Audio Steganography in Wavelet Domain – A Survey

Jisna Antony
M. Tech student
Department of Computer Science and Engineering
MES College of Engineering
Kuttippuram, India

Sobin c. c
Asst. Professor
Department of Information Technologies
MES College of Engineering
Kuttippuram, India

Sherly A. P
Asst. Professor
Department of Computer Science and Engineering
MES College of Engineering
Kuttippuram, India

ABSTRACT
Steganography is the art and science of writing hidden messages such that the existence of a secret communication is known only to the sender and receiver. For hiding messages different types of media are used. Audio steganography uses audio as the cover media. Commonly used techniques for audio steganography are temporal domain and transform domain techniques, where the frequency domain techniques and wavelet domain techniques come under transform domain. Under temporal domain the techniques include LSB encoding, parity coding and echo hiding. Under frequency domain the different techniques are tone insertion, phase coding and spread spectrum technique. This paper makes a discussion on audio steganography techniques. Among the techniques studied wavelet domain shows high hiding capacity and transparency. In wavelet domain different techniques are applied on the wavelet coefficients to increase the hiding capacity and perceptual transparency. The paper mainly concentrates on a survey on audio steganography in wavelet domain.

General Terms
Audio Steganography, Wavelet

Keywords
Audio steganography, Wavelet, Lifting scheme, Speech steganography, Int2Int wavelets, Efficient wavelet masking.

1. INTRODUCTION
Steganography is the art and science of writing hidden messages so that no one apart from the sender and receiver even realizes that there is a hidden message. Steganography [1], [2] hides the data in an unremarkable media so that it doesn't arouse any suspicion. Steganography, cryptography and watermarking does the same purpose i.e., protecting the information. In steganography the existence of the hidden message is not known, but in cryptography the secret message is converted into a different form.

A cryptographic message is a meaningless stream of words or bits, which are rare in a computer world and hence it arises suspicion. Cryptography protects the contents of a message whereas steganography protects the messages and the communication parties. Digital watermarking hides information in a carrier signal; the hidden information should, but does not need to contain a relation to the carrier signal. Both steganography and digital watermarking uses steganographic techniques to embed data covertly in noisy signals. But whereas steganography aims for imperceptibility to human senses, digital watermarking tries to control the robustness as top priority. Today, computer and network technologies provide easy-to-use communication channels for steganography.

For hiding information, a cover media is required. Secret message is hidden in the redundant portion of the cover media. After embedding the secret message, we get a stegano media. The block diagram is shown in figure 1. This stegano media must have the same characteristics as the cover media, i.e., it must be imperceptible from cover media. Three requirements for any steganographic system are perceptual transparency, hiding capacity and robustness [1]. Perceptual transparency means that the stegano medium must be imperceptible from the cover medium i.e., the cover object and the stegano object must be perceptually indiscernible. The transparency of cover object and stegano object is often computed using Peak Signal to Noise Ratio (PSNR) values. Hiding capacity refers to the amount of information that can be hidden in a cover media. It is measured in bits per pixel for images and bits per second for audio. When capacity is increased beyond a limit, transparency decreases. Robustness refers to its ability to withstand with attacks. A good steganographic system must satisfy all these requirements.

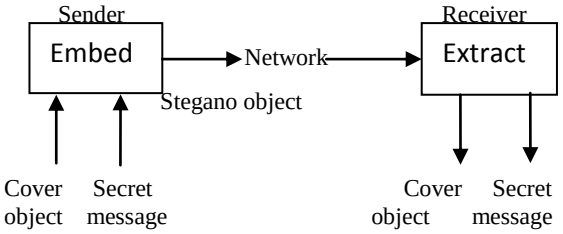

Figure 1: Block diagram for audio steganography

Depending upon the cover media, there are different types of steganography [3], [4], like text, image, audio, video and protocol steganography. Among the different types of steganography audio steganography is more important because audio and video can carry more redundant information compared to other media. Increase of audio traffic on the internet shows that audio can be transferred at lower rates. Also audio steganography can also be applied to video steganography since video is a combination of audio and image. But embedding data in audio is considered to be more difficult compared to other media because of high precision of human auditory system.

