WAVE STEGANOGRAPHIC ALGORITHM BASED ON WAVELET TRANSFORM +

خوارزمية فن الاختزال للموجة باستخدام التحويل المويجي

Ali Jaber Abdul Wahab *

Abstract:

In this paper, the secret color image is embedded in speech signal using discrete wavelet transform (DWT) and wavelet packet transform (WPT). The information bits are hidden in the LSB (s) of discrete speech signal. Different number of parameters are changed such as level of wavelet transform (level 2 and 4), types of wavelet family (daubechies (Db4), coiflet (coif4) and symmlet (sym4)), the length of filter (Db2, Db6, Db8, Db10) and the number of bits in the LSB (1bit, 2bits and 3bits) where information to be hidden in it. The results show that WPT gives better performance than DWT.

المستخلص:

في هذا البحث تم تضمين و اخفاء صورة ملونة داخل اشارة صوتية بأستخدام طريقتين لتحويل المويجـــة، الطريقـــة الاولى هي تحويل المويجةالمقطعة(DWT) بينما كانت الطريقة الثانية هي التحويل الحزمي للمويجة(WPT). حيث تم اخفاء بتات المعلومات في البتات الاوطئ مرتبة(LSB's) من اشارة الصوت المقطعة. حيث تم تغيير العديـــد مـــن المتغيرات وهي تشمل مستوى تحويل المويجة (٢ و ٤)، نوع عائلة المويجـــة ((Db4),(coif4),(sym4))، طـــول الفلتر المستخدم (Db2, Db6, Db8, Db10) و اخيرا عدد البتات الاوطئ مرتبة(LSB) (1bit, 2bits , 3bits) و التي سيتم اخفاء المعلومات فيها. حيث ان النتائج اظهرت ان اداء التحويل الحزمي للمويجة(WPT) افـــضل مـــن تحويل المويجة المقطع(DWT).

Introduction

Steganography is a technology that hides a message within an object, a text, picture, audio or video. It is often confused with cryptography, not in name but in appearance and usage [1]. The word steganography comes from the Greek steganos (covered or secret) and graphy (writing or drawing) and means, literally, covered writing [2]. The general idea of hiding some information in digital content has a wider class of applications that go beyond steganography, the techniques involved in such applications are collectively referred to as information hiding. For example, an image printed on a document could be annotated by metadata that could lead a user to its high resolution version [3].
When hiding information inside Audio files the technique usually used is low bit encoding which is somewhat similar to LSB that is generally used in Images.
Spread Spectrum is another method used to conceal information inside of an audio file. This method works by adding random noise to the information signal that is concealed inside a carrier and spread across the frequency spectrum. Echo-data hiding is yet another method of hiding information inside an audio file. This method uses the echoes in sound files in order

+ **Received on 11/10/2010 , Accepted on 10/7//2011**
* **Assist Lech ./College of Electrical and Electronic Techniques**

to try and hide information. By simply adding extra sound to an echo inside an audio file, information can be concealed [4], [5].

Wavelet Transform:

1- Discrete Wavelet Transform (DWT)

The general form of an L-level DWT is written in terms of L detail sequences, $d_j(k)$ for j=1,2,….,L, and the L-th level approximation sequence, $c_L(k)$ as follows [5]:

$$f(t) = \sum_k c_L(k)\, \phi_L(t) + \sum_{j=1}^{L} \sum_k d_j(k) \psi_j(t) \tag{1}$$

Where $\phi_L(t)$ is the L-th level scaling function and $\psi_j(t)$ for j=1,2,…,L are wavelet function sequences for L different levels, and k is the frame index.

The approximation and detail sequences at level j are given by [6]:

$$c_{j+1}(k) = \sum_m h_o(m-2k) c_j(m) \tag{2}$$

and

$$d_{j+1}(k) = \sum_m h_1(m-2k) c_j(m) \tag{3}$$

Equations (2) and (3) state that approximation sequence at higher scale (lower level index), with the wavelet and scaling filters, $h_o(t)$ and $h_1(t)$ respectively, can be used to calculate the detail and approximation sequences (or discrete wavelet transform coefficients) at lower scales.

