An Audio Steganography Scheme Robust to White Gaussian Noise Attack

Ei Thin Su
University of Computer Studies, Yangon
eithinsu.ucsy@gmail.com

Abstract

An innovative audio Steganography scheme based on optimize algorithm as Probabilistic Global Search Lausanne (PGSL) in modified discrete cosine transform domain (MDCT) is proposed. Inaudibility and robustness are two main important requirements for audio Steganography. The major contribution of the proposed system is to embed the secret data in the best MDCT coefficients against secret message detection from unauthorized users. Embed the secret data by employing original signal transformation with MDCT and best point searching process through PGSL Algorithm to achieve high inaudible and robustness. Results of the experimental test using clean utterances from Texas Instruments Massachusetts Institute of Technology (TIMIT) databases and music files show error-free data recovery from noise addition attack stego audio. Robustness of the system is evaluated with BER (bit error rate). Spectrogram and Informal listening test also showed that the distortions in the stego audio from the original cover speech and music files is imperceptible and inaudible even with high payload.

Keywords: MDCT, PGSL Algorithm, Secret data communication.

1. Introduction

Steganography is the science of concealing a covert message for a specific purpose [1] [2]. Recently, data hiding using Steganography techniques have been employed in covert communication of information [13-14], It is employed in copyright authentication [7] [10], fingerprinting [3], forensics [9] [12], broadcast monitoring [6], medical [4][5][9], land consolidation [11], commercial interests [15],[16],[17] and military applications [8], etc.

Today, fast improvement of the Internet and digital information revolution caused major changes in the overall culture. Information hiding using audio is one of the most effective ways to protect our privacy and secret communication [16]. Imperceptibility, robustness and capacity are the most important requirements for any data hiding scheme. The strength of secret data communication using data hiding lies inaudibility and imperceptibility to be unnoticeable by the human ear. Payload Capacity: Indicates the amount of hidden bits as much as possible. Robustness: Evaluates the audible distortion due to signal manipulation attacks that the correct secret data recovery rate in attacked stego audio.

According to the implementation process of data hiding process, secret data embedding domain can be divided into time and transform domain. Hidden bits are embedded directly into the time signal samples in time domain. Transform domain is considered better than that of the time domain in most data hiding scheme to withstand attacks.

Thus, proposed system considered transform domain and the transformation technique is done by applying MDCT. The result of our presented paper is organized as follows: Section 2 introduced related work of the proposed system. Problem Statements and Contributions are presented in Section 3.We briefly summarize the necessary background

Theory in Section 4. Data Embedding and Extraction scheme based MDCT and PGSL is described in Section 5, describe Evaluation Experiments and Results in Section 6, the paper is successfully concluded by expressing the application areas to apply proposed system in Section 7.

2. Related Work

Secret data communication using data hiding scheme is an active research area that has been strongly motivated to solve the privacy protection problem. The previous methods in data hiding based audio Steganography and watermark are Low-bit Encoding, Phase coding, Echo data hiding, Spread Spectrum coding, Psycho acoustical masking property of Human Auditory System and bit modification of audio sound files have also been proposed.

In ***Low-bit encoding***, the secret message is substituted with the least significant bit (LSB) of each sample of the audio cover file. This method is simple and can be used to embed larger messages, but cannot protect the hidden message from small modifications. Thus, it is rarely used in the real commercial applications. ***Phase coding*** is based on the phase components of sound as imperceptible which the differences between cover and stego audio to the human ear. Message bits are encoded as phase shifts in the phase spectrum of a digital signal [25] [26]. This leads to inaudible encoding. A characteristic feature of Phase coding is the low data transmission rate that the secret message is encoded only in the first segment of the audio signal. An increase in the length of the segment would have a ripple effect by altering the phase relations between the frequency components of the segment, thus making detection easier. Hence, the Phase Coding method is used only when a small amount of data embedded.

In ***Echo hiding***, secret message is embedded by introducing an echo into the discrete audio signal. Echo hiding allows for a higher data transmission rate than least significant bit insertion method [26] [27] [28] [29] [30] [31] [32]. ***Spread Spectrum coding*** method spreads the secret message across the frequency spectrum of the audio signal. It is robust than Phase coding and Low-bit coding but use complex algorithms [25]. In [20], the author proposed Genetic-Algorithm Based Approach for audio Steganography that have been withstand attacks like noise addition than simple LSB method. In [21], the method of audio Steganography by Cepstrum Modification is presented, which combines psycho acoustical masking property of HAS with the decorrelation property of speech cepstrum, that achieved high imperceptible embedding and correct data recovery but low payload. In bit modification of audio sound files, embedding and extraction procedure are identical in most previous system till now. No new innovation.