Steganalysis [1], [2] is the detection of steganographic encryption. Steganalysis identifies steganography by inspecting various parameters of stegano media. Unlike cryptanalysis, where the existence of a secret message is sure from the structure of the message itself, in the case of steganalysis this may not be true. Various techniques are applied for steganalysis. The properties of electronic media are changed after hiding any object to it. This results in degradation in terms of quality and usual characteristics of the media. In this way steganography can be detected. For image steganography the commonly used technique is visual detection. An intelligent system trained with statistics of host and stegano signals is used in most of the steganalysis test. Steganalysis is followed by steganography attacks. Steganography attacks [1] involve detecting, extracting, and destroying a hidden object of stegano media.

The remainder of the paper is organized as follows: Section 2 discusses commonly used audio steganography techniques. Section 3 makes a survey on audio steganography in wavelet domain. Section 4 concludes the survey presented here, following which references are given in section 5.

2. AUDIO STEGANOGRAPHY TECHNIQUES

2.1 Temporal Domain

2.1.1 LSB: LSB [5], [6] is one of the earliest and simplest methods for hiding information in audio signals. It is the commonly used technique for audio steganography. In LSB encoding, the least significant bits of the cover media/original audio is altered to include the secret message. Even though this is a simple method, an attacker can easily extract the secret message from the stegano object.

2.1.2 Parity coding: Parity coding technique [3], [4] operates on a group of samples instead of individual samples. Here individual samples are grouped and parity of each group is calculated. For inserting message bit one by one, check the parity bit of a group of samples. If the parity bit and message bit matches do nothing. Otherwise change the LSB's of any one of the individual samples in that group to make the parity bit equal to the message bit.

2.1.3 Echo hiding: In echo hiding [7] method data is embedded in the echo part of the host audio signal. The echo is a resonance added to the host signal and hence the problem with the additive noise is avoided here. While using echo hiding three parameters are to be considered: they are initial amplitude, offset (delay), and decay rate, so that echo is not audible. The main disadvantage of this method is lenient detection and low detection ratio. Due to its low embedding rate and low security no researches are going on echo hiding technique.

2.2 Frequency Domain

Frequency domain techniques and wavelet domain technique comes under transform domain. The main techniques under frequency domain are: tone insertion, phase coding and spread spectrum technique.

2.2.1 Tone insertion: Frequency masking property is exploited in tone insertion method [8]. A weak pure tone is masked in the presence of a stronger tone. This property of inaudibility is used in different ways to embed information.

2.2.2 Phase coding: Phase coding method [9] is based on the fact that the phase components are not audible to human as noise components. This method embeds the secret message bits as phase shift in the phase spectrum of the original audio signal. The method tolerates better signal distortion, better robustness but it does not survive low pass filtering. Here the secret message is inserted only at the phase vector of the first signal segment.

2.2.3 Spread spectrum technique [10]: This technique takes the advantage of masking property of HAS. A masking threshold is calculated using a psycho-acoustic model. The spread signal now lies below the masking threshold. Apart from phase shifting, here the secret message is distributed along with the host signal. Here the final signal occupies a bandwidth which is more than what is actually required for transmission.

2.3 Wavelet Domain

Wavelet domain [11] is suitable for frequency analysis because of its multi-resolution properties that provides access to both most significant parts and details of spectrum. Wavelet domain techniques works with wavelet coefficients. Upon applying the inverse transform, the stegano signal can be reconstructed.

