Audio Hiding Using Wavelet Transform with Amplitude Modulation

Waffaa S. Ahmed* and Loay E. George**
*Department of Shari'a, College of Education for Women, Al-iraqia University, Baghdad-Iraq.
**Department of Computer Science, College of Science, Baghdad University, Baghdad-Iraq.
*E-mail:waffaa_shehab@yahoo.com.
**E-mail: loayedwar57@yahoo.com.

Abstract

This research focus on using wavelet method for hiding secret data in the audio signal. The audio signal is transformed to frequency domain by using Haar wavelet transform. The wavelet coefficients are utilizied to host secret data with exact retrieval capability. For embedding process the magnitude modulation mechanism was applied. The introduced method makes checking for extracted hidden data to be sure that the host block can hold secret data, and can exactly reproduce the hidden data. The performance of hiding method was tested using the fidelity measures Mean Square Error (MSE), and Peak Signal to Noise Ratio (PSNR) to describe the overall error rate. Hiding rate was computed to asses the embedding power of this method. The test results indicated that the proposed system can give maximum hiding rate which is %5.10 when the value of threshold is 6, quantization step is 8 and block size is 16. The maximum PSNR is 42.94 when the block size value is 16, threshold value is 6, quantization step is 2, and don't make subjective deterioration in quality of audio.

Keywords: Audio steganography, Transform coding, Wavelet coding, Modulation.

Introduction

Data hiding has got great attention as a promising method that plays a complementary role to the conventional cryptographic techniques. Digital steganography is based on the fact that artifacts like bitmaps and audio file contain redundant information. The idea behind digital steganography is replace some of redundant information with other data [1]. One should not underestimate the difficulty of hiding data in audio signals. Bender explained the depth of the problem when he noted that the human auditory system (HAS) operates over a wide dynamic range. The HAS perceives over a range of power greater than one billion to one and a range of frequencies greater than one thousands to one. The HAS is fairly poor with respect to its differential range. This means that loud sounds tend to mask out softer sounds easily [2].

One of the techniques that used for hiding secret data in multimedia covers is wavelet transform. Wavelet domain based techniques are becoming very popular because of the developments in the wavelet stream in the recent years [3]. Information is embedded in the wavelet domain by modifying selected wavelet coefficients of the host image, and by using wavelet transform one can hide data into wave audio file cover. Wavelet transform is used in some audio application, such as it is used for audio denoising, classification, identifying transient speech, also it is used in compression methods [4,5].

Many methods have been introduced for hiding data in audio digital media. Sridevi et al in 2009, have proposed enhanced audio steganography (EAS) system. Which applied both in steganography and cryptography. It uses a powerful encryption algorithm as the first level of security, and in the second level, it uses a more powerful modified least significant bit (LSB) algorithm to encode the message into audio [6]. Alkawaz in 2007, have developed two hiding schemes for embedding secret data bits in fourier (specifically, phase) coefficients of the audio signal. The first scheme called Hiding in Voiced Blocks (HVB) hides the secret data in voiced blocks only, and the second scheme called Hiding in Checked Blocks (HCB) hides secret data in blocks which can host the data and later could retrieve data from them without error [7].

The aim of this research is to design and implement a steganography system that hides secret data in the blocks of audio data using wavelet transform domain, only to audio block can preserve the secret data are chosen as host blocks.

1. The Proposed System

The structure of the proposed system is shown in Fig.(1). It consists of two basic modules: hiding and extraction modules. The input to this system are the cover file (wave file), and secret file (binary file). These inputs are processed in the hiding module with various operations are applied to produce stego wave file. The stego audio could passed through the extraction module, and after passing a set of operations, the secret data is retrieved. The following sub-sections explain the action of each stage in the both modules of proposed system:

1.1 Hiding Module

The hiding module consists of a sequence of processes. As a first stage, the cover audio data are partitioned into blocks of samples. The block size must previously be predefine by the user, and must be not so large to avoid the occurrence of high computational complexity. These blocks are classified to voice and unvoiced blocks, only the voice blocks are used as host block.

