Color Image Hiding In Cover Speech Signal By Using Multi-resolution Discrete Wavelet Transform

BY

Lecturer Dr. Samir J.AL- Muraab
Electrical Engineering Dept.
Babylon University
Email:Dr_S_J_AL-Muraab@Yahoo.com

Asst. lecturer Haider I. AL–Mayaly
Electrical Engineering Dept.
Babylon University
Email: haider_almayaly @ yahoo.com

ABSTRACT:

Data hiding, a form of steganography, embeds data into digital media for the purpose of identification, annotation, security, and copyright. The goal of steganography is to avoid drawing suspicion to the transmission of a hidden message. Digital audio provides a suitable cover for high-throughput steganography. At 16 bits per sample and sampled at a rate of 44100 Hz, digital audio has the bit-rate to support large color image messages. In addition, audio is often transient and unpredictable, facilitating the hiding of messages. In this paper a high robustness system against the attackers in hiding of color images is presented. We used the multi-resolution discrete wavelet transform in hiding process. The JPEG format type for color images and WAV format for speech cover signal that used in test of system .Programs and graphics are executed by using MATLAB version 6.5 programs.

أخفاء صورة ملونة في أشارة غطاء كلامية بأستخدام تحويل المويجات المنفصلة متعددة الدقة

د. سمير جاسم المرعب
مدرس

حيدر إسماعيل الميالي
مدرس مساعد

قسم الهندسة الكهربائية / كلية الهندسة / جامعة بابل

خلاصة البحث:

أخفاء البيانات ، شكل من أشكال فن الاختزال، يُضمّنُ البياناتَ إلى الاوساط الرقميةِ لغرضِ التعريفِ، تذييل، أمن، وحقوق طبع. إنّ هدفَ فن الاختزالِ أنْ يَتجنّبَ سَحْب الشكِّ إلى الرسالة المخفية المرسلة. التسجيل الصوتي الرقمي يُزوّدُ غطاء مناسب لفن الاختزالِ وبطاقة استيعابيةِ عالية للمعلومات المخفية. عند 16خانة رقمية لكلّ عيّنةِ ونسبة تردد التقطيع 44,100 هيرتز، التسجيل الصوتي الرقمي لَهُ نسبةُ استيعابية جيدة لدَعْم رسائلِ الصورةِ الملونةِ الكبيرةِ. بالإضافة الى ذلك فأن التسجيل الصوتي عابر ومتقلّبُ في أغلب الأحيان، لذلك يُسهّلُ إخفاء الرسائلِ. في هذا البحث تم تقديم نظام عالي المتانة ضدّ المهاجمين في عمليات إخفاء الصور الملونة. إستعملنَا تحويل المويجات المنفصلة المتعدّدة الدقة في عملية الاخفاء. إنّ نوعَ صيغةِ الصور الملونة التي إستعملتْ في إختبار المنظومة هي JPEG وصيغةِ الموجةِ لإشارةِ الغطاءِ الكلامية هي WAV . البرامج والرُسوماتِ نفذت بإستعمال نسخةِ برنامج مُنَفّذِ Matlab نسخة 6.5.

KEYWORDS: Multi-Resolution Discrete Wavelet Transform MRDWT, Hiding System, Show Hiding System

1- INTRODUCTION

Over the past few years, increasingly sophisticated techniques for information hiding (steganography) have been rapidly developing (see [1–3] for general reviews). These developments, along with high-resolution carriers, pose significant challenges to detecting the presence of hidden messages. There is, nevertheless, a growing literature on steganalysis. [4–7] While much of this work has been focused on detecting steganography within digital images, digital audio is a cover medium capable of supporting high-throughput steganography; sampled at 44,100 Hz with 16 bits per sample. 10% of the cover signal could be used to the length of hiding massage in high level security information hiding, and 40% of the cover signal could be used to the length of hiding massage in low level security information hiding. These percentages are very large with respect to the other cover types.

In this work a color image of JPEG format used as a massage that is hiding in speech cover signal of WAV format. The size of image used in system proportional with the length of speech and the security of system increase with the small size of image and large length of speech cover signal. Because of speech is often transient and unpredictable, we obtain a high level system security.

