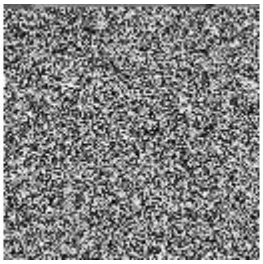

Fig. 6 Image retrieved by extracting LSBs directly which looks like noise

Table 3 gives the results of the proposed method based on XOR operation of the LSBs. It can be seen from these results and the results of the parity method, that the values of the performance parameters are quite similar.

Table 3. Results of Proposed method (using XORing of LSBs)

Cover	Secret	MSE	PSNR	SNR
Guitar	Pingpong	4.17E-10	190.12	69.85
	Chimes	4.67E-10	189.63	69.36
	Newmail	4.65E-10	189.65	69.38
	Text1	7.34e-011	197.67	77.40
	Text2	2.46e-010	192.42	72.14
	Image	1.68e-010	194.07	73.81
Triangle	Pingpong	2.07E-10	193.16	69.32
	Chimes	4.67E-10	189.63	65.78
	Newmail	4.60E-10	189.69	65.85
	Text1	3.66e-011	200.68	76.84
	Text2	1.24e-010	195.41	71.56
	Image	3.34E-10	191.08	67.23
Speech	Pingpong	2.12E-10	193.07	62.71
	Chimes	4.67E-10	189.63	59.26
	Newmail	4.64E-10	189.66	59.3
	Text1	3.76e-011	200.56	70.21
	Text2	1.24e-010	195.41	65.04
	Image	3.41E-10	191	60.64
Average		**2.84E-10**	**192.92**	**68.09**

In both of the proposed methods, only LSB had been used for embedding purpose either considering parity of the cover samples or by performing XOR operation on LSBs. But if we think of extending these approaches to increase the capacity of cover audio by utilizing more than just the LSB for data embedding, it may prove more beneficial. However, the approach based on parity cannot be extended to utilize multiple LSBs, but the other approach based on XOR method can be very well extended further to use multiple LSBs for data embedding.

One such approach of using 8 LSBs considering XOR operation has been discussed below. In this, XOR is performed on 16th bit and 8th bit, 15th bit and 7th bit, 14th bit and 6th bit, 13th bit and 5th bit, 12th bit and 4th bit, 11th bit and 3rd bit, 10th bit and 2nd bit, and 9th bit and 1st bit as shown in Fig.7.