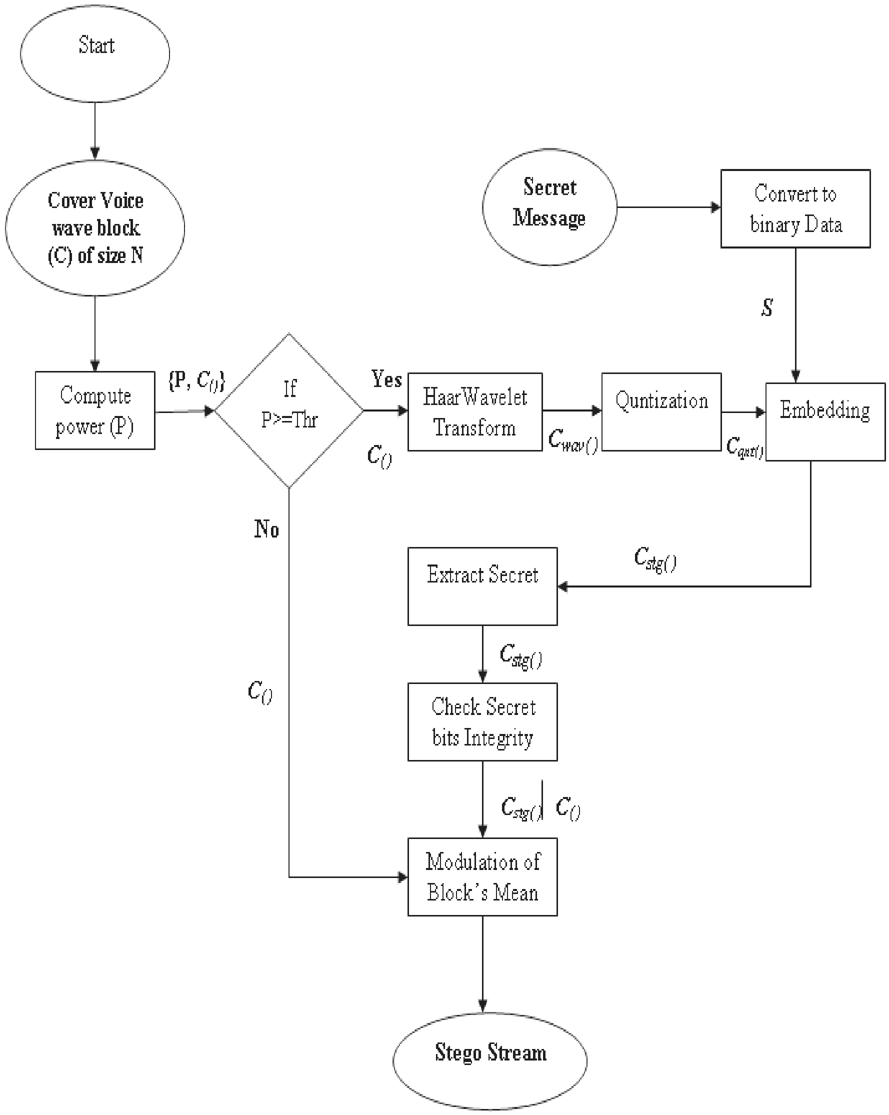

Fig. (1) The layout of hiding and extraction stages.

A. Voice/Unvoice Classification

The hiding method in this research depends on hiding in the voice area of the cover sound and ignores unvoiced area. The implemented voiced and unvoiced classification was based on measuring the voice power for each block of cover wave data, and compare this power with a predefined threshold value (*Thr*) to decide if the tested block is considered voiced block or not.

B. Wavelet Transform

Wavelet transform is used to convert the spatial domain into frequency domain. Haar Wavelet Transform is the simplest variant of wavelet transform. In this transform the low frequency wavelet coefficients are generated by averaging each pair of adjacent audio samples value, and high frequency coefficients are generated by taking half of the difference of the same two samples. The two produced bands are called the low frequency (L) and high frequency (H). The L-consists of low frequency wavelet coefficients, it contains the significant part of information of the spatial domain. While H-consists of high frequency coefficients, which contain the details of the spatial domain [8]. It is decided for the secret data to be hidden in high frequency domain.

C. The Quantization

Quantization is a process comes after forward haar wavelet transform, it is applied on the high frequency coefficients. The type of applied quantization was the uniform scalar quantization, where each sample is treated individually. The high frequency coefficients are uniformly quantized in order to establish suitable uniform slack spaces with their values, these slack space are utilized as hosting areas, to hide secret data. The gaps made between the quantized values of high frequency coefficients should be enough to carry the added values of secret data. The uniform scalar quantization was done using following equation:

$$C_{qnt}(i)= Q\times \text{round}\left(\frac{C_{wav}(i)}{Q}\right) \quad \text{..........................(1)}$$

Where:

$C_{wav}(i)$ is the i^{th} original wavelet (detail) coefficients.