The method of [22] has proposed increasing the capacity of LSB-based audio Steganography using a Novel Embedding Method. They described that robustness against noise addition or MPEG compression better than standard LSB method. In [23], the author presented Telephony Speech Enhancement by Data Hiding to improve intelligibility and perceived quality of telephone speech. Data is hiding based on the perceptual masking principle; the inaudible spectrum components within the telephone bandwidth can be removed without degrading the speech quality. They showed that robust to quantization errors and channel noises.

Our proposed system aims to reduce the distortion over the host cover audio due to data embedding and enhances fidelity. Covert data is hidden in the best coefficients of MDCT domain on cover speech through optimization based problem transformation method through PGSL algorithm.

3. Contributions

The main contribution of the proposed system is searching the best position by using optimize PGSL algorithm to embed secret data with secure and against unauthorized detection of the secret message. PGSL can support minimum searching time. Thus the system save time complexity and meet low computation time in

data embedding. PGSL algorithm of the global searching technique wasn't use in these problem domains till now. This approach is the new in the literature of information hiding for audio watermark, audio Steganography and multimedia content protection applications.

4. Background Theory

MDCT domain and PGSL Algorithm used in Data Embedding and Extraction of proposed system are represented detail as follows:

4.1. Modified Discrete Cosine Transform

Modified discrete cosine transform (MDCT) domain is used to overcome reconstruction error in data embedding process. MDCT is type-IV discrete cosine transform (DCT-IV).MDCT provides better energy compaction than DCT. MDCT can be directly estimated the capacity of a transformation scheme by its ability to pack input data into as few coefficients as potential.

MDCT guarantee without audible distortion in the reconstructed audio. It avoids errors between block boundaries. It has the principle of time-domain aliasing cancellation (TDAC) [18] [19]. So it can be done perfect reconstruction. Because of this advantage, we can fulfill imperceptible issues that the differences between the original and stego audio in the proposed scheme by applying MDCT transform-domain in general. Thus, MDCT is selected as the secret data embedding and extraction domain in our proposed system.

4.2. PGSL Algorithm

PGSL (Probabilistic Global Search Lausanne Algorithm) is the global optimization algorithm**.** It has been developed by Raphael and Smith in 2002. It is applied in the field of structural engineering to solve the optimization problems for bridge diagnosis. It is a direct search algorithm, to find the minimum of user defined objective function by using global sampling [24]. In PGSL, optimal solutions can be defined through focusing search around sets of good solutions. Tests on benchmark problems having multi-parameter non-linear objective functions exposed that PGSL is advanced algorithms and performs better than genetic algorithms. Moreover, PGSL performs better than other approaches even increased the problem size.

In the fields of diagnosis, design and control, PGSL has already established to be worthful for engineering tasks. PGSL helps to solve complicated problems in a high convergence speed. It has the advantage of fewer search parameters. In PGSL Algorithm, a uniform probability density function is accepted for entire search space in the beginning of search. When good solutions are found, increased the probabilities in these regions. Better sets of points are found in the neighborhood of good sets of points. So, search space is more focus on an area of best points then the convergence is achieved.

Searching the best point in PGSL Algorithm through four nested cycles: Sampling Cycle, Probability Updating Cycle, Focusing Cycle and Subdomain Cycle. Sampling Cycle: Number of samples is randomly generated according to the current probability density function. Each point is evaluated by user-defined objective function and selected the best point. Probability Updating Cycle: The sampling cycle is brought up number of probability updating cycle times. After each iteration, the probability density function of each variable is modified using the probability-updating algorithm. This ensures that the sampling frequencies in regions containing good solutions are increased and regions containing bad points are decreased. Focusing Cycle: Search is focused on the interval containing the best solutions after a number of Probability Updating Cycle, by further subdivision of interval. Subdomain Cycle: Search space is increasingly narrowed by selecting a subdomain of smaller size center on the best point after each Focusing Cycle. In proposed system, best points of the selected frames are seeked based on PGSL Algorithm to embed the secret message with imperceptibility.

In figure 1, four nested loops of PGSL algorithm is displayed.