Our statistical model begins by converting the color image matrix to one dimension vector then decomposing the cover function by using multi-resolution wavelet transform and then insert the information vector of image in noisy parts of signal after segmentation process. The recovering of hidden image system also presented. Many tested example used in hiding and recovering hidden message systems.

2- Multi-resolution Formulation of Wavelet System

Multi-resolution Analysis uses the wavelet transform to decompose a data series in a cascade from the smallest scales to the largest. At each scale, there are two components: the Smooth (or low-pass filtered) data series, and the Details (or high-pass) data series [8].

This multi-resolution formulation is obviously designed to represent signals where a signal event is decomposed into finer and finer detail, but it turns out also to be valuable in representing signals where a time-frequency or time-scale description is desired even if no concept of resolution is needed.

A space that is particularly important in signal processing is called $L^2(\Re)$. This is the space of all functions *f(t)* with a well-defined integral of the square of the modulus of the function. The " *L* " signifies a Lebesque integral, the "2" denotes the integral of the square of the modulus of the function, and "$\Re$" states the independent variable of integration is a number over the whole real line. A function g(t) to be a member of that space is denoted: $g \in L^2(\Re)$or simply $g \in L^2$.

In order to develop the wavelet expansion we will need the idea of an expansion set or a basis set. If we start with the vector space of signals, *S*, then if any $f(t) \in S$ can be expressed as

$$f(t) = \sum_k a_k \varphi_k(t) \qquad (1)$$

the set of functions $\varphi_k(t)$ are called an expansion set for the space *S*. If the representation is unique, the set is a basis. Alternatively, one could start with the expansion set or basis set and defines the space *S* as the set of all functions that can be expressed by $f(t) = \sum_k a_k \varphi_k(t)$. This is called the span of the basis set.

In several cases, the signal spaces that we will need are actually the closure of the space spanned by the basis set. That means the space contains not only all signals, which are the limit of these infinite expansions. The closure of space is usually denoted by an over-line [9].

The aim of the MRDWT is to decompose the discrete time signal into basis functions, called the wavelets, to give us a good analytic view of the analyzed signal. The decomposition process is divided into stages, called levels or depths. At each depth, different time and frequency resolution is taken (high frequency resolution means lower time resolution and vise versa). This variable resolution is done by using building blocks, or wavelets, which are derived from an original wavelet, called the mother wavelet. The signal is decomposed by using dilated and shifted versions of the mother wavelet [10].

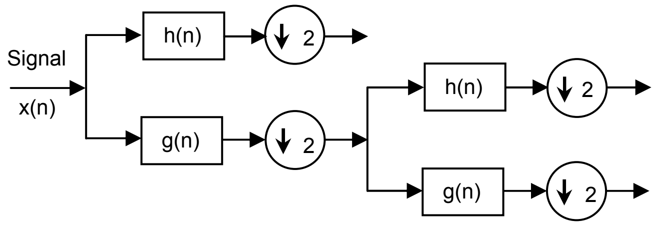
Calculating wavelet coefficients at every possible scale is a fair amount of work, it turns out, that if we choose scales and positions based on powers of two – so – called *dyadic* scales and positions – then our analysis will be more efficient and just such an analysis form the discrete wavelet transform (DWT).

An efficient way to implement this scheme using filters was developed in 1988 by Mallat. This is a practical filtering algorithm yields a *fast wavelet transform* [11].

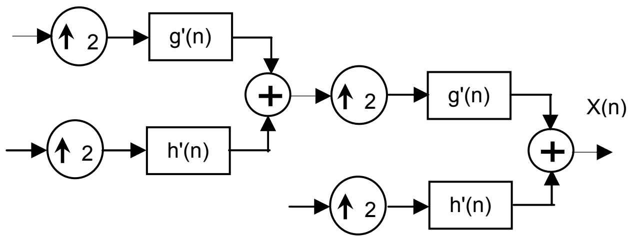
For many signals, the low frequency content is the most important part. It is what gives the signal its identity. For example, consider the human voice. If you remove the high frequency components, the voice sounds different, but you can still tell what's being said. However, if you remove enough of the low – frequency components, you hear vgibberish [12].