Q is the quantization step.

$C_{qnt}(i)$ is the corresponding i^{th} quantized coefficients.

N is the block size.

and i= N/2 , …….., N-1

D. The Embedding

In this stage, the hiding of secret block is done on the voiced blocks, exclusively. So, the implementation of Haar wavelet transform and quantization process (to prepare the slack space) is done on the voiced blocks. The embedding process is done by adding or subtracting a value (Δ) to quantized coefficients. The value of (Δ) should be less than half the value of quantization step, in order to avoid the occurrence of a jump from certain quantization bin to the adjacent one (previous or next) as shown in Fig.(2). The application of addition or subtracting processes depends on the value of the secret bit (whether it is 0 or 1). The quantity (Δ) is added to quantized coefficient (C_{qnt}) if the secret data value (*S*) is 0, or it is subtracting if (*S*) is 1, as follows:

$$C_{wstg}(u)=\begin{cases} C_{qnt}(u) - \Delta & \text{if } S(i) = 1 \\ C_{qnt}(u) + \Delta & \text{if } S(i) = 0 \end{cases} \quad \text{..........(2)}$$

Where:

$C_{qnt}(u)$ is the u^{th} quantized coefficient in the block.

$C_{wstg}(u)$ is the u^{th} host coefficient in the block.

S (i) is the i^{th} secret bit.

N is the block size.

u= N/2,….., N-1.

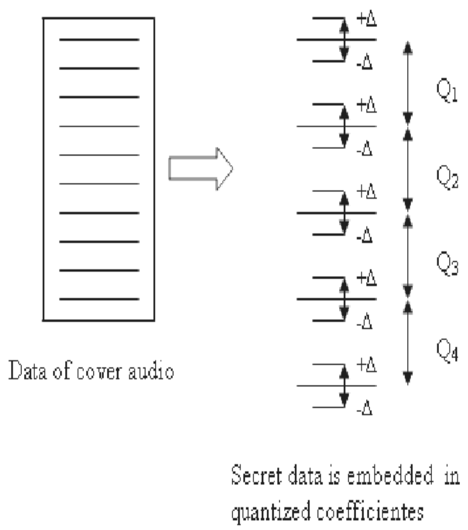

Fig. (2) The Embedding Stage.

1.2 Extraction Module

The extraction operation is similar to that mentioned in hiding stage but it is arranged in reverse manner, example instead of embedding the secret data it is extracted.

A. Stego-bit Extraction

In this stage Fig. (3) the forward wavelet transform applied on the stego coefficients (C_{stg}) and extraction operation applied on the same coefficients, and secret data extracted.

The extraction of the secret bits (either 1 or 0), is done according to the following equation:

$$\acute{S}(i) = \begin{cases} 0 & if\ C_{wav}(u) > C_{qnt}(u) \\ 1 & Otherwise \end{cases} \quad ....(3)$$

Where:
$C_{wav}(u)$ is the u^{th} Haar wevelet coefficient of the stego audio block
$C_{qnt}(u)$ is the u^{th} quantized coefficients
$\acute{S}(i)$ is the i^{th} retrieved secret data
and u= N/2,……, N-1

B. Integrity Check and Mean Modulation

In this method Fig.(3) not all the embedded secret bits are retrieved correctly, a little amount of secret bits may retrieved wrong. For this reason, the mean value for each block computed and then checked, if odd or even value. If it is even then the block is considered empty from secret data, otherwise it is considered a host block.