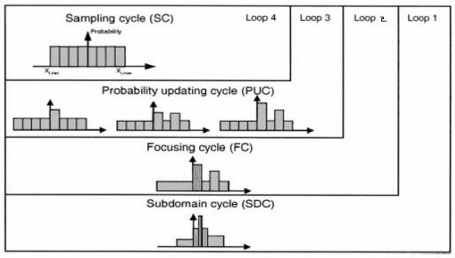

Figure 1. Four nested loops of PGSL Algorithm

5. Proposed System based PGSL

In this section Pre-Processing, Data Embedding, Extraction and System Robustness with noise attack are described.

5.1. Pre-Processing

Clean host speech from the Texas Instruments Massachusetts Institute of Technology (TIMIT) database with 16 bit resolution and 16000 sample rate are used as the cover host. Framing with non-overlapping and windowing are applied on the cover speech. Secret message as text is converted into binary sequences {1, 0} as covert message to embed into the original cover audio.

5.2. Embedding process

The embedding technique is presented in the following.

Input: - speech utterance from TIMIT dataset for cover host
- Secret text as covert message

Step 1. Cover speech samples are transformed to MDCT domain.
Step 2. Searching the best point in the selected frames of transformed speech by defining objective function and parameter values in PGSL.
Step 3. For further security, cover coefficients are modified with gain factor.
Step 4. Embed secret data into best coefficients of frames which are optimal positions through PGSL.
Step 5. Reconstruct the stego audio by applying inverse modified discrete cosine transform (IMDCT) to form stego audio.

Output: Stego audio

The flow of data embedding process is demonstrated in figure 2.

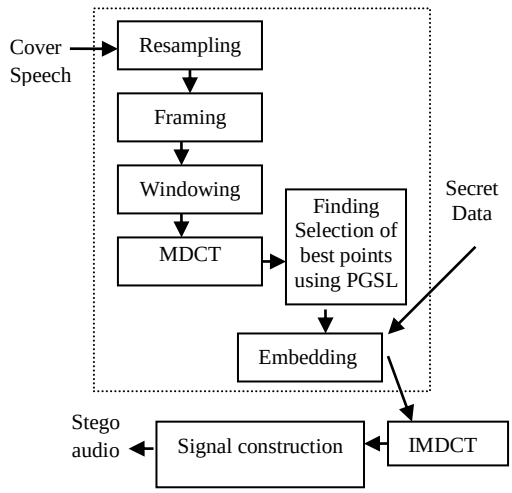

Figure 2. Embedding Process

5.3. Extraction process

Proposed extraction technique can be seen as the following steps.

Input: Stego audio

Step 1. Divide frame from the stego audio as embedding process.
Step 2. Select embedded frames length and then transform it with MDCT.
Step 3. According to the embedding depth, estimate the secret data as binary 0 or 1 with defined threshold value.
Step 4. Convert extracted binary data into character sequence as secret text message.

Output: Detected character sequence as secret text message

The flow of data extraction process is described in figure 3.

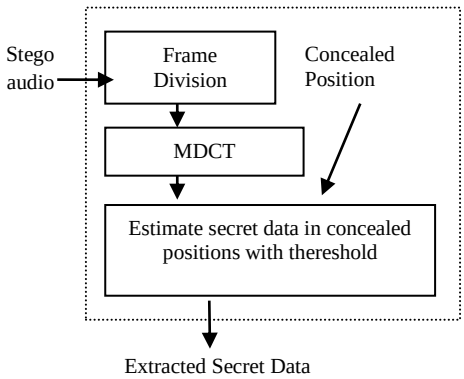

Figure 3. Extraction Process

6. Evaluation Experiments & Results

Experiments were performed to test the performance of PGSL based proposed system. This section first describes the performance criteria by the results obtained by using Spectrogram and Bit Error Rate (BER).

6.1 Spectrogram

Spectrograms are color-based visualizations of the evolution of the power spectrum of a speech signal through time. It is widely used by speech and audio engineers to view the frequency content of speech signal. The general framework of the proposed imperceptibility test using Spectrogram is showed in figure 4.

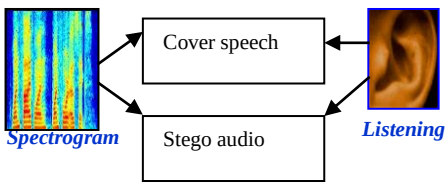

Figure 4. Test Imperceptibility and Inaudibility between original cover host and stego audio

6.2 Bit Error Rate (BER)

Robustness is the main benchmarks used for performance evaluation of data hiding scheme. Robustness of the proposed system is measured with ***BER*** (bit error rate).Compute **BER** as the number of bit errors divided by the total number of bits in the embedded signal to determine how many of received bits are in error from attacked stego audio. The flow of the proposed BER measure is described in figure 5.

