Hiding Text in Audio Using LSB Based Steganography

K.P.Adhiya Swati A. Patil
CSE Dept. SSBT's COET Bambhori,Jalgaon,Bambhori,India
Email: swati.patil251@gmail.com

Abstract
A Steganographic method for embedding textual information in WAV audio is discussed here. In the proposed method each audio sample is converted into bits and then the textual information is embedded in it. In embedding process , first the message character is converted into its equivalent binary. The last 4 bits of this binary is taken into consideration and applying redundancy of the binary code the prefix either 0 or 1 is used. To identify the uppercase, lower case, space ,and number the control symbols in the form of binary is used. By using proposed LSB based algorithm, the capacity of stego system to hide the text increases. The performance evaluation is done on the basis of MOS by taking 20 samples and comparison of SNR values with some known and proposed algorithm.

Keywords: LSB, WAV, MOS , control symbols, stego system , SNR.

1. Introduction

As the need of security increases only encryption is not sufficient. So stegnograpghy is the supplementary to encryption. It is not the replacement of encryption. But Steganography along with encryption gives more security to data. The word steganography is of Greek origin and means "concealed writing" from the Greek words stegnos meaning "covered or protected", and graphei meaning "writing". Steganography is the technique to hide the information in some media so that third party can't recognize that information is hidden into the cover media.. That media may be text, image ,audio or video. The information that to be hidden is called stego and the media in which the information is hidden is called host. The stego object can be text, image, audio or video. When the information is hidden into the audio then it is called Audio steganography. The process of Steganography is as shown in Figure1. The random selection of the samples used for embedding introduces low power additive white Gaussian noise (AWGN). It is well known from psychoacoustics literature [3] that the human auditory system (HAS) is highly sensitive to the AWGN.

Hiding information into a media requires following elements [6]

- The cover media(C) that will hold the hidden data
- The secret message (M), may be plain text, cipher text or any type of data
- The stego function (Fe) and its inverse (Fe^{-1})
- An optional stego-key (K) or password may be used to hide and unhide the message.

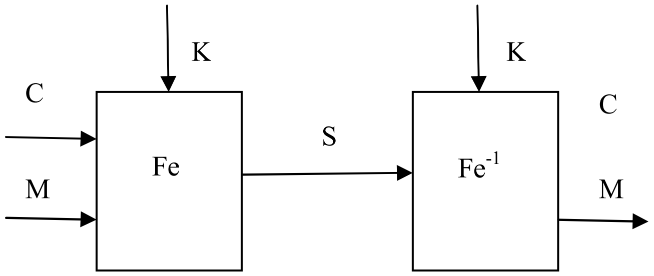

Figure 1: The Steganographic operation

Because the size of the information is generally quite small compared to the size of the data in which it must be hidden (the cover text), electronic media is much easier to manipulate in order to hide data and extract messages. Secondly, extraction itself can be automated when the data is electronic, since computers can efficiently manipulate the data and execute the algorithms necessary to retrieve the messages. Because degradation in the perceptual quality of the cover object may leads to a noticeable change in the cover object which may leads to the failure of objective of steganography.

If in one application we want to achieve confidentiality, than we have two alternatives: encryption or steganographic techniques for protection against detection (see Figure 2).

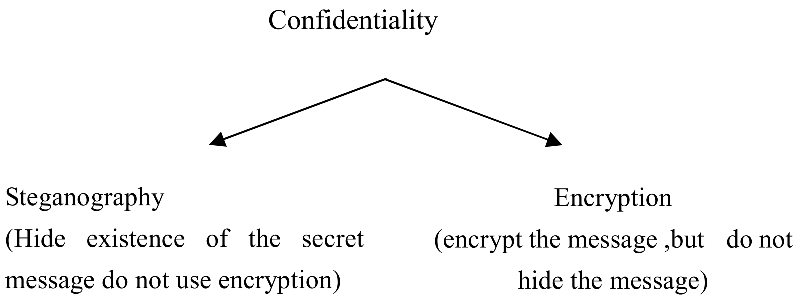
Confidentiality
Steganography
(Hide existence of the secret message do not use encryption)
Encryption
(encrypt the message ,but do not hide the message)

Figure 2 Achieving confidentiality

An effective steganographic scheme should posses the following desired characteristics [9]:

Secrecy: A person should not be able to extract the covert data from the host medium without the knowledge of the proper secret key used in the extracting procedure.