The scaling coefficients are related to wavelet coefficients by:

$$h_1(k) = (-1)^k h_o(R-k) \tag{4}$$

Where R is a finite odd length of quadrature mirror filter.

Figure (1) shows filter bank of discrete wavelet transform. The symbol ↓2 is down-sampler (decimator) that it takes a signal s(n) as input and produces an output of s(n)=s(2n), which means half of the data is discarded [6].

The wavelet reconstruction process consists of up-sampling and filtering up-sampling process of lengthening signal component by inserting zeros between samples. In other words, the input signal is stretched twice its original length and zeros are inserted in the even numbered samples. The inverse discrete wavelet transform is illustrated in Figure (2). [6].

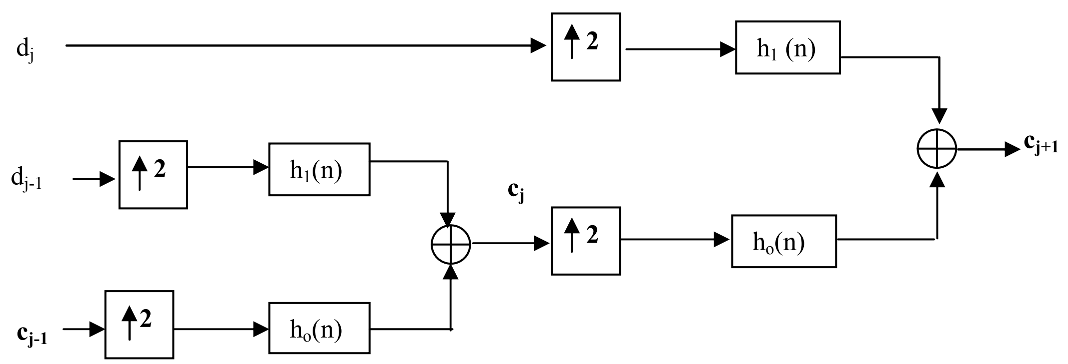

Figure (2) Two stage two-band tree signal synthesis tree of IDWT

2- Wavelet Packet Transform (WPT)

Wavelet packet transform is a wavelet transform where the signal is passed through more filters than the DWT. Wavelet packets are the particular linear combination of wavelets. They form bases which retain many of the orthogonality, smoothness, and localization properties of their parent wavelets. The coefficients in the linear combinations are computed by a recursive algorithm making each newly computed wavelet packet coefficient sequence the root of its own analysis tree. In the DWT, each level is calculated by passing the previous approximation coefficients though a high and low pass filters. However, in the WPD, both the detail and approximation coefficients are decomposed [6].

For n levels of decomposition the WPT produces 2^n different sets of coefficients (or nodes) as opposed to (n +1) sets for the DWT. However, due to the down sampling process the overall number of coefficients is still the same and there is no redundancy. Figure (3) shows the wavelet packet decomposition for 2 levels.

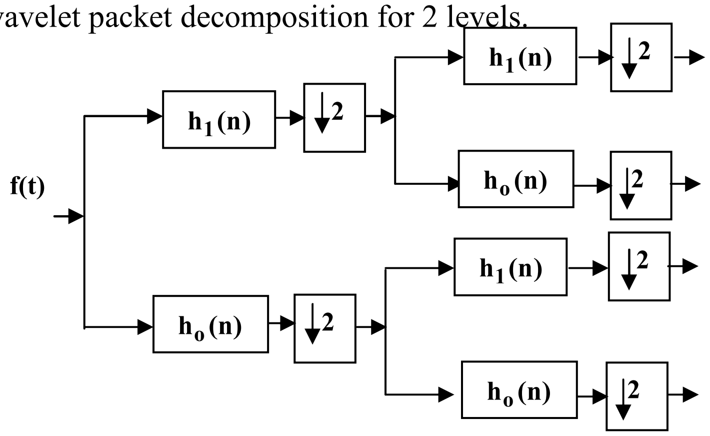

Figure (3) Wavelet Packet Decomposition Tree

The proposed Wave Steganographic System

1- Embedding process of an image into a speech signal

The block diagram of the proposed embedding process of a secrete image into a speech signal is shown in Figure (4).