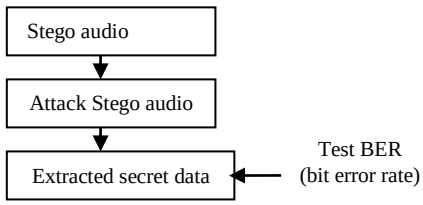

Figure 5. Evaluate Robustness of the system

6.3 Evaluation

The unavailability of data hiding scheme to address the three issue of audio Steganography is mainly important in the research community. As it is uneasy to address all issue at once. The results shown in this system were obtained using Spectrogram and Bit Error Rate to verify high imperceptible and robustness of the system even with high payload. Thus, the performance of the proposed system has been analyzed by two experiments.

6.3.1 Experiment I

The system was evaluated using Clean-host utterances from the TIMIT dataset spoken by both female and male persons with 16000 sample rate and 16 bit resolution as the cover speech to embed one bit in each of the first selected frames. Covert data are defined as secret text. The quality degradation of secret data embedding is evaluated by: (1) Compare waveform of cover and stego audio; (2) objective distortion measure as transparency test using Spectrogram; (3) subjective Informal listening test.

1*) Transparency*

Subjective Informal listening tests have shown that most ordinary audience can not distinguish the difference between the original cover audio and the stego audio signal. To further objectively demonstrate that the proposed system fulfilled the transparency requirement according to the signal waveform and Spectrogram result between the stego and cover host audio at different embed bit rate. Results of comparing waveforms between original and stego audio are described in Fig. 6. Objective distortion measures are showed in Fig. 7 and 8 by using Spectrogram. In these figures shows the distortions between cover (before data embedding) and stego audio (after embedded secret data). At all of the test results; the cover host and stego audio are similar and cannot be discriminated even higher bits rate. Thus, high imperceptible is achieved from these testing.

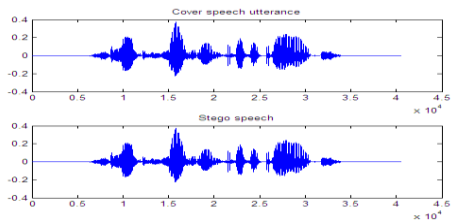

Figure 6. Waveforms of the original clean speech (above) and Stego (bottom)

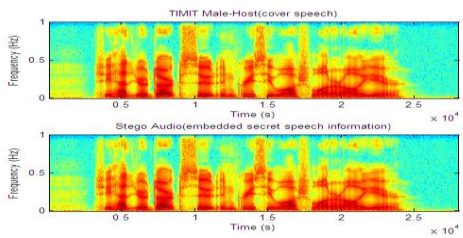

Figure 7. TIMIT utterance host cover audio (above) and Stego speech (bottom)

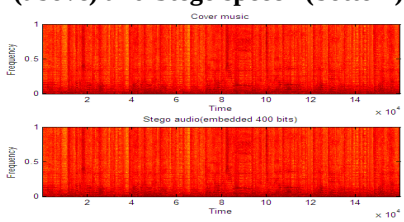

Figure 8. Spectrogram of Cover Music (Original) and Stego Audio

6.3.2 Experiment II

The following observation has been tested for ***robustness test:***

To test robustness, use seven cover speech from TIMIT Dataset and secret text as binary sequence length of 100 bits and 205 bits are used tested data. Results of robustness against noise attack are showed in Table 1. In table 2, three cover music and secret text length as binary sequence of 100, 300 and 400 bits are used to test robustness.

Add white Gaussian noise with snr 10 dB to 80 dB of audio power is added into stego audio, the BER of the recovered data is 0(zero).

The accuracy 100 % and robustness of the system without attacks and with additive white Gaussian noise attack is showed in table 1 and table 2. High accuracy of data recovery rate is satisfied.

