Developed Method of Information Hiding in Video AVI File Based on Hybrid Encryption and Steganography

Ashawq T. Hashim*, Dr.Yossra H. Ali**
& Susan S. Ghazoul*

Received on: 7/10/2009
Accepted on: 5/1/2011

Abstract

This paper produces a development of an AVI **H**iding **I**nformation **S**ystem (HIS) based on steganography techniques to prevent intruders to obtain the transmitted information. This work is based on a combination of steganography and cryptography techniques to increase the level of security and to make the system more complex to be defeated by attackers. In this work AVI file is separated into two parts, video and audio. The video is a stream of frames; each frame is stored as a bmp file image and a number of frames required or needed to be used as a cover are chosen.

The algorithm that is used for encryption is the Type-3 Feistel Network of The 128-bits block size improved Blowfish encryption it is a symmetric uses a variable-length up to 129 bytes, making it useful for both domestic and exportable use and a variable-length key would make cryptoanalysis more difficult for potential attackers.

Two methods of hiding are used in this work, the first method is the Least Significant Bit (LSB) and the second is the Haar Wavelet Transform (HWT). The proposed HIS system was tested using standard subjective measures such as Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR). All of the measures obtained as the test results indicate good results for PSNR (above 50db) and they increase when the number of frames used as a cover increases.

الطريقة المطورة لإخفاء المعلومات في الملف الصوتي الصوري المتداخل المعتمدة على التهجين للتشفير والإخفاء

الخلاصة

الهدف من هذا البحث هو تطوير نظام اخفاء معلومات (HIS) في الملفات الصوتية الصورية المتداخلة(AVI) مرتكز على تقنيات الاخفاء لمنع المتطفلين من الحصول على المعلومات المرسلة. هذا البحث مستند على مركب من التشفير و اخفاء المعلومات لزيادة درجة السرية ولجعل النظام اكثر تعقيدا ليهزم من قبل المهاجمين. في هذا البحث, ملف صوتي صوري متداخل من نوع(AVI) , قسم هذا الملف الى جزئين فديو و صوت. الفديو عبارة عن سيل من الهياكل الصورية, كل هيكل يخزن كملف صورة من نوع(BMP) , ثم يتم اختيار عدد الهياكل الصورية المطلوبة لغرض استخدامها كغطاء.

الخوارزمية المستخدمة للتشفير هي Type-3 Feistel Network of The 128-bits). block size improved Blowfish encryption) وهي خوارزمية تشفيرمتماثلة تستخدم مفتاح متغير الى129 bytes, يجعل منها خوارزمية مفيدة لكل من الاستخدام المحلي والدولي والطول المتغير للمفتاح يجعل تحليل الشفرة اكثر صعوبة للمهاجمين.

هنالك طريقتين استخدمت للاخفاء في هذا البحث, الطريقة الاولى هي الثنائيات الاقل اهمية(LSB) والطريقة الثانية هي نظام التحويل المويجي (Haar Wavelet Transform).النظام(HIS) المقترح تم

***Control and Systems Engineering Department, University of Technology/Baghdad**
****Computer Sciences Department , University of Technology/Baghdad**

اختباره باستخدام القياسات المعتمدة مثل(MSE)و((PSNR). كل القياسات الناتجة من الاختبارات تشير الى نتائج جيدة (اكثر من ٥٠ ديسي بيل) و هذه القيم تزداد عند زيادة عدد الهياكل المستخدمة كغطاء.

Introduction

Digital communication has become an essential part of infrastructure nowadays, a lot of applications are Internet-based and in some cases it is desired that communication be made secret. Consequently, the security of information has become a fundamental issue, many of the techniques are available to achieve this goal some of them are the Encryption and the steganography techniques. Using cryptography, the information is transformed into some other gibberish form and then the encrypted information is transmitted. Steganography is a process that involves hiding a message in an appropriate carrier for example an image or an audio file. The carrier can then be sent to a receiver without anyone else knowing that it contains a hidden message [1].

Steganography is mostly used on computers with digital data being the carriers and networks being the high speed delivery channels. Modern steganography is generally understood to deal with electronic media rather than physical objects and texts. This makes sense for a number of reasons. First of all, because the size of the information is generally (necessarily) quite small compared to the size of the data in which it must be hidden (the *cover file*), electronic media are 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. Electronic data also often includes redundant, unnecessary, and unnoticed data spaces which can be manipulated in order to hide messages [1].

AVI File

In general, AVI files contain multiple streams of different types of data. Most AVI sequences will use both audio and video streams a standard package to allow its simultaneous playback. A simple variation for an AVI sequence uses video data and does not require an audio stream. Specialized AVI sequences might include a control track or MIDI track as an additional data stream. The control track could control external devices such as an MCI videodisc player. The MIDI track could play background music for the sequence [2].

Framework of Steganography Model

In general, the basic framework of the steganography model is illustrated in Figure (1).