It is for this reason that, in wavelet analysis, we often speak of approximations and details. The approximations are the high – scale, low – frequency components of the signal. The details are the low – scale, high – frequency components of the signal. The decomposition process can be made by filtering the signal by LPF and HPF, then the signal is down-sampled by two [13].

The reconstruction process can be made by up-sampling f and *d* in equations (2) and (3), and filtering them with g (n) and h (n) [14]. Fig, (1) shows the wavelet decomposition and reconstruction tree.



(a)Decomposition



(b)Reconstruction

Fig (1): Wavelet Decomposition & Reconstruction For Two Level

$$d^{(j-1)}(n) = \sum_k f^{(j)}(k)\, h(k-2n) \qquad (2)$$

$$f^{(j-1)}(n) = \sum_k f^{(j)}(k)\, g(k-2n) \qquad (3)$$

Where: $f^{(j)}$ = the original signal.

$f^{(j-1)}$ = the approximation.

$d^{(j-1)}$ = the details.

g(n) is the low pass filter (LPF) & h(n) is the high pass filter (HPF) impulse responses.

3- METHOD OF HIDING SYSTEM

The color image hiding in speech system phase includes two main stages: pre-hiding stage and hiding stage, the general block diagram is shown in fig.(2).

3-1 Pre-hiding stage

This stage includes read and sampling for carrier speech signal and read and arrangement for color image. The detail of this stage illustrated in text below :

- Read and sampling for carrier speech signal: This job is implemented by reading WAV file format of carrier speech signal by using one of two ways; the first way is to read speech file signal from allocation in computer, the second way is to record the speech from the microphone , after limiting the recording time and sampling frequency. In this work we used sampling frequency as 44.1KHZ and recording time may be taken any number of seconds.
- Read and Arrangement for Color Image: we read a color image from allocation in computer, then we obtained on a matrix of (*M*N*3*) elements where *M* is the number of rows in image matrix, *N* is the number of columns in matrix of color image, and 3 represents the number of main colors (RGB). The values of matrix elements are in range (0-255), and the range of WAV speech vector are in the range (-1 to 1), therefore we convert the data of image to double data type and divide all images matrix elements on 256 to make the image matrix elements in range of speech signal. The second step after image matrix reading we arranged the image matrix as a vector. This process achieved by making each row after previous until reach to the last row of the first dimension, then repeat the same process to the second and third layer of matrix respectively. In this stage we obtained on the first secret key contains from the image size and information vector length.

3-2 Hiding Stage

The general block diagram of hiding stage is shown in fig. (2).The hiding stage includes wavelet decomposition operation, image information insertion, and wavelet reconstruction.

- Wavelet Decomposing Process: in this process we decompose the input speech signal, that sampled by 44.1KHZ by using MRDWT, the result of decomposition is two signals or two parts, the detail parts and the approximation part. The approximate represents nearly the original speech signal because the most information of the original signal are in low frequencies part (approximation part) while the details part represents the noisy part of the signal therefore the insertion of the image information vector operation must be in the details part to keep the specification of the output hearing speech signal. The mother wavelet function type for each level must be secret to represent the second secret key in hiding system.

- Insertion of Image Information: this process takes the image vector from the pre-hiding stage, then framing the vector to the number of frame *(*N frame, and M elements in each fram*)*, this number entered from the user of system to obtain on the third secret key. Then insert these frames in details parts by replacing *K(i)*N* frame from the detail part of each level by the same number of frames of image information vector, where *K*(i) is the percentage of frame number that required to insert in each level, and i=1,2….L is the level number. The new two keys (*M,N*) and *K(i)* must be secret. The most of insertion frames achieved in the first level of DWT because the most noise exist in the details of this level, and the most of remainder frames insert in second level and so on. In the insertion operation for each level each frame has a starting position to put the first element and then arrange the other elements one after other. These starting positions also must be secret because the position is a one of the keys used in this work.