Table 1. Robustness and Accuracy on Noise attack using seven host utterances

Cover Speech Length (sec)	Frame size	Sample rate	Secret Message Length (bits)	noise (attack)	Bit error rate (BER)	Correct Data Recovery rate (%)
2.9273	256	16 KHz	100	Snr 30 dB	0	100
2.8045	256	16 KHz	100	Snr 40 dB	0	100
2.3796	256	16 KHz	100	Snr 50 dB	0	100
3.0594	256	16 KHz	100	Snr 60 dB	0	100
2.8989	256	16 KHz	100	Snr 70 dB	0	100
2.8328	256	16 KHz	100	Snr 80 dB	0	100
7.5	480	16 KHz	205	Snr 25 dB	0	100

Table 2.Evaluate Robustness with noise attack (snr 20, 60, 80 db) over three songs

Music (data-size)	320000	180000	1209572
Sample-rate	16 kHz	16 kHz	16 kHz
Frame_ size	800	600	256
Num_ frame	400	300	4.72E+03
Threshold	0	0	0
Secret-bit	400	300	100
Run-time	39.109019 sec	20.003412 sec	11.183278 sec
Attack power(dB)	20	60	80
BER (bit error rate)	0	0	0

Table 3: Comparison of the proposed system with previous method on noise attack

Comparison	**Embed Secret bits**	**Attack Noise addition**	**BER**
Audio Steganography by Cepstrum Modification (2005-IEEE,ICASSP)	**205 bits**	**25 dB**	**5 ~ 8**
Proposed system	**205 bits**	**25 dB**	**0**

7. Conclusion

A new secret data communication using information hiding techniques by imperceptible ways over cover audio has already demonstrated. The system was tested with different speech from TIMIT dataset and variety of music styles from music CD. The advantages of proposed system are clear and significant. The effectiveness of this system is verified through experiment test. By using PGSL, the proposed system can fulfill the good inaudibility of stego audio from cover speech with less processing time. Achieve robustness against distortion constrained attacks as additive noise. The measurement of perceived audio quality demonstrated to address three issue of data hiding. The last but not the least, the system can be valuable for multimedia authentication for music copyright protection, verify integrity and secret data communication.

References

[1] J.R. Krenn, “Steganography and Steganalysis”, January-2004.

[2] Niels Provos, Peter Honeyman, Hide and Seek: Introduction to Steganography (2003).

[3] Z. Brahimi, H. Bessalah, A. Tarabet, and M. K. Kholladi, Selective Encryption Techniques of JPEG2000 Codestream for Medical Images Transmission, *WSEAS Transactions on Circuits and Systems*, Vol. 7, No. 7, 2008, pp. 718-727.

[4] C. Chemak, J.-C. Lapayre, and M.-S. Bouhlel, A New Scheme of Image Watermarking Based on 5/3 Wavelet Decomposition and Turbo-Code, *WSEAS Transactions on Biology and Biomedicine*, Vol. 4, No. 4, 2007, pp. 45-52.

[5] W.-P. Fang, Combining Copyright Protection and Data Hiding – A Sensitive Transformation Approach, *Proceedings of the 6th WSEAS International Conference on Signal Processing, Computational Geometry & Artificial Vision*, Greece, August 21-23, 2006, pp. 45-49.

[6] M.-S. Wang and W.-C. Chen, DCT-domain Copyright Protection Scheme Based on Secret Sharing Technique, *Proceedings of the 7th WSEAS International Conference on Signal Processing, Computational Geometry & Artificial Vision*, Greece, August 24-26, 2007, pp. 107-111.

[7] Y.I. Khamlichi, M. Machkour, K. Afdel, and A. Moudden, Medical Image Watermarked by Simultaneous Moment Invariants and Content-Based for Privacy and Tamper Detection, *6th WSEAS International Conference on Multimedia Systems & Signal Processing*, China, April 16-18, 2006, pp. 109-113.

[8] A. Samcovic, and J. Turan, Digital Image Watermarking by Spread Spectrum, *Proceedings of the 11th WSEAS International Conference on*

Communications, Greece, July 26-28, 2007, pp.29-32.

[9] X. Wang, Y Ou-Yang, and H.-M. Gu, A Remote Sensing Image Self-Adaptive Blind Watermarking Algorithm Based on Wavelet Transformation, *Proceedings of the 7th WSEAS International Conference on Signal, Speech and Image Processing*, China, September 15-17, 2007, pp. 76-82.

[10] Gary C. Kessler, An Overview of Steganography for the Computer Forensics Examiner, *Forensics Science Communications*, Vol. 6, No. 3, July 2004.

[11] C. Hosmer, and C. Hyde, Discovering Covert Digital Evidence, *Digital Forensic Research Workshop (DFRWS)*, August 2003.

[12] N. Provos, P. Honeyman. "Hide and seek: an introduction to steganography", *IEEE Security & Privacy Magazine*, vol. 1, Issue 3, pp. 32-44, May-June 2003.

[13] K. Bailey, K. Curran. "An evaluation of image based steganography methods", *Multimedia Tools and Applications*, vol. 30, Issue 1, pp. 55-88, July 2006.