This model consists of two main processes, namely the *embedding process* and the *extracting process*. The main function of the embedding process is to hide the secret message, called *embedded message*, in a given cover, called *cover-file*. In hidden communication techniques, the cover-file is no more than an innocent (unrelated to the embedded message) piece of information that is used to hide the secret information. A secret key, called *stego-key* is used in the embedding process such that it makes

the embedded message computationally infeasible to extract without possessing this key. The output of the embedding process is called *stego-file*, which is the original file holding the hidden secret message. This output becomes, at the other end, the input of the extracting process, in which the embedded message is extracted from the stego-file to complete the hidden communication process. Since the stego-key is used in the embedding process, it needs to be used in the extracting process [3].

Type-3 Feistel Network of The 128-bits block size improved Blowfish encryption algorithm:

The algorithm takes four 32-bit plaintext data words A, B, C, D as input and produces four 32-bit ciphertext data words A, B, C, and D. The cipher is word-oriented, in that all the internal operations are performed on 32-bit words. This algorithm is a type-3 Feistel network iterated simple function 16 times (Fig. 2) [6].

This cipher uses a variety of operations to provide a combination of high security, high speed, and implementation flexibility. It uses also four key dependent (S-box) tables of 255 32-bit words to provide good resistance against linear and differential attacks, as well as good avalanche of data and key bits [6].

In each round the output of F-function is the input to E-function then, one data word will be used as the input to the E-function and the three output words from the E-function are added or XORed to the other three data words. In addition, the source word is rotated by 13 positions to the left. The algorithm uses the same structure of F-function of previous Blowfish algorithm [6].

The E-function:

The E-function takes as input one data word and uses two more key words to produce three output words. In this function three temporary variables will be used, denoted below by L, M and R a shown in figure (3) [6].

The Overall System Model

The overall system model can be described as shown in Figure (4).

Figures from (5 to 10) show the flowcharts of the whole proposed system.

Embedding the encrypted message into the video stream is done in two methods:

1. Least Significant Bit Embedding

It is one of the basic and easily implemented image steganography methods. This is done by embedding one bit from the encrypted data into one pixel of the cover; the given bit embeds in the Least Significant Bit of the blue byte of this pixel [7].

2-Haar Wavelet Transform Embedding Method: The frequency domain transform applied in this thesis is the Haar-Discrete Wavelet Transform probably the simplest and best known of the wavelet transforms. It consists of a series of averaging and difference steps. The operation can be divided into two steps: one is the horizontal operation and the other is the vertical one. Haar Discrete Wavelet Transform procedures are described as follows:

Step 1: At first, scan the pixel from left to right in horizontal direction.

Then, do the addition and subtraction operation on neighboring pixels and then store the sum on the left and the difference on the right. Repeat the operation until all rows are processed. The pixel sums represent the low frequency part of the original image and denoted as symbol L. The pixel

differences represent the high frequency part of the original image and denoted as symbol H.

Step 2: Secondly, scan the pixels from top to bottom in vertical direction. Then, do the addition and subtraction operation on neighboring pixels and store the sum on the top and the difference on the bottom. Repeat this operation until all the columns are processed. Finally 4 sub-bands denoted as LL, HL, LH and HH respectively are created. The LL sub-band is the low frequency part and looks very similar to the original image [7].

Experimental Results and System Evaluation

Since the essential goal of steganography is the concealing of the fact that a secret message is transmitted, then it is very important to make the stego-video to be as close as to the cover-video. In fact, imperceptibility of the stego AVI reflects how much it is affected due to embedding process, in other words, imperceptibility can be decided by measuring that effect. In the proposed system, the MSE and PSNR measurements are adopted.

Mean Squared Error (MSE)

The (weighted) mean squared error between the cover image and the stego-image can be used as one of the measures to assess the relative perceptibility of the embedded message.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\|C(i,j) - S(i,j)\|^2 \quad (1)$$

where *m* and *n* are the number of rows and number of columns respectively of the cover image, C (*i, j*) is the pixel value from the cover image, S (*i, j*) is the pixel value from the stego-image [6].

Peak Signal to Noise Ratio (PSNR)

The phrase **Peak Signal-to-Noise Ratio**, often abbreviated into **PSNR**, is an engineering term for the ratio between the maximum possible power of a *signal* and the power of corrupting *noise* that affects the fidelity of its representation. PSNR is most commonly used as a measure of quality of reconstruction in image compression. It is most easily defined via the **MSE** [8].

The PSNR is defined as:

$$PSNR = 10\cdot\log_{10}\left(\frac{MAX_I^2}{MSE}\right) = 20\cdot\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right) \quad (2)$$

Here, MAX_I is the maximum pixel value of the image. When the pixels are represented using 8 bits per sample, this is **255**. More generally, when samples are represented using linear **PCM** with *B* bits per sample, MAX_I is 2^B-1.