- Wavelet Reconstruction Process: After complete vector insertion operation we are beginning by wavelet reconstruction process this process was illustrated in section two. The output from this process is the speech signal similar to the original speech signal and we can verify from this by hearing the original and output signal (must be the same).

4- COLOR IMAGE INFORMATION RECOVERY

The second phase is the color image information recovery or show hidden message. This phase system includes: the MRDWT decomposition and information vector extraction stage and image reconstruction from information vector stage. Fig. (3) shows the general block diagram of show hidden massage system.

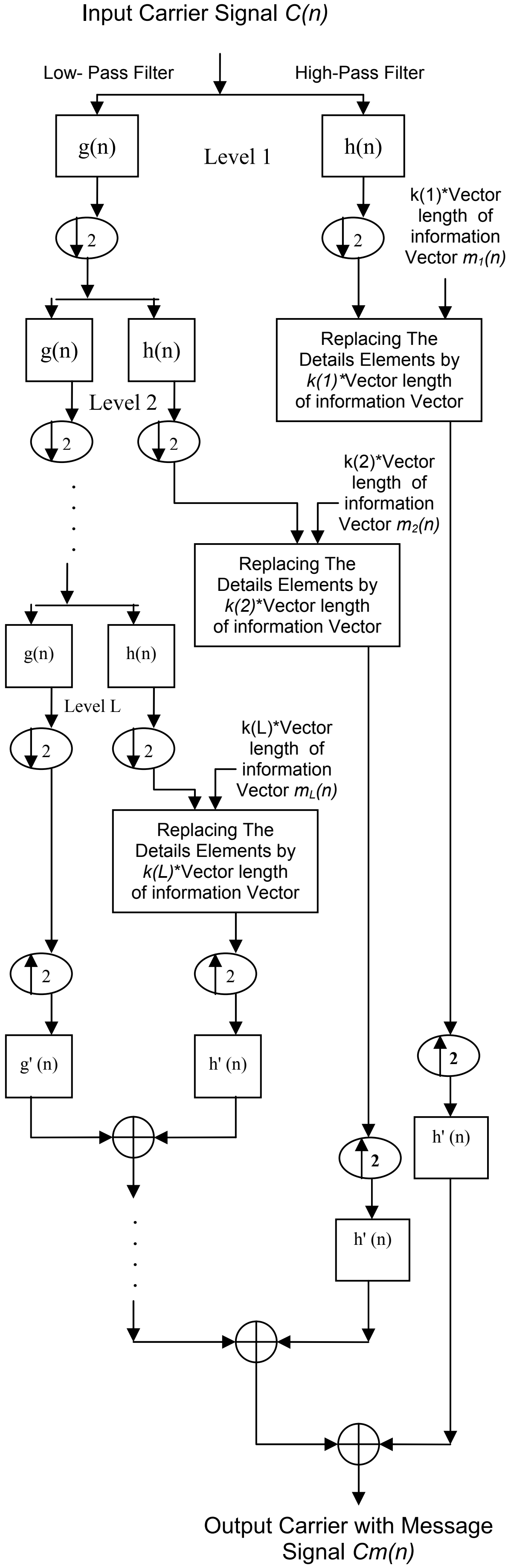


Fig. (2): The General Block Diagram of The Hiding System

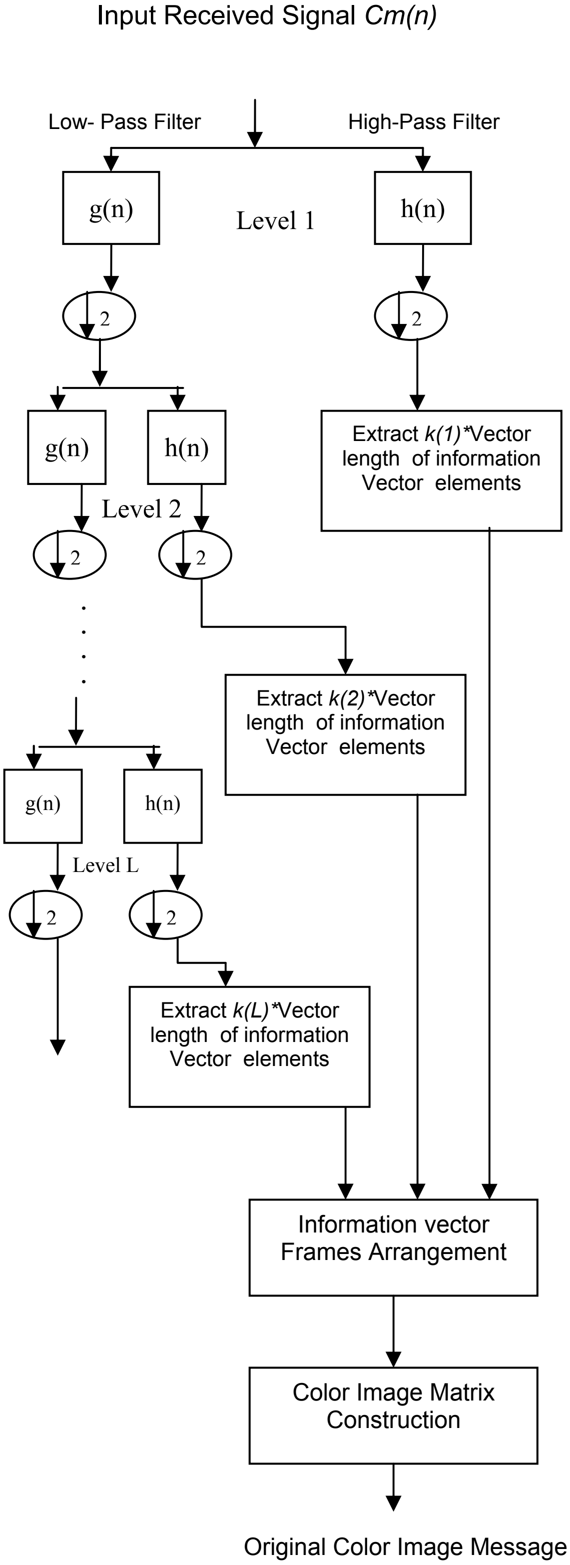


Fig. (3): The General Block Diagram of The Show Hidden System

4-1 MRDWT DECOMPOSITION STAGE

This stage is similar to the DWT decomposition used in hiding stage. The decomposition of *L* level DWT is done in this stage, then extracted *K(i)* frame of image information vector from each level of details parts, where: *i* is the level number. After previous process we arrange the extracting frames in their locations to obtain the final image information vector.

4-2 Color Image Matrix Construction

In this stage reconstruct the color image matrix from information vector by using the image size key and the inverse structure of the information vector production. The result from this stage is the original color image without any recognized distortion in its displayed.

5- SYSTEM SECURITY ANALYSIS