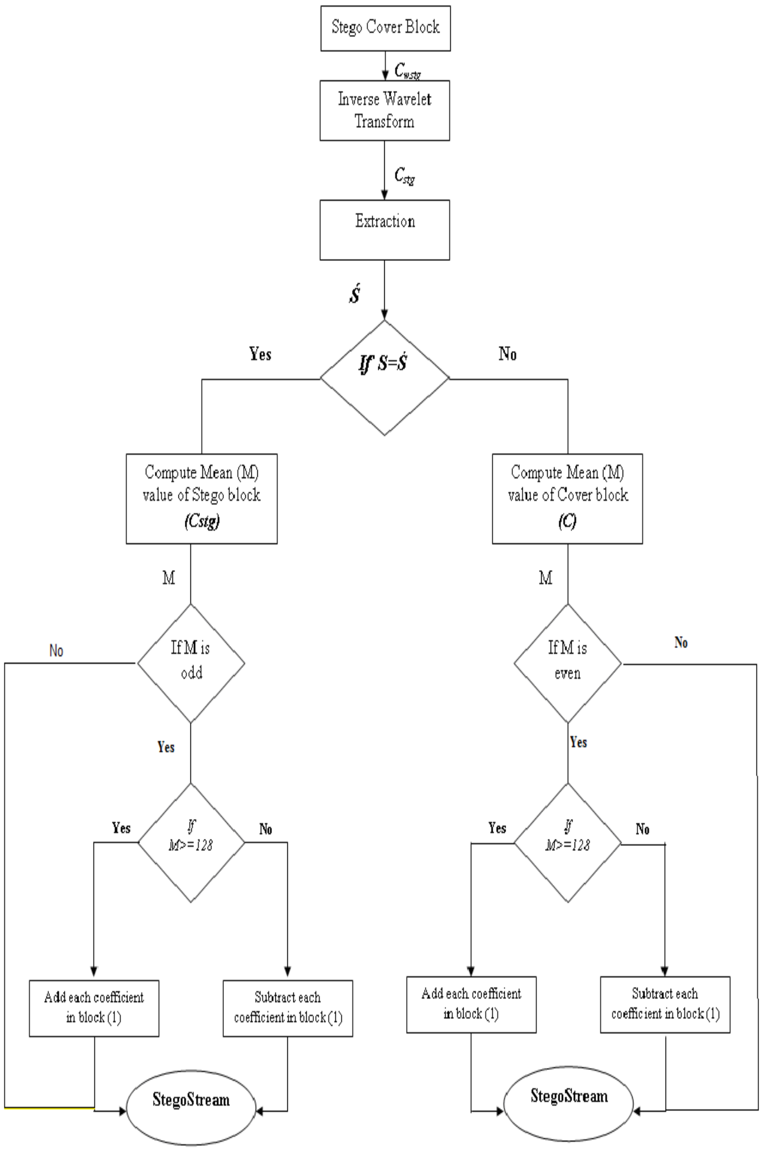

Fig. (3) Integrity Check and Mean Modulation Stage.

2. Test Measures

The measures used in this research to assess the performance of the considered hiding methods are Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) for the stego cover signal and the Hiding Rate (HR), they defined as follows:

$$MSE = \frac{1}{M}\sum_{i=1}^{M}(S_i - S_i^{'})^2 \text{ ....(4)}$$

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) \text{ ....(5)}$$

Where ,

S_i is the i[th] sample in the audio cover data.

$S_i^{'}$ is the corresponding i[th] sample in the audio stego-cover data.

M is the total number of audio samples.

$$HR = \frac{S_t}{8 \times M} \times \%100 \text{ ....(6)}$$

Where,

S_t is the maximum possible number of inserted bits in the cover media.

M is the size of cover media in bytes.

3. Test Results

In this research, the control parameters that have significant effects on the hiding rate and ratio of the correctly retrieved secret bits: are the threshold value, quantization step, block size. In the following sub sections the effects of the above effective parameters on the performance of the hiding method are discussed:

3.1 Threshold Value

The decision of voiced/unvoiced sound classification depends on the value of threshold (Thr). If the average energy of the tested audio block is higher than threshold (Thr) then the block is classified as voiced block, otherwise it is classified as unvoiced. Then the voiced blocks are used as host area for hiding the secret data. Table (1) illustrates the effect of threshold (Thr), embed the data of (*Doc2Se*c) file has a size is 41.5 kB in the cover (*music1*) audio cover file whose size is 70.3 kB.

Table (1)
Effect of the threshold (Thr) on the performance of hiding method, (N=16, Q=8).

Thr	*Hiding Rate (HR)*	*MSE*	*PSNR*
1	4.35	5.27	40.91
2	4.25	5.22	40.95
4	3.98	5.10	41.05
6	3.66	4.97	41.15

3.2 Quantization Step

Quantization process provides suitable slacked spaces to hide secret data. The main parameter used to manage the amount of slack space is the quantization step (Q). If the value of Q is increased then the slack space used for hiding increases, and peak signal to noise ratio decreased, but the same time it causes an increase in the distortion level of the stego-cover signal. Table (2) presents the effect of quantization step (Q) on the hiding method, the contents of the (DocSec) file where the size = 19.5 kB, was taken as secret data, and (music2) audio cover file which the size of it = 70.3 kB.

Table (2)
Effect of quantization step (Q) on the performance of hiding method (N= 16, Thr= 6).