Imperceptibility: The medium after being embedded with the covert data should be indiscernible from the original medium. One should not become suspicious of the existence of the covert data within the medium.

High capacity: The maximum length of the covert message that can be embedded should be as long as possible.

Resistance: The covert data should be able to survive when the host medium has been manipulated, for example by some lossy compression scheme .

Accurate extraction: The extraction of the covert data from the medium should be accurate and reliable. Basically, the purpose of steganography is to provide secret communicate like cryptography.

2.Audio Steganography

Like the document images, the sound files may be modified in such a way that they contain hidden information, like copyright information; those modifications must be done in such a way that it should be impossible for a pirate to remove it, at least not without destroying the original signal. The methods that embeds data in sound files use the properties of the Human Auditory System (HAS). The HAS perceives the additive random noise and also the perturbations in a sound file can also be detected. But there are some “holes” we can exploit. While the HAS have a large dynamic range, it has a fairly small differential range.

2.1Technique for Data Hiding in Audio

There are four techniques for hiding data in Audio as following:

2.1.1 Least Significant Bit (LSB) Encoding:

Least significant bit (LSB) coding is the simplest way to embed information in a digital audio file. By substituting the least significant bit of each sampling point with a binary message, LSB coding allows for a large amount of data to be encoded. In LSB coding, the ideal data transmission rate is 1 kbps per 1 kHz. In some implementations of LSB coding, however, the two least significant bits of a sample are replaced with two message bits. This increases the amount of data that can be encoded but also increases the amount of resulting noise in the audio file as well. A novel method which increases the limit up to four bits by Nedeljko Cvejic, Tapio Seppben & mediaTeam Oulu at Information Processing Laboratory, University of Oulu, Finland .[5]

To extract a secret message from an LSB encoded sound file, the receiver needs access to the sequence of sample indices used in the embedding process. Normally, the length of the secret message to be encoded is smaller than the total number of samples in a sound file. One must decide then on how to choose the subset of samples that will contain the secret message and communicate that decision to the receiver. One trivial technique is to start at the beginning of the sound file and perform LSB coding until the message has been completely embedded, leaving the remaining samples unchanged. This creates a security problem, however in that the first part of the sound file will have different statistical properties than the second part of the sound file that was not modified. One solution to this problem is to pad the secret message with random bits so that the length of the message is equal to the total number of samples.

There are two main disadvantages associated with the use of methods like LSB coding. The human ear is very sensitive and can often detect even the slightest bit of noise introduced into a sound file, Second disadvantage however, is that this is not robust. If a sound file embedded with a secret message using either LSB coding was resample, the embedded information would be lost. Robustness can be improved somewhat by using a redundancy technique while encoding the secret message. However, redundancy techniques reduce data transmission rate significantly.

2.1.2 Phase Coding

Phase coding addresses the disadvantages of the noise inducing methods of audio steganography. Phase coding relies on the fact that the phase components of sound are not as perceptible to the human ear as noise is. Rather than introducing perturbations, the technique encodes the message bits as phase shifts in the phase spectrum of a digital signal, achieving an inaudible encoding in terms of signal-to perceived noise ratio. Original and encoded signal are as shown in Figure 3.