Cachin [15] gave a formal information-theoretic definition of the security of steganography system. The main idea is to refer to the selection of a cover as random variable *C* with probability distribution P_c. The embedding of a secret message can be seen as a function defined in *C*; Let P_s be the probability distribution of $E_k\ (c, m, k)$,that is a set of all a stego-object. Then $P_s(c) = 0$. in order to calculate P_s, probability distributions on *K* and *M* must be imposed. Using the definition of the relative entropy $D(P_1//P_2)$ between two distributions P_1 and P_2 defined on a set *Q:*

$$D(P_1 \backslash\backslash P_2) = P_1(q)\log_2 \frac{P_1(q)}{P_2(q)} \qquad (4)$$

The above equation measures the inefficiency of assuming that the distribution is P_2 where true distribution is P_1 the impact of the embedding process on the distribution P_c can be measured. Specifically, we define the security of a steganography system in term of $D(P_c \backslash\backslash P_s)$. Let *G* be the a steganography system, P_s the probability distribution of the stega-covers sent via the channel, and P_c the probability distribution of *C*. *G* is called ε-secure against passive attackers, if

$$D(P_c \setminus\setminus P_s) \le \varepsilon \qquad (5)$$

And perfectly secure if $\varepsilon = 0$ [16].
Since $D(P_c \setminus\setminus P_s)$ is zero if and only if both probability distribution are equal then we can conclude that a steganography system is (theoretically) perfectly secure, if process of insertion of a secret message in a cover does not alter the probability distribution of *C.* Because of the random speech cover signal the probability of both P_s and P_c is nearly equal and the entropy $D(P_c \setminus\setminus P_s) \to 0$. These are random in cover signal to make the hiding system very robustness against the passive attackers and the speech cover in this work satisfy this condition.