[14] A.W.Naji, A.A.Zaidan, B.B.Zaidan, Shihab A, Othman O. Khalifa, "Novel Approach of Hidden Data in the (Unused Area 2 within EXE File) Using Computation Between Cryptography and Steganography ",International Journal of Computer Science and Network Security (IJCSNS) , Vol.9, No.5 , ISSN : 1738-7906, pp.

[15] A.A.Zaidan, Fazidah. Othman, B.B.Zaidan, R.Z.Raji, Ahmed.K.Hasan, and A.W.Naji," Securing Cover-File without Limitation of Hidden Data Size Using Computation between Cryptography and Steganography ", World Congress on Engineering 2009 (WCE), The 2009 International Conference of Computer Science and Engineering, Proceedings of the International Multi Conference of Engineers and Computer Scientists 2009, ISBN: 978-988-17012-5-1, Vol.I, p.p259-265.

[16] A.A.Zaidan, A.W. Naji, Shihab A. Hameed, Fazidah Othman and B.B. Zaidan, " Approved Undetectable-Antivirus Steganography for Multimedia Information in PE-File ",International Conference on IACSIT Spring Conference (IACSIT-SC09) , Advanced Management Science (AMS), Listed in IEEE Xplore and be indexed by both EI (Compendex) and ISI Thomson (ISTP), Session 9, P.P 425-429.4-300.

[17] Pal S.K., Saxena P. K. and Mutto S.K. "The Future of Audio Steganography". Pacific Rim Workshop on Digital Steganography, Japan, 2002.

[18] Princen, J. P., Bradley, A. B.,"Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, No. 5, October 1986.

[19] Princen, J. P., Johnson, A. W., Bradley, A. B., "Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation," IEEE International Conference on Acoustics, Speech, and Signal Processing, 1987,Dallas, USA, pp. 2161-2164.

[20] M. Zamani, A. A. Manaf, R. B. Ahmad, A. M.Zeki, and S. Abdullah[5] "A Genetic-Algorithm-Based Approach for Audio Steganography ", *World Academy of Science, Engineering and Technology* 54 2009.

[21] K. Gopalan "Audio Steganography by Cepstrum Modification "*ICASSP, IEEE*, 2005.

[22] N. Cvejic & T. Seppanen "Increasing Robustness of LSB Audio Steganography Using a Novel Embedding Method ", *IEEE* 2004.

[23] S. Chen ; H. Leung, H. Ding, "Telephony Speech Enhancement by data hiding " *Instrumentation and Measurement, IEEE Transactions* on, Feb 2007.

[24] B. Raphael, I.F.C. Smith, "A direct stochastic algorithm for global search "Applied *Mathematics and Computation* 146 (2003) 729–758.

[25] W. Bender, W. Butera, D. Gruhl, R. Hwang, F. J. Paiz, S. Pogreb, "Techniques for data hiding", IBM Systems Journal, Volume 39 , Issue 3-4, July 2000, pp. 547 568.

[26] Samir Kumar Bandyopadhyay, Debnath Bhattacharyya, Poulami Das, Debashis Ganguly and Swarnendu Mukherjee, "A tutorial review on Steganography", International Conference on Contemporary Computing (IC3-2008), Noida, India, August 7-9, 2008.

[27] W. Bender, D.Gruhl, and N. Morimoto, "Techniques for data hiding" *IBM Systems journal*, 1996, 35(3/4): 131 336.

[28]M.Cooperman and S.Moskowitz, "Steganographic method and device",*U.S. Patent* 5, 913,004, Mar.,1997.

[29] J. Huang, Y. Wang, and Y.Q. Shi, "A blind audio watermarking algorithm with self-synchronization", *Proc. Of IEEE, Int. Sym. On Circuits and Systems*, vol. 3 pp. 627-630, 2002.

[30] D. Gruhl, A. Lu and W. Bender, "Echo Hiding", *Proc.Of 1st Information Hiding Workshop*, LNCS vol. 1174, Berlin, Germany: Springer – Verlag, pp. 295 -315, 1996.

[31] P. Bassia, I. Pitas, and N. Nikoladis, "Robust audio watermarking in time domain", *IEEE Transactions on Multimedia*, vo1. 3, pp 232-241, 2001.

[32] Kim, H.J., and Choi Y.H., "A novel echo-hiding scheme with backward and forward kernels",

IEEE Transactions on circuits and Systems for video and technology, vo1. 13, 2003.