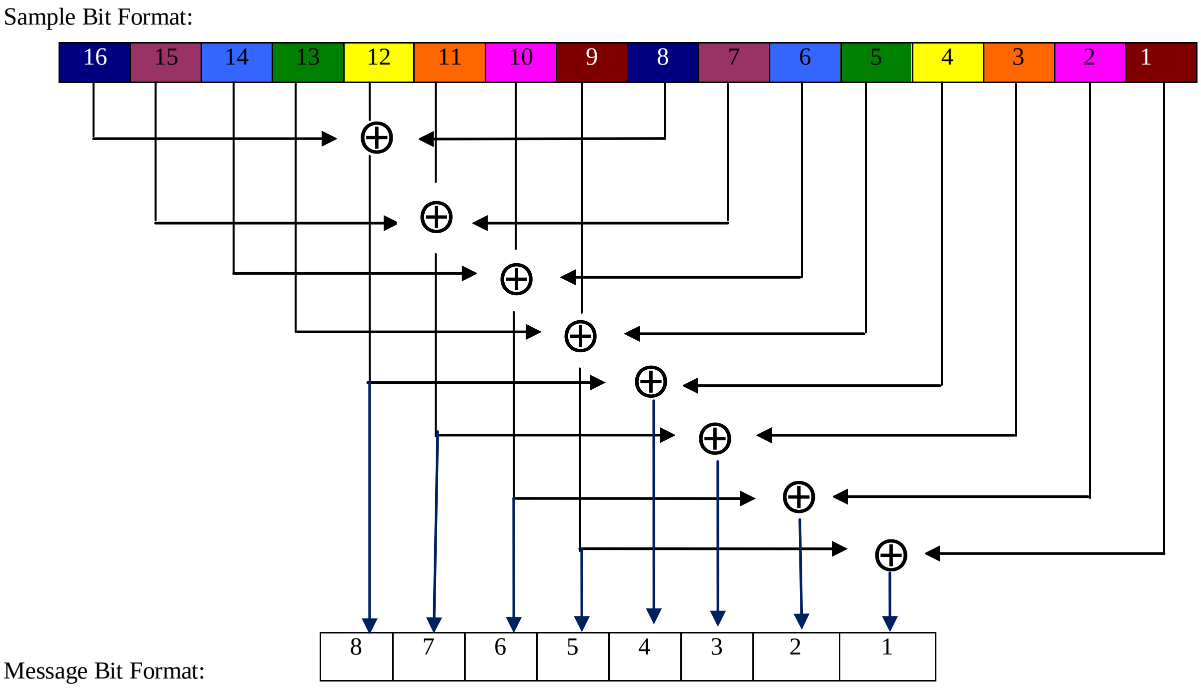

Fig.7 Diagrammatic representation of procedure for using XOR with multiple LSB's

Depending upon the result of XOR operation and the message bit to be embedded, the 8 LSBs can be used for data embedding. This increases the capacity of the cover for data embedding. To clearly understand the above mentioned approach, an example is presented below.

Consider the bits in the binary representation of a sample of cover audio and the message bit to be embedded is as given below. Table 4 gives the tabular representation explaining the procedure for data embedding using above approach.

Original Sample bits: 1000000000000001

Message bits: 10100010

Modified Sample bits: 1000000000**1**000**10**

Table 4. Data embedding procedure for multiple LSBs

Bit 1	Bit 2	XOR result	Message Bit	Action
1 (16)	0 (8)	1	1	No change
0 (15)	0 (7)	0	0	No change
0 (14)	0 (6)	0	1	Flip Bit 2
0 (13)	0 (5)	0	0	No change
0 (12)	0 (4)	0	0	No change
0 (11)	0 (3)	0	0	No change
0 (10)	0 (2)	0	1	Flip Bit 2
0 (9)	1 (1)	1	0	Flip Bit 2

In above table, the numbers in the brackets in the first 2 columns indicate the position of the bits of digitized samples of cover audio. As can be seen from the table, action takes place on bit2 of second column; it is either flipped or unchanged as shown in last column, as these bits forms the 8 LSBs.

The retrieval of bits is done by performing XOR operation on bits as done in the embedding process, and the result of the XOR operation will give the message bits back.

The MSE value for the given example where three bits have been changed comes out to be 9.39e-07. If we assume that all 8 bits are changed during the data embedding process, then the MSE value is 6.08e-05. However, all 8 bits being changed during embedding has very least probability of occurrence. Table 5 gives the results of the approach discussed above.

Table 5 Results of XOR method for multiple LSB's

Cover	MSE	PSNR	SNR
Guitar	1.49e-06	154.57	34.30
Triangle	8.88e-07	156.84	32.99
Speech	8.05e-07	157.27	26.91

The above discussed approach has been experimented on 3 cover signals. It can be seen from the above table that the MSE values are better than the estimated value which considers all 8 bits being changed during the embedding process.

5. CONCLUSION

The paper proposes two methods using LSB coding along with encryption to hide information (audio, image and text) in digital audio files. In the first method, the information is hidden by altering LSBs indirectly considering parity of samples of cover audio. In the second method, information is hidden by performing XOR operation on LSBs. In both these methods, direct LSB extraction will only result in noise. Thus, by using encryption along with steganography, these methods provide an additional level of security. From experimental results, it is seen that the proposed methods are effective. From listening tests, no difference is found between the original audio signal and the stego audio signal. The hidden information is recovered without any error. A novel idea which is an extension to the second proposed method based on XORing of LSBs that uses multiple LSBs has also been presented. This approach increases the capacity of the cover audio by as much as 8 times and also provides robust encryption. This will give great security and the embedded message cannot be extracted without the knowledge of the embedding process.

6. REFERENCES

[1] Shahreza S.S. and Shalmani M.T.M., “Adaptive wavelet domain audio steganography with high capacity and low error rate”, in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, pp: 1729 – 1732, 2008.

[2]“audio steg: overview”, Internet publication on www.snotmonkey.com *http://www.snotmonkey.com/work/school/405/overview.html*.

[3] Nedeljko Cvejic , “ Algorithms for audio watermarking and steganography”, *http://herkules.oulu.fi/isbn9514273842/isbn9514273842.pdf,*

[4] Gary C. Kessler, “Steganography: Hiding Data Within Data”, *http://www.garykessler.net/library/steganography.html*, September 2001.