6- RESULTS And Discussions

Both hidden and show hidden systems are tested by using many color images as a message information and many speech signal as a carrier signal. The demonstrated example below illustrate the process of the two systems.

Demonstration Example:

The carrier speech signal had shown in figure (4) with sampling frequency 44.1 KHZ, and the color image shown in figure (5) with size 100*82.

The hiding system phase is beginning by producing image information vector shown in figure (6). Then carrier signal and image information vector are segmented into fifth segments. Each one of them with length shown in table (1). Then each segment from Carrier speech signal is decomposed by using DWT for approximation ; details signals are shown in figure (7). The first column in figure (7) represents the approximation signals where first row in this column represent the approximation coefficients for the first segment and the second row represents the approximation coefficients for the second segment and so on. The second column in figure (7) represents the details coefficients of the first level of DWT for each segment beginning by first which row represents the first segment and ending by fifth which row represents fifth segment. Then the third column in figure (7) represents the details coefficients in the second level of DWT for each segment.

After the above processing insert each one of information vector segment in one speech carrier segment details coefficients. In each carrier speech segment, there are two details signals for two levels of DWT; therefore each image information vector segment will insert in two details coefficients segments with starting elements position and segments arrangements shown in table (1). The elements embedded operation are achieved by replacing some of details coefficients elements through parts of image information vector segment; the value of this part is also shown in table (1). Figure (8) shows the details signals after insertion operation the figures in the first column shows the details signals after insertion operation of the first level starting by first row to represent the first segments and ending by the fifth row, and the second column for the second level. Then the wavelet reconstruction operation is applied on the approximation signal shown in the first column of figure (7) and the result of insertion process signal is shown in first and second columns of figure (8), where each row for these figures are used to produce one output segment, the arranging of all producing segments produces the output speech signal similar to the original carrier signal and contains color image message information vector. Figure (9) shows the final speech signal that is required in transmitting, this signal is similar to the original carrier speech signal after has been tested by hearing from more than one person.

The show hiding system phase segmenting the received signal (hiding system output signal) into five segments (similar to the number of segments used in hiding phase) with length equal to the lengths of segments used in table (1), then two levels wavelet decomposition processing applied to each segments with DWT filters types shown in table (1) to obtain on two details coefficients signals and one approximation coefficient signal for each segment. These signals similar to the coefficients signals that explained in figure (7). From the two levels details coefficients for each segment we can extract the color image information vector segment that embedded in hiding phase in this segment by using the same starting positions and the same lengths of information vector segments that the same in hiding phase. These operations repeated until extracting all image information segments. Finally we arrang the produced segments of information vector in one vector to obtain the same color image information vector that explained in figure (6-a) in hiding phase system. The resulting vector converts to matrix of three dimensions with the same size of original color image message to give in display the same original color image shown in figure (5) with out any distortion.

After complete the demonstrating example explained we must note many of important notes in the system keys and system complexity, and the system processing time in each phase of system.

The keys used in the system are many and complex. The first key in hiding system is the color image size and the information vector length. The second key is the number of information vector segments. The third key is the information vector segments lengths. The fourth key is the speech carrier segments lengths. The fifth key is the number of levels used in DWT and the data type of each level and for each segment of carrier speech signals. The sixth key is the information vector segments insertion in the current carrier speech segment and with any percentage of its length for each level of details coefficients. The seventh key is the starting location of data embedded in each level. All these keys are used in the hiding system and they gave the system high security against the attackers, because the uncertainty that occur from the huge possibilities may occur in estimating of the correct solution. This security which we are talked about it with assumption of the knowledge of attackers in all details of hiding system's algorithm except the keys. This uncertainty make the attacker needs very long time to reach the correct solution.