Table 1. Summary of audio steganography techniques

Method	Strength	Weakness
LSB	Simple	Easy to extract
Parity coding	More robust than LSB	Easy to extract
Echo hiding	Avoids problem with additive noise	Low capacity
Tone insertion	Exploits masking property	Low embedding capacity
Phase coding	Robust	Low capacity
Spread spectrum	Increases transparency	Occupies more bandwidth
Wavelet domain	High hiding capacity and transparency	Lossy data retrieval

Signal-to-Noise Ratio (SNR) is used to evaluate the performance of the audio steganography techniques. SNR value indicates the distortion amount induced by embedding data in the cover audio signal. SNR [7] value is given by the following equation:

$$SNR_{dB}=10\log_{10}\left(\frac{\sum_{n=1}^{N}|s_c(m,n)|^2}{\sum_{n=1}^{N}|s_c(m,n)-s_s(m,n)|^2}\right) \quad (1)$$

Where $s_c(m,n)$ is the cover audio signal and $s_s(m,n)$ is the stegano audio signal such as: m = 1, …M and n = 1, …N, where M is the number of frames in milliseconds and N is the number of samples in each frame.

Table I shows a summary of audio steganography techniques. Among the different techniques studied, wavelet domain shows high hiding capacity compared to other domain. Also subjective experiments in wavelet domain showed that this hiding scheme is acoustically more transparent compared to other domain. Hence we are concentrating on wavelet domain.

3. LITERATURE SURVEY

Nedeljko Cvejic, et al. [11], proposes a paper based on wavelet domain. In this paper the audio signal is converted into wavelet coefficients. Figure 2 shows decomposition of signal into low and high frequency components. The wavelet coefficients are then scaled using the maximum value inside the given subband, before they are converted into binary string. The data bits are hidden in the least significant bit portion of wavelet coefficients. At the receiver side the secret bits are extracted from the LSB portion of the wavelet coefficients. This is possible because of perfect reconstruction properties of filter banks. The discrete wavelet transform decomposes the signal into low pass and high pass components. This paper uses Haar filter, because it is the only quadrature mirror filter that has finite impulse response. On applying the inverse transform the original signal can be reconstructed. This paper proposes a general method of using wavelet domain for audio steganography. The method shows large advantage in terms of hiding capacity and transparency compared to other domain.

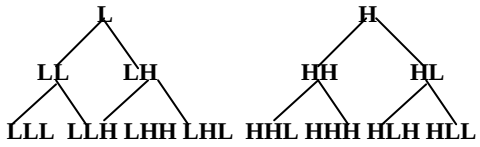

Figure 2: Signal decomposition

Sajad Shirali Shahreza, et al. [12], proposes a paper in wavelet domain based on lifting scheme. In this method, lifting scheme is used to create perfect reconstruction Int2Int wavelets. The paper first converts the signal into wavelet coefficients. The problem with the normal wavelets is that when applying them on an integer signal such as a speech, the resulted coefficients may not be integer. To avoid this problem the paper uses a lifting scheme to produce Int2Int wavelets. Int2Int means that if the input signal is integer, the wavelet coefficients are also integer. So there is no need to scale the coefficients and convert them to binary representation. Thus Int2Int wavelets avoid the problem with non-integer values.

Here the data bits are hidden according to the capacity of wavelet coefficients. Bigger coefficients hide more data bits than smaller coefficients. For hiding the data bits in wavelet coefficients, find the biggest power of 2 (say p) which is smaller than the wavelet coefficient. This means that p bits can be hidden in that wavelet coefficient. For transparency, data bits are hidden in detailed coefficients. The data is hidden in the LSB's of wavelet coefficients. The paper doesn't hide data in the silent parts. Hence the problem with the silent parts is avoided. The method shows high hiding capacity and zero error in hiding/unhiding process. The main advantage of this method is its high hiding capacity and high perceptual transparency. The method also shows zero error rate compared to other methods.