Q	*MSE*	*PSNR*
2	3.29	42.94
4	4.27	41.82
6	5.23	40.94
8	6.67	39.88

3.3 Block Size

Table (3) shows the effect of the audio block size (N), we can notice that numbers of inserting bits are increased when value of (N) become large, and peak signal to noise ratio decreased, but the same time it causes an increase in the distortion level of the stego-cover signal. The secret data was the content of the (DocSec) file has a size is 19.5 kB embed in audio (music3) file whose size is 80 kB

Table(3)
Effect of the block size (N) on the performance of hiding method (Thr=6, Q=8).

Block Size (N)	*Hiding Rate (HR)*	*MSE*	*PSNR*
8	5.00	7.03	39.66
12	5.06	7.90	39.15
16	5.10	8.63	38.76

4. Conclusion

Hiding in voiced block sample is more suitable to avoid noise occurrence, large threshold value filters in only the high power blocks of the cover audio signal, and this will avoid the unvoiced blocks, and consequently increases the probability of correct retrieved bits. But it decreases the hiding rate, large quantization step provides more slacked space to hide secret bits, but it causes more distortion in cover audio and the integrity check and mean value of the voiced block is imposed to ensure that all the embedded secret bits are retrieved correctly.

References

[1] Mark, N.; "*MP3Stego: Hiding Text in MP3 Files*"; SANS Institute Reading Room site; GIAC paper. P5; 2001. www.sans.org/reading_room/.../mp3stego-hiding-text-mp3-files_550

[2] Bender, W.; Gruhl, D.; Morimoto, N.; Aiguolu; *"Techniques for Data Hiding"*; IBM Systems Journal; Vol. 35, nos. 3&4; pp. 313-336; 1996.

[3] Mohamed, I.; Moawad, M.; Salah, D.; and Fatma, H.; "*Wavelet Data Hiding Using Achterbahn-128 on FPGA Technology*"; Ubicc Journal-Special Issue of IKE'07 conference; Vol. IKE'07 – Special Issue; 2008.

[4] Matko, S.; Luki, B.; and Hrvoje, D.; "*White Noise Reduction of Audio Signal Using Wavelet Transform with Modified Universal Threshold*"; Proceeding of the 4th WSEAS International Conference on Signal Processing, Robotics and Automation; p.279, Vol.2, Issue.2; 2005.

[5] Daniel, M. R.; "*Identification of Transient Speech Using Wavelet Transform*"; M.Sc. Thesis Graduate faculty of school of Engineering, University of Pittsburgh; 2005.

[6] Sridevi; Adamodaram; and Svl, N.; "*Efficient Method of Audio Steganography By Modified LSB Algorithm and Strong Encryption Key with Enhanced Security"*; Journal of theoretical and applied information technology; Vol.5, no.6; pp. 768-771; 2009. www.Jatit.org.

[7] Hiba, M. S. Alkawaz; *"Low Rate Hiding in Audio Data Using Phase Domain"*; Computer Science Department, College of Science, Al-Nahrain University; 2007.

[8] HS Manjunatha R.; Raja KB; "*High Capacity and Steganography Using Discrete Wavelet Transform"*; International Journal of Computer Science and Security (IJCSS); Vol. 3; pp. 462-472; 2009.

الخلاصة

يهدف البحث إلى استخدام طريقة التحويل المويجي لإخفاء بيانات سرية في الإشارة الصوتية. في البدء الإشارة السمعية تحول إلى المدى الترددي باستخدام تحويل المويجي هار. المعاملات المويجية استخدمت لإخفاء البيانات السرية مع مقدرة استرجاعها بشكل كامل. في عملية الإخفاء استخدمنا آلية تضمين القيمة. في هذه الطريقة عملنا اختبار على البيانات المخفية للتأكد إذا ما كان هذا المقطع يحوي على بيانات سرية أم لا، وإمكانية استرجاعها بشكل كامل. في النهاية تم اختبار أداء طريقة الإخفاء باستعمال مقاييس الدقة، متوسط مربع الخطأ (MSE) ونسبة الإشارة إلى الضوضاء (PSNR) لتحديد نسبة الخطأ وتم استخدام نسبة الإخفاء لتقييم قوة الإخفاء. نتائج الاختبار حددت أعلى قيمة متحققة لنسبة الإخفاء هي ٥.١% عندما كانت قيمة threshold هي ٦ وقيمة quantization step هي ٨ وحجم المقطع هو ١٦. وأعلى PSNR هو ٤٢.٩٤ مع حجم مقطع ١٦ وقيمة threshold هي ٦ و quantization step هي ٢.