The complexity of system came from the complexity of the algorithm stages as well as the arbitrary of the chosen speech carrier signals that prevents the attacker has got any pointers about the frequently of using this cover, therefore the attacker depends on only the little frequently covers that make his job difficult.

Another important feature in this system is the short processing time for hiding and show hiding system. The hiding processing time is less than one minute, while the processing time of show hiding system is less 30 second, although the program used is MATLAB 6.5 and processor with speed 1700 MHZ.

Remains one point, it is the capacity of the cover. Because of the high sampling rate used in this system, the number of samples is 44100 samples for each recording second. The increasing of recording time makes the system able to hiding large color images.

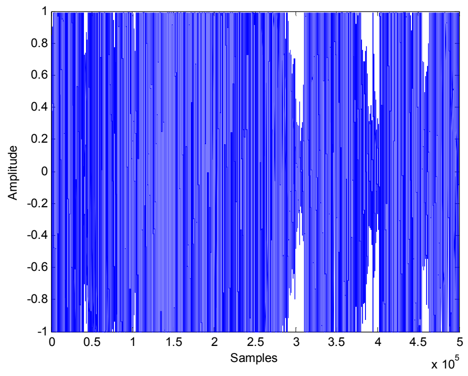


Fig.(4) : The Carrier Speech Signal.

Fig.(5): The Color Image Message that Used In Test

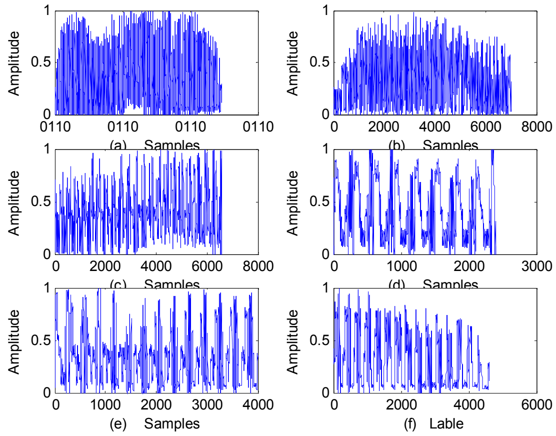


Fig.(6): The Image Information Vector and Theirs Segments. (a) Original Vector, (b) First Segment, (c) second Segment, (d) Third Segment, (e) Fourth Segment, (f) Fifth Segment.

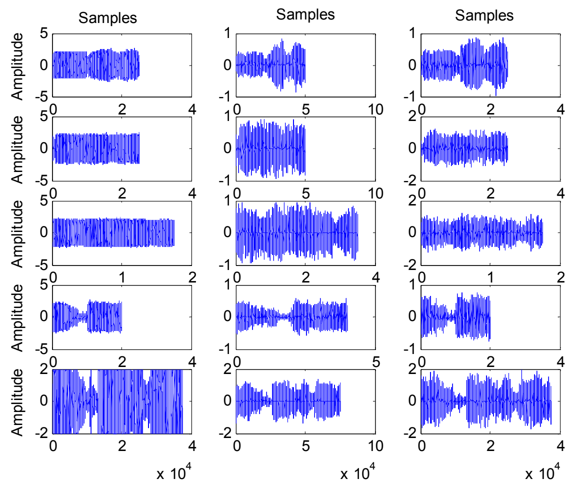


Fig.(7) : The DWT Coefficients for each segment.