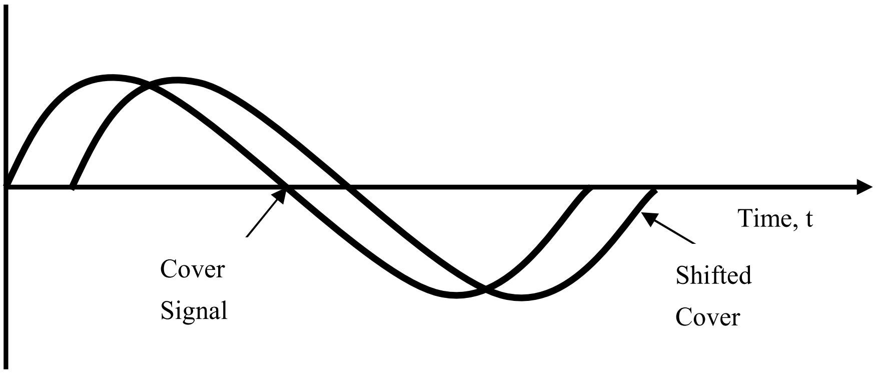

Figure 3. illustrate the original cover signal and encoded shifted signal of phase coding technique.

Phase coding is explained in the following procedure:

- The original sound signal is broken up into smaller segments whose lengths equal the size of the message to be encoded.
- A Discrete Fourier Transform (DFT) is applied to each segment to create a matrix of the phases and Fourier transform magnitudes.
- Phase differences between adjacent segments are calculated.
- Phase shifts between consecutive segments are easily detected. In other words, the absolute phases of the segments can be changed but the relative phase differences between adjacent segments must be preserved. Therefore the secret message is only inserted in the phase vector of the first signal segment as follows:

$$\textbf{Phase\_new} = \begin{cases} \pi / 2 \text{ if message bit } =1 \\ \pi / 2 \text{ if message bit } =0 \end{cases}$$

- A new phase matrix is created using the new phase of the first segment and the original phase differences.
- Using the new phase matrix and original magnitude matrix, the sound signal is reconstructed by applying the inverse DFT and then concatenating the sound segments back together.

To extract the secret message from the sound file, the receiver must know the segment length. The receiver can then use the DFT to get the phases and extract the information. One disadvantage associated with phase coding is a low data transmission rate due to the fact that the secret message is encoded in the first signal segment only. This might be addressed by increasing the length of the signal segment. However, this would change phase relations between each frequency component of the segment more drastically, making the encoding easier to detect. As a result, the phase coding method is used when only a small amount of data, such as a watermark, needs to be concealed.

2.1.3 Echo Hiding

In echo hiding, information is embedded in a sound file by introducing an echo into the discrete signal. Like the spread spectrum method, it too provides advantages in that it allows for a high data transmission rate and provides superior robustness when compared to the noise inducing methods. To hide the data successfully, three parameters of the echo are varied: amplitude, decay rate, and offset (delay time) from the original signal. All three parameters are set below the human hearing threshold so the echo is not easily resolved. In addition, offset is varied to represent the binary message to be encoded. If only one echo was produced from the original signal, only one bit of information could be encoded. Therefore, the original signal is broken down into blocks before the encoding process begins. Once the encoding process is completed, the blocks are concatenated back together to create the final signal. To extract the secret message from the stego-signal, the receiver must be able to break up the signal into the same block sequence used during the encoding process. Then the autocorrelation function of the signal's cepstrum (the cepstrum is the Forward Fourier Transform of the signal's frequency spectrum) can be used to decode the message because it reveals a spike at each echo time offset, allowing the message to be reconstructed. For a discrete signal f(t), an echo f(t-dt), with some delay can be introduced to produce the stego signal s(t)=f(t) + f(t-dt)[2].

2.1.4 Spread Spectrum

In the context of audio steganography, the basic spread spectrum (SS) method attempts to spread secret information across the audio signal's frequency spectrum 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. Two versions of SS can be used in audio steganography: the direct-sequence and frequency-hopping schemes. In direct sequence SS, the secret message is spread out by a constant called the chip rate and then modulated with a pseudorandom signal. It is then interleaved with the cover-signal. In frequency-hopping SS, the audio file's frequency spectrum is altered so that it hops rapidly between frequencies.

The SS method has the potential to perform better in some areas than LSB coding, parity coding, and phase coding techniques in that it offers a moderate data transmission rate while also maintaining a high level of robustness against removal techniques. However, the SS method shares a disadvantage with LSB and parity coding in that it can introduce noise into a sound file.

3.Proposed Technique

The existing algorithm hided the text in image. In proposed technique the algorithm will be implemented for Audio Signal to hide text.