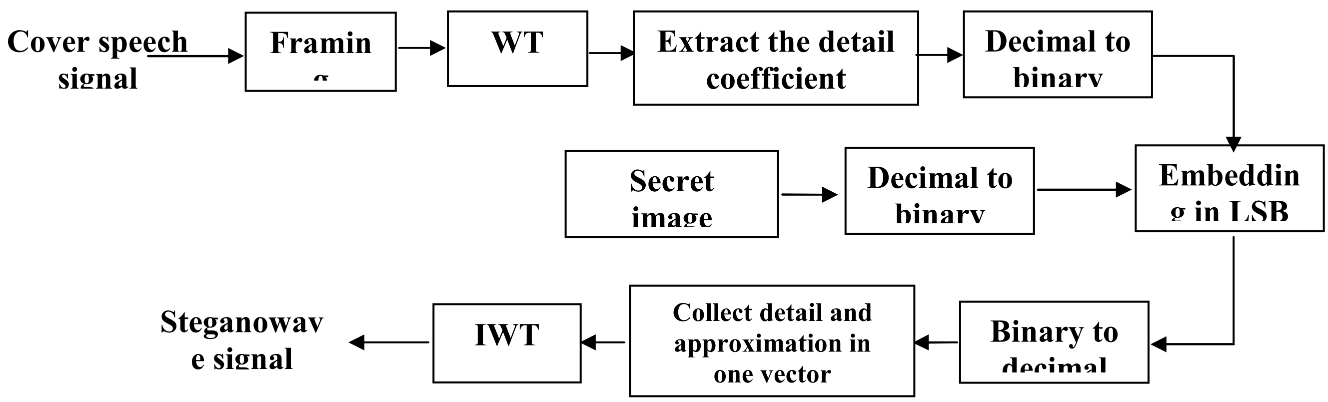

Figure (4) Embedding process of a steganowave signal

The cover speech signal is sectioned into frames (typical value of frame length is 256 ... signal, ... binary signal of 16 bits. During that the secret image which needed to be hidden is converted into stream of bits and will be embedded in the least significant bit (s) of the covered signal. Then the cover signal is back to decimal signal. Finally the detail and approximate coefficients are passing through inverse wavelet transform (IWT) to get the steganowave signal.

2- Extracting process of an image from a speech signal

The block diagram of extracting process of a secrete image from a speech signal is shown in Figure (5).

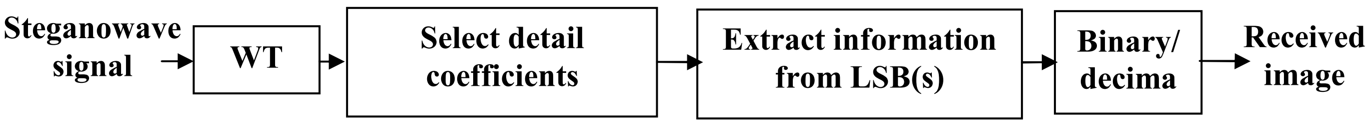

Figure (5) Block diagram of extracting process

To recover the secret image from the steganowave signal, the steganowave frame is passing through the wavelet transform (either DWT or WPT) and after that the detail coefficients are selected and the secret binary image is extracted from the LSB(s). Finally the results of the stream bits are converted to decimal signal to get the original secret image.

Performance Measures:

Number of quantitative parameters can be used to evaluate the performance of speech coder in terms of both reconstructed signal quality after decoding and compression scores. The following parameters are used [5].

a- Signal to Noise Ratio (SNR)

This is defined per frame as:

$$SNR = 10 \log_{10} \frac{\sum_{n=0}^{N-1} x_r^2(n)}{\sum_{n=0}^{N-1} \left(x(n) - x_r(n)\right)^2} \qquad (5)$$

Where

N: frame size.

x(n): actual speech.