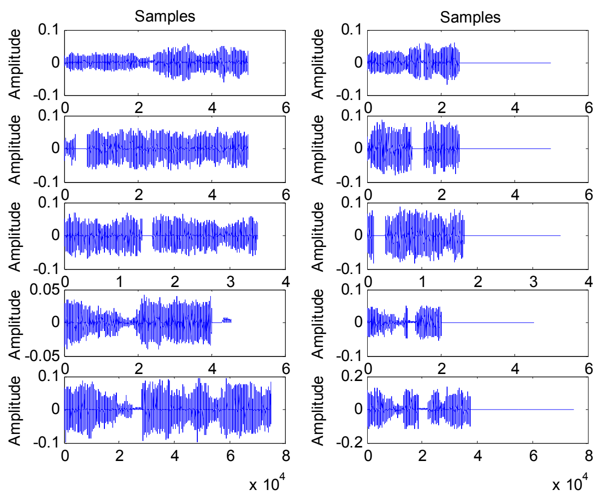


Fig.(8): The Details Coefficients After Insertion Operation

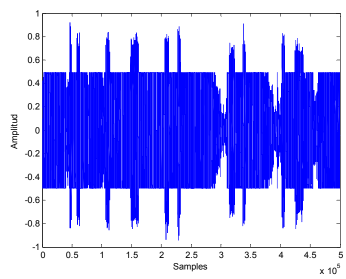


Fig.(9): The Final Signal Produced From The Hiding Phase

Table (1): The Hiding Specification Keys

Carrier Segments name	T_1	T_2	T_3	T_4	T_5
Length Of each Segments	100000	100000	70000	80000	150000
Information Vector Segments Name	V_1	V_2	V_3	V_4	V_5
Length of each Segments	7000	6600	2400	4000	4600
DWT Type for the First Level	Db_8	Db_4	Db_2	Db_{16}	Haar
DWT Type for the Second Level	Db_4	Db_8	Haar	Db_1	Db_{32}
Information Vector Segments Inserted	V_3	V_2	V_4	V_5	V_5
The Percentage of Inserted Segment Length in Level 1	0.5	0.5	0.5	0.5	0.5
The Percentage of Inserted Segment Length in Level 2	0.5	0.5	0.5	0.5	0.5
The Starting Position Of embedded elements For the First Level	23187	3187	13999	3325	24890
The Starting Position Of embedded elements For the Second Level	14500	12140	1287	111111	18765

8- REFERENCES

[1] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, “Information hiding—a survey,” *Proceedings of the IEEE* 87, pp. 1062–1078, July 1999.

[2] N. F. Johnson and S. Jajodia, “Exploring steganography: Seeing the unseen,” *IEEE Computer* 31(2), pp. 26–34, 1998.

[3] R. J. Anderson and F. A. P. Petitcolas, “On the limits of steganography,” *IEEE Journal on Selected Areas in Communications* 16, pp. 474–481, May 1998.

[4] J. Fridrich and M. Goljan, “Practical steganalysis of digital images—state of the art,” *Proceedings of the SPIE Photonics West* 4675, pp. 1–13, 2002.

[5] N. F. Johnson and S. Jajodia, “Steganalysis: The investigation of hidden information,” *Proceedings of the 1998 IEEE Information Technology Conference,* pp. 113–116, 1998.

[6] J. Fridrich, M. Goljan, and D. Hogea, “Steganalysis of JPEG images: Breaking the F5 algorithm,” *5th International Workshop on Information Hiding* , 2002.

[7] S. Lyu and H. Farid, “*Detecting hidden messages using higher-order statistics and support vector machines,”* 5th International Workshop on Information Hiding, 2002.

[9] R. Wilson, "*Multiresolution Image Modelling*", Electronic and Comunication Journal, pp.89-97, April 1997.

[11] S. Mallat , “ *A Theory For Multiresolution Signal Decomposition : The Wavelet Transform*” , IEEE Trans. Pattern Anal. Machine Intell. Vol.11 No.7 ,PP. 674-693,1989.

[13] F. Reyadh. AL-Hashimi, "*Text-Independent Speaker Recognition Using Wavelet And Neural Networks"* M. Sc. Thesis, Collge of Engineering, University of Baghdad, Baghdad, November, 2002.

[14] E. C. Ifeachor, B. W. Jervis, " *Digital Signal Processing*: *a Practical Approach* ", Addison – Wesley, 1996.

[15] C. Cachin, *“An Information-Theoretic Model for Steganography”,* Proceedings of 2nd Workshop on Information Hiding, MIT Laboratory for Computer Science, May 1998.

[16] S. Katzenbeisser, and Fabien A. P. Petitcolas, “*Information Hiding Techniques for Steganography and Digital Watermarking*” Rolf Oppliger, Series Editor, 2000.