Parul Shah, et al. [13], proposes an adaptive wavelet packet based audio steganography using data history. In this paper, data is embedded by adaptively modifying wavelet packet coefficients of host audio signal. The paper adopts a scheme which consists of five main parts: Wavelet Packet Decomposition (WPD), binary mapping using trend detection, pattern classification, embedding covert data by pattern modification and inverse wavelet packet. WPD decomposes the audio signal into wavelet coefficients. Binary mapping is done by using four steps. First sampling is done on the wavelet coefficients in selected subband to evenly distribute the secret message. Then, as a first level of encryption these coefficients are converted from 1-D to 2-D. Blocks for embedding the secret message are selected by using a pseudo random sequence. Then trend detection is used for mapping these blocks to binary sequence. After pattern classification a pattern matrix and corresponding pattern value is obtained. For embedding the data bits (pattern value), the blocks are modified according to the pattern matrix to get the pattern value. The stegano audio file is obtained after applying inverse transform. The paper shows superior SNR values, with good hiding capacity and speed. Also this method exhibits zero bit error in recovered data which is one of the most desired features of any steganographic system.

S Nehete, et al. [14], made a detailed comparison between DWT and Discrete Cosine Transform (DCT). The paper proposes a method for digital audio steganography with security. Security is provided using cryptography. The embedding algorithm proposed in the paper is explained below. First the signal is decomposed up to its third level. Then quantization is performed on decomposed signal. The secret data are then hidden in audio coefficients and stegano audio signal is constructed. The characteristics of this method are imperceptibility, robustness, large payload. The paper is implemented in integer wavelet domain using quantization to reduce embedding error.

Haider Ismael Shahadi and Razali Jidin [15], proposes a high capacity and inaudibility audio steganography scheme. The algorithm is based on discrete wavelet packet transform with adaptive hiding in least significant bits. Here the input signal is segmented into G segments. The secret message is also segmented into G segments. Then the following steps are repeated G times to hide each secret message segment in one cover segment. The cover signal is decomposed into wavelet coefficients and each detail signals is scaled according to its maximum value and number of bits per sample. For each sample, the algorithm determines the number of bits that can be safely hidden. This forms an embedded position contents vector. The second step uses a bits block matching process to find the most matching embedded position content vector for a secret message segment. Then secret message is embedded and a stegano-key vector is constructed. In the next step the stegano-key is embedded in lowest frequency details signal which makes the stegano-key more resistant against

distortion. Then stegano signal is reconstructed. The algorithm has got high hiding capacity and excellent output quality.

Dora M. Ballesteros L and Juan M Moreno A [16], proposes a paper in wavelet domain based on Efficient Wavelet Masking (EWM). The paper mainly concentrates on speech in speech hiding. EWM is a steganography model which adapts the secret message to the host signal. It uses two principles: the efficient adaptation and masking property of Human Auditory System (HAS). The efficient adaptation model says that, 'any speech secret message may seem similar to a speech host message if its wavelet coefficients are sorted'.
The two speech signals, host speech and secret message, are decomposed by discrete wavelet transform to its wavelet coefficients and adequate scaling is applied in each case. The coefficients with low amplitudes are set to zero, since they does not provide significant energy to the signal. Then the secret message is sorted according to the host speech signal. The paper then uses indirect LSB substitution for hiding speech in speech. The main advantage of using this indirect method is that the number of bits replaced is less than in direct form. It is shown that the proposed model has high hiding capacity than spread and shift spectrum method and its statistical transparency are also higher than others. Classical frequency masking schemes are not suited for speech in speech hiding. The method in the paper is best suited for speech in speech hiding. Also it uses all of the host coefficients to hide the secret message, instead of a selected group of coefficients in other methods. The method also uses a secret key which adds additional security.

The method in [12] hides the data bits according to the capacity of wavelet coefficients. The method [15] uses a bits block matching process to get the most matching host coefficient to embed the secret message. The method in [16] also uses a similar method, called efficient sorting to get the most matching coefficient. Almost all the papers discussed above hides the data bits without affecting the transparency of the stegano audio. But the main advantage of the method [16] is that, it uses an indirect LSB replacement which reduces the size of secret message and thereby increases the hiding capacity compared to the previous methods. In [16] it is shown that by indirect LSB replacement, a fourteen bit data is reduced to five bit data.