The algorithm based on the redundancy of bits in binary code of numbers, lowercase and uppercase alphabets.

If we look at the binary code of numbers from 0 to 9, A to O, O to P, a to o and a to p the last 4 bits are different and first 4 bit are similar. So any number and alphabet can be represented by the last 4 bits and adding either ‘0’ or ‘1’ at the first position. To differentiate whether the character is number, uppercase alphabet or lowercase alphabet control symbols are used which is of the same type as that of number or alphabet.

For special symbols like !, “ , # , $, %, & , (, ,) , *, + , ‘, - , . , / is also observed and these special symbols can also be embedded in WAV file.

When embedding the textual information in any audio file, first the audio signal is converted into bits. Then the message to be embedded is converted from above strategy[4]. By applying LSB algorithm, the message is embedded into 16 bits or 8 bits audio sample. The performance is evaluated by applying LSB algorithm at different position i.e 1LSB, 2LSB and so on. At the receiver side, the first five bytes are taken, if these bytes are same as our control symbols bytes then the next character case is defined.

Encoding Algorithm and Decoding Algorithm

Encoding Algorithm

1. Input the text to be embedded.
2. Convert the text into 5 bit code by checking the redundancy in binary code of alphabets and numbers.
3. Read WAV audio file as cover file.
4. Select audio sample and hide the converted 5 bit code of the text in WAV file using LSB algorithm.
5. Repeat till the whole message can be embedded in audio.

Decoding Algorithm

1. Read the stego-object i.e. cover audio after encoding.
2. Extract the message by reading the control symbols in samples and reading LSB.

3. Select all samples and store all LSB position bits in array.
4. Divide the array into number of rows and columns
5. Display the secret message.

4. Experiment

This Steganpgraphy is implemented in Matlab 7. To measure the performance of proposed method , MOS (Mean Opinion Score strategy is used. Mean value is calculated by asking people about the difference in the original wav file and embedded wav file.This rating is done on 5 point scale. The LSB algorithm is tested for 1, 2, 3,4 ,5,6,7LSB position.

In order to evaluate the sound quality after embedding the secret image into audio files, two type of test are carried out: MOS and signal to noise ratio w.r.t. sample size, sample rate, bitrate, size of audio.

4.1 Mean Opinion Score (MOS)

Subjective quality evaluation for the text hiding in audio has been done by listening tests involving twenty persons. The audio files are categorized as per number of bits per sample, number of channels. The entire tests have been carried out at each category of sound files. Initially in the first part repeatedly presented the audio clips with hidden text and audio clips without hidden text into it, in random order to the listeners. Listeners were asked to determine which one is the audio with text hidden in it and without it. Here we have also calculate the mean opinion score of audio where different LSB positions of audio file i.e. simply 1 or 2 or 3 or 4 or 5 th LSB positions, and up to fourth 4th LSB positions were used for hiding image into it. To calculate the mean opinion score, five point scales is used by the individual after listening the music file and final mean of all scores is M.O.S. Mean Opinion Score for all four categories of sound are as shown in Table1. This five-point scale is defined in the following manner as given under, using a 5-point impairment scale:

5: Imperceptible
4: Slightly Perceptible but not noisy
3: Slightly noisy
2: Noisy
1: Very Noisy

Table1 shows that for 8 bit Brake wav file, 16 bit with 1 mono channel Originalrekam2 sound file, 16 bit 2 channel windows Shutdown audio file and 16 bit handel sound. Form table 1 proposed algorithm is close for 16 bit audio. For 8 bit audio with 1 mono channel, 11 KHz the algorithm is closer up to 5 LSB, for 16 bit audio with 1 mono channel ,22 KHz result is close up to 7 LSB, for 16 bit audio with mono channel,8Khz sample rate the best result up to 3 LSB and for 16 bit , 2 channel, 22 KHz sound the result is good upto 5 LSB. Using the statistical tool SPSS the mean value for 4 types of sounds up to 5 bits LSB are calculated as shown in table1.Table 2 shows the SNR for the audio samples and table 3 shows the comparison of SNR values with some known algorithm. [7]