$x_r(n)$: reconstructed speech.

b- Peak Signal to Noise Ratio (PSNR)

$$PSNR = 10 \log_{10} \left[\frac{NX^2}{\|X - Xr\|^2} \right] \qquad (6)$$

X is the maximum absolute square value of the signal x, and

$\|X - Xr\|^2$ is the energy of the difference between the original and reconstructed signal.

c- Mean Square Error

$$MSE = \frac{\sum_{n=0}^{N-1} \left(x(n) - xr(n)\right)^2}{N} \tag{7}$$

d- Normalized Root Mean Square Error

$$NRMSE = \sqrt{\frac{\overline{(x(n) - x_r(n))^2}}{\overline{(x(n) - \mu_x(n))^2}}} \tag{8}$$

$\mu_x(n)$ is the mean of the speech signal.

e- Correlation

$$COR = \frac{\sum_{n=0}^{N-1} x(n)\, x_r(n)}{\sqrt{\sum_{n=0}^{N-1} x^2(n) \sum_{n=0}^{N-1} x_r^{\,2}(n)}} \tag{9}$$

<u>Simulation Results:</u>

A cover speech signal is sampled at 8 KHz with 16 bits/sample. This signal is framed using 256 samples per frame. Therefore, there are 32 frames per second. Two techniques are studied, the first technique is based on Discrete Wavelet Transform (DWT) and the second technique is based on Wavelet Packet Transform (WPT). Both techniques are used at different levels and different types of wavelet families. The secrete image signal is jpeg type with size of 116 x160. Figures (6), (7) show the cover speech signal and secrete image respectively.

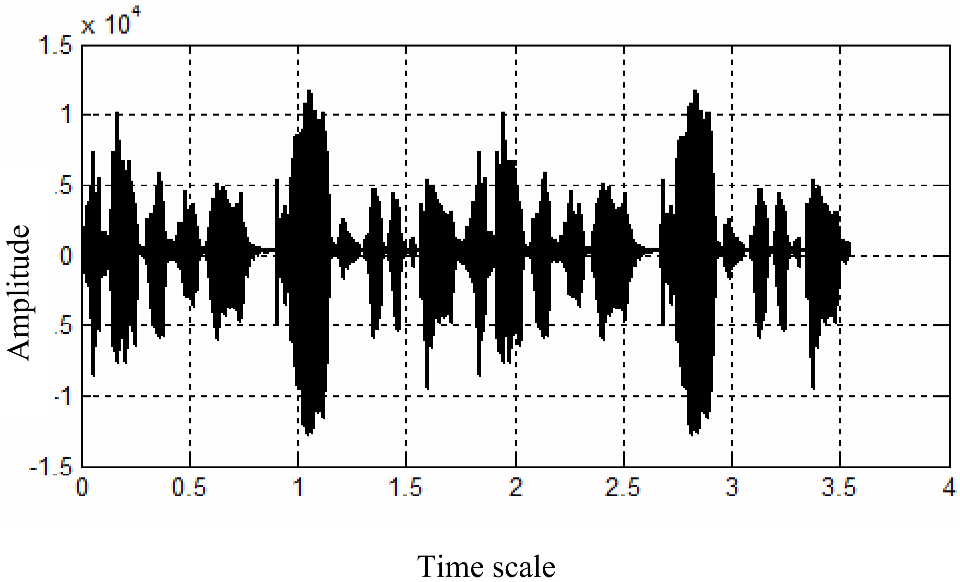

Figure (6) the cover speech signal

Figure (7) the secret image signal

Note : the value of NRMSE in the following tables appeared to be very low compared to the other values of(MSE, SNR,PSNR) this gap is a result of the normalized value for root mean square error formula as calculated from equation(8).

- Table (1) shows the effect of changing the levels of DWT and WPT taking into consideration the information will be inserted into the last LSB only with Db8.

Table (1) Comparison between DWT and WPT using different levels

measure type	DWT		WPT	
	Level 2	Level 4	Level 2	Level 4
SNR	34.3572	33.8576	49.0542	50.1795
PSNR	164.9112	164.3837	194.8734	195.0589
NRMSE	5.4054e-004	5.7439e-004	1.7168e-005	1.6805e-005
COR	1.0000	1.0000	1.0000	1.0000
MSE	0.8811	0.9949	8.8883e-004	8.5166e-004

- Table (2) shows the effect of embedding bits at the last two or three LSBs when the number of levels equals four and Db8 are considered.