[5] C. Parthasarathy and Dr. S.K.Srivatsa, “Increased Robustness Of Lsb Audio Steganography By Reduced Distortion Lsb Coding” 2005.

www.jatit.org/volumes/research-papers/Vol7No1/9Vol7No1.pdf,

[6] Dr.H.B.Kekre and A.A.Archana, “Information hiding using LSB technique with increased capacity”, *International Journal of Cryptography and Security*, vol. 1, No.2, October 2008.

[7] Nicola Cocchiaro, *http://stegui.sourceforge.net/intro.html*.

[8] Zamani M., Ahmad R.B., Manaf A.B.A., Zeki A.M., “An Approach to Improve the Robustness of Substitution Techniques of Audio Steganography”, *in Proc. IEEE International Conference on Computer Science and Information Technology*, ICCSIT pp: 5-9, 2009.

[9]“audio steg: methods”, Internet publication on www.snotmonkey.com *http://www.snotmonkey.com/work/school/405/methods.html*

[10] Soumyendu Das, Subhendu Das, Bijoy Bandyopadhyay and Sugata Sanyal, “Steganography and Steganalysis: different approaches”, *http://www.tifr.res.in/~sanyal/papers/Soumyendu_Steganography_Steganalysis_different_approaches.pdf*.

[11] N. Cvejic, T. Seppanen, "Increasing Robustness of LSB Audio Steganography using a novel embedding method", *in Proc. IEEE Int. Conf Info. tech.: Coding and Computing*, Vol. 2, pp.533-537, April 2004.

[12]Samir K Bandyopadhyay, Debnath Bhattacharyya, Debashis Ganguly, Swarnendu Mukherjee and Poulami Das, “A Tutorial Review on Steganography”.

http://www.jiit.ac.in/jiit/ic3/IC3_2008/IC3-2008/APP2_21.pdf

Author Biographies

Dr.H.B.Kekre has received B.E.(Hons.) in Telecomm. Engineering from Jabalpur University in 1958, M.Tech (Industrial Electronics) from IIT Bombay in 1960, M.S.Engg.(Electrical Engg.) from University of Ottawa in 1965 and Ph.D. (System Identification) from IIT Bombay in 1970. He has worked over 35 years as Faculty of Electrical Engg. and then HOD Computer Science and Engg. at IIT Bombay. From 13 years, he was working as a Professor and Head in the Department of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. He is currently Senior Professor working with MPSTME, NMIMS University, Mumbai. His areas of interest are Digital Signal Processing and Image Processing. He has more than 300 papers in National/ International Conferences/ Journals to his credit. He was Senior Member of IEEE. Presently he is Fellow of IETE and Life Member of ISTE Recently six students working under his guidance have received best paper awards and two of his students have been awarded Ph.D. degree by NMIMS University. Currently 10 research scholars are pursuing Ph.D. program under his guidance.

Ms. Archana Athawale has received M.E. (Computer Engg.) degree from V.J.T.I., Mumbai University in 1999, currently pursuing Ph.D. from NMIMS University, Mumbai. She has more than 9 years of teaching experience. Currently working as Assistant professor in Department of Computer Engg. at Thadomal Shahani Engineering College, Mumbai. Her area of interest is Image Processing, Signal Processing and Computer Graphics. She has 35 papers in National International Journals, IEEExplore, ACM, Springer, International and Natinal Conference to her credit.

Ms. B. Swarnalata G. Rao has received B.Tech (Computer Engg.) from Babasaheb Ambedkar Technological University, Lonere, Raigad in 2002. Currently pursuing M.E. (Computer Engg.) from Thadomal Shahani Engineering College, Mumbai. She has more than 6 years of teaching experience. She is currently working as Assistant professor in Department of Computer Engg. at MPSTME, NMIMS University, Mumbai.

Ms. Uttara Athawale has received M.C.A. degree in 2002 and M.Sc (Mathematics) degree in1994 from Shivaji University , She has 8 years of teaching experience. Currently working as Lecturer in Bharati Vidyapeeth Institute of Management and Information Technology CBD, Belapur Navi Mumbai. Her area of interest is Image Processing, Signal Processing and Computer Graphics. She has 1 paper in International Conference to her credit.