Table1: MOS for Proposed algorithm

MOS at 1,2,3,4,5,6,7 LSB Position				
Methods	Brake	Originalrekam2	Handel	Shutdown
1 LSB	4.3	3.9	4.3	4.0
2 LSB	4.6	4.3	4.3	4.0
3 LSB	4.1	4.1	4.3	4.0
4 LSB	4.0	4.0	4.1	3.9
5 LSB	3.9	4.2	4.1	4.0
6 LSB	3.9	4.2	4	4.0
7 LSB	3.5	4.2	4	4.0

The second approach for testing is the quality evaluation using signal to noise ratio. For above 4 samples SNR is as shown in the following table. SNR is calculated by following formulae [8].

Mer (mean error rate) = coverfilebits – embeddedfilebits / coverfilebits
sizeinfo=size of the cover file
snr=(20*log10(sizeinfo / mer))

Table 2 : SNR for 4 sample audio

Audio	SNR	Size	Sample size	Sample rate
Brake	58.33	5781	8	11
Handel	73.75	73113	16	22
Originalrekam2	69.36	44100	16	22
WindowsXP shutdown	73.45	70641	16	22

Table 3: Performance Comparison with some known and similar technique

Researchers/ Algorithm	**SNR**
Pooyan and Delforouzi	39
Cvejik and Sepannen	42
Cvejik and Sepannen	39
Bao and Ma	36
Mansour Sheikhan	49
Proposed Algorithm	68

To see the result in terms of number of bytes required the 4 samples of 16 bit WAV audio is taken and its number of bytes to hide the text is measured. e.g To hide 5 character ordinary algorithm takes 5*8=40 bytes in Matlab and proposed algorithm takes 5*5=25 bytes.

5. Conclusion

In the proposed steganographic system, 16bitWAV and 8bitWAVaudio file are supported and the secret message can be hidden in the audio file with less storage capacity. The existing system requires the large storage capacity as the message is stored as it is ,so the proposed method requires less storage capacity as it requires less storage space instead of 8 bit code. Proposed algorithm gives better result for 16 bit wav audio as compared to 8 bit .

References

[1]C. Yeh, C. Kuo, (October 1999)Digital Watermarking Through Quasi M- Arrays, Proc. IEEE Workshop On Signal Processing Systems, Taipei, Taiwan, Pp. 456-461.

[2] Dr. D Mukhopadhyay, A Mukherjee, S Ghosh, S Biswas, P Chakarborty (2005.) An Approach for Message Hiding using Substitution Techniques and Audio Hiding in Steganography, IEEE

[3] E. Zwicker, “Psychoacoustics”, Springer Verlag, Berlin, 1982.

[4] Mazen Abu Zaher Modified Least Significant Bit (MLSB) Published by Canadian Center of Science and Education Vol. 4, No. 1; January 2011

[5] Nedeljko Cvejic, Tapio Seppben(,IEEE 2002) Increasing The Capacity Of LSB-Based Audio Steganography

[6] Steganographic Techniques and their use in an Open-Systems Environment- Bret Dunbar, The information Security reading Room, SANS Institute 2002http://www.sans.org/reading room/whitepapers/covert/677.php

[7]Mansour Sheikhan et al, High Quality Audio steganography by Floating Substitution of LSBs in Wavelet Domain, world applied science Journal IDOSI publications , 2010

[8] Y.V.N.Tulasi et al.,Steganography -Security through Images

[9] On Embedding of Text in Audio – A case of Steganography Pramatha Nath Basu, 2010 International Conference on Recent Trends in Information, Telecommunication and Computing

K.P.Adhiya is an associate professor and HOD of Computer science and engineering department of SSBT's COET Bambhori Jalgaon, Maharshtra, India. He has twenty one years teaching experience and pursuing PhD from North Maharashtra University, Jalgaon.

Swati A. Patil is a research scholar in SSBT's COET Bambhori. She has seven years of teaching experience. Currently she is working as lecturer in G. H. Raisoni Insitute of Information Technology, Jalgaon.