4. CONCLUSION

Steganography is the process of hiding information in host signal. Audio steganography uses audio as the cover media. Commonly used audio steganography techniques and their advantages and disadvantages are discussed in this paper. Among the different techniques reviewed, wavelet domain shows high hiding capacity and high perceptual transparency. Hence we are concentrated on wavelet domain. Then a survey is made on audio steganography in wavelet domain. Based on this survey the efficient wavelet masking technique proposed in [16] satisfies almost all the requirements for a good steganographic system. The method in [16] also shows better performance in steganalysis test, compared to other methods. As a future work lifting scheme can also be used with efficient wavelet masking scheme. The main problem with the wavelet domain is obtaining high quality data at the receiver side. A small change in the host coefficients value will affect the retrieving procedure. Hence an embedding method used in wavelet domain must also focus on quality data at the receiver side.

5. REFERENCES

[1] S. Das, B. Bandyopadhyay and S. Sanyal, " Steganography and steganalysis: different sapproaches", Cornell University Library, 2011.

[2] S. K. Bandyopadhyay, B. Bhattacharyya, D. Ganguly, S. Mukherjee, P. Das, "A tutorial review on steganography", International Conference on Contemporary Computing, 2008.

[3] P. Jayaram, H. Ranganatha, and H. Anupama, "Information hiding using audio steganography-a survey", International Journal of Multimedia and its Applications, 2011.

[4] H. Kekre, A. Athawale, S. Rao, and U. Athawale, "Information hiding in audio signals", International Journal of Computer Applications, IJCA, vol. 7,no. 9, pp. 14-19, 2010.

[5] M. Asad, J. Gilani, and A. Khalid, "An enhanced least significant bit modification technique for audio steganography", 2011 International Conference on Computer Networks and Information Technology (ICCNIT), IEEE, 2011.

[6] K. Bhowal, A. Pal, G. Tomar, and P. Sarkar, "Audio steganography using GA", 2010 International Conference on Computational Intelligence and Communication Networks (CICN), IEEE, 2010.

[7] F. Djebbar, B. Ayad, H. Hassmam, and K. Abed-Meraim, "A view on latest audio steganography techniques", 2011 International Conference on Innovations in Information Technology (IIT), IEEE, 2011.

[8] K. Gopalan and S. Wenndt, "Audio steganography for covert data transmission by imperceptible tone insertion", Proceedings of Communications Systems and Applications, IEEE, 2004.

[9] M. Nutzinger and J. Wurzer, "A novel phase coding technique for steganography in auditive media", 2011 Sixth International Conference on Availability, Reliability and Security (ARES), IEEE, 2011.

[10] S. Md, B. Vijaya, and V. Shiva Nagaraju, "An optimized method for concealing data using audio steganography", International Journal of Computer Applications, 2011.

[11] N. Cvejic and T. Seppanen,"A wavelet domain lsb insertion algorithm for high capacity audio steganography", Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop, IEEE, 2002.

[12] S. Shahreza and M. Shalmani, "High capacity error free wavelet domain speech steganography", IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2008.

[13] P. Shah, P. Choudhari, and S. Sivaraman, "Adaptive wavelet packet based audio steganography using data history", IEEE Region 10 and the Third international Conference on Industrial and Information Systems, ICIIS, IEEE, 2008.

[14] S. Nehete, S. Sawarkar, and M. Sohani, "Digital audio steganography using dwt with reduced embedding error and better extraction compared to dct", Proceedings of the International Conference & Workshop on Emerging Trends in Technology, ACM, 2011.

[15] Haider Ismael Shahadi and Razali Jidin, "High capacity and inaudibility audio steganography scheme", 2011 7th International Conference on Information Assurance and Security (IAS), IEEE, 2011.

[16] D. Ballesteros L and J. Moreno A, "Highly transparent steganography model of speech signals using efficient wavelet masking", Expert Systems with Applications, Elsevier, 2012.