Table (2) Comparison between the last two and three LSBs

measure type	Last two bits		Last three bits	
	DWT	WPT	DWT	WPT
SNR	32.6910	53.3492	32.2571	47.5137
PSNR	158.6813	187.6460	152.9133	182.7859
NRMSE	0.0011	3.9454e-005	0.0022	6.9040e-005
COR	1.0000	1.0000	1.0000	1.0000
MSE	3.6985	0.0047	3.9578	0.0144

- Table (3) shows the effect of changing the length of Daubechies filter on WPT when the information inserted at the last LSB(1 bit) and the number of levels is four

Table (3) the effect of changing the filter length of daubechies

Measure type	Db2	Db6	Db8	Db10
SNR	47.9104	44.6614	53.3492	46.7956
PSNR	180.6016	181.2373	187.6460	183.4304
NRMSE	8.8779e-005	8.2514e-005	3.9454e-005	6.4102e-005
COR	1.0000	1.0000	1.0000	1.0000
MSE	0.0238	0.0205	0.0047	0.0124

- Table (4) shows the differences between three types of wavelet family (daubechies (Db4), coiflet (coif4), and symmlet (sym4)) on WPT when the information is inserted at the last LSB(1 bit) and the number of levels is four.

Table (4) the different types of wavelet family using WPT

Measure type	Db4	Coif4	Sym4
SNR	44.6288	46.9043	44.3063
PSNR	180.8988	183.2193	181.1389
NRMSE	8.5793e-005	6.5679e-005	8.3454e-005
COR	1.0000	1.0000	1.0000
MSE	0.0222	0.0130	0.0210

Conclusions:

The following are a summary of the conclusion remarks.

- WPT gives better performance than DWT, this result from table (1) by comparing the two methods for two levels(level 2 & level 4)
- Increase number of levels in DWT will be degraded the performance slightly. While in WPT, the performance results will be increased slightly as shown in table (1).
- The capacity of hidden information can be increased by hidden information in the last two or three LSB instead of only last LSB but the performance measurement is decreased as shown in table (1), (2) and (3).
- The different error functions that is used in table (2) shows that WPT method is better than DWT in the case of hiding the information in the last three bits rather than last two bits.
- Increasing the length of filter is not a necessary condition to improve the performance measurement but from Table (3) Db8 gives the best performance, this is called optimum solution.
- The system is affected by the wavelet family types. Coiflet filter gives the best performance as compared to the other types of wavelet family.
- The correlation always equals 1 because of the NRMSE error is very low and the reconstructed speech is almost the same actual speech
- The relationship between the error parameters is:

SNR ↑ PSNR ↑ NRMSE ↓ MSE ↓

References:

1- Alain C. Brainos II, “A Study of Steganography and the Art of Hiding Information” - East Carolina University – 2004.
2- Jonathan Watkins, “Steganography - Messages Hidden in Bits”, Multimedia Systems Coursework, Department of Electronics and Computer Science, university of Southampton, SO17 1BJ, UK - 15th December 2001. http://mms.ecs.soton.ac.uk/mms2002/papers/6.pdf
3- Mehdi Kharrazi. Husrev T. Sencar, and Nasir Memon, “Image Steganography: Concepts and Practice”, Department of Electrical and Computer Engineering, Department of Computer and Information Science Polytechnic University, Brooklyn, NY 11201, USA - April 22, 2004.
4- Aelphaeis Mangarae, “Steganography FAQ”, Zone-H.Org, China - March 18th 2006.
5- Muhalim Mohamed Amin, Subariah Ibrahim, Mazleena Salleh, Mohd Rozi Katmin, “INFORMATION HIDING USING STEGANOGRAPHY” University Technology Malaysia, Department of Computer System & Communication Faculty of Computer Science and Information system, 2003
6- Fadhil S. Hasn, “Speech Compression using Discrete Wavelet Transform with Optimum Differential Pulse Code Modulation (DPCM)”, IJCCCE, vol.8, NO.1, 2008.
7- K.P. Soman, K.I. Ramachandran, “Insight into Wavelets from Theory to practice”, Second Edition, PHI, 2005.
8- Michał Dyrek. Jakub Gałka. Bartosz Ziółko, “Measures on Wavelet Segmentation of Speech“, International Journal of Circuits, systems and signal processing. Issue 2, Vol. 1, 2007.