For color image with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three. The larger PSNR dB value is the higher the image quality is (which means there is only little difference between the cover-image and the stego-image). On the contrary, a small dB value of PSNR means the great distortion between the cover-image and the stego-image [8].

The Similarity Test

The similarity test is the correlation between the cover-image and stego-image. Correlation is one of the best known methods that evaluate the degree of closeness between two functions. This measure can be used to determine the extent to which the

original image and the stego-image are close to each other, even after embedding data.When the stego-image is perceptually similar to the original cover-image; then the correlations equals one [8].

Pearson Correlation Coefficient (Corr) is given by;

$$Corr = \frac{\sum\sum (S - \bar{S})(C - \bar{C})}{\sqrt{\sum\sum (S - \bar{S})^2 \sum\sum (C - \bar{C})}} \quad (3)$$

where $\bar{S} = \frac{\sum\sum S}{MN}$.

S: stego-image. C: cover-image.

Experimental Tests

The tests are done on two types of hiding methods using three types of secret files (text, image and audio).

The two hiding methods were used to hide, they are (Least Significant Bit and Haar Wavelet Transform), in LSB hiding method the embedding is done in the least position, while in HWT hiding method embedding that is used to hide in video stream is done in the LL sub-band of the wavelet coefficients, while in audio stream one level of HWT coefficients is used, the message data will be hidden in the LSB of these sample coefficients that hide limited block of data in limited bytes.

The proposed system is implemented on nine message samples (3- pictures, 2- texts and 2-audio files); these samples differ in size and in format.

Implement on the (GLOBE. avi) file as shown in figure (11).

a) In video stream part:

The size of the video stream is 3.46 MB. It consists of 107 frames each has the same format (BMP file) and the size of each frame is (320 X 240) pixels.

The experimental result examples will be shown in table 1.

b) In audio stream part:

The format of the audio part is WAVE file and the size is 3.58MB.

Audio sample rate=44KHZ. Number of channels=2(stero).

Audio sample size=16 bits.

The experimental result examples will be shown in table 2.

Robustness Test

A simple equation that may be used to check the system robustness and the performance of the system is by computing the difference between the extracted messages without any attack or modification and between extracted messages after some attacks and signal processing.

$$\text{Error diff} = \frac{\text{Extracted message without attack} - \text{Extracted message after attack}}{\text{Message size}} * 100\% \quad (4)$$

It is important to notice that in order for an attack to be effective, it has to eliminate the message without visibly altering the carrier file.

Some types of these attacks are taken form [9]. Table (3) contains some kinds of attacks and their effects.

Results Discussion

1- The video part is considered as a stream of images, so the method of hiding information used in this project is to hide in these image frames.
2- When the size of message increases the number of frames used as a cover increases. From the studying of the whole system, it's clear that when the secret file size increases MSE increases and the PSNR decreases. This result is obtained by applying different secret message

sizes. But even with a minimum number of frames the MSE is reasonably small.

3- The time of execution depends on the following parameters:
a-The method of hiding: when LSB method is used, the time of execution is smaller compared with the HWT method used because the time wasted through decompose and reconstruction operations done for each cover frame.
b-The place of hiding: Hiding into audio stream using LSB method takes longer time compared with hiding the same message into videostream using the same method, this is because the format and the construction of the WAVE file.
c-The size of the secret message: When the message size increases the time of execution is increased too.

4- Hence a hacker must know the following in order t extract the embedded message from the stego-file:
a. Algorithm to extract the message from the image. (stego algorithm)
b. Encryption algorithm.
c. Correct password for algorithm.
With these increased levels of protection using encryption algorithm, the proposed system for steganography is stronger to attacks than any other existing system that does not use encryption.

5- The results obtained from the correlation test indicate that the stego-file is similar to its corresponding cover since correlation values approach one. This proves that the system is secure.

6- The hiding information in audio stream limits the amount of data that could be hidden, since the audio file size is generally not too big.

7- With LSB method there is more space to hide since the entire frame pixels used to hide while in HWT method is only one sub-band of wavelet coefficients used for hiding so hiding data rate using LSB is bigger than when using HWT method.

8- As seen in result in Table (3), when noise coefficient exceeds 0.005, the message cannot be extracted completely without error. One can observe that hiding using HWT method is more robust than hiding using LSB method.

9- Because of the simplicity of implementing LSB hiding method, the time of execution of this method is considered smaller than HWT method.

Conclusions

The proposed system offers high embedding data rate and a high level of security in term of transmitting the resultant stego-AVI file without rising any suspicious. After studying the whole system, one can conclude the following:

1- No technique of information hiding can ensure perfect secrecy; however, by combining steganography with other techniques, such as cryptography, a higher chance of success can be achieved.
2- The output stego-file remains of the same size as the original file. It is also rarely affected after hiding the information according to subjective measures (seeing and hearing) or the objective measures (MSE and PSNR) as it appears in the results given in experimental result. These results prove that the goal of steganography is achieved where the stego-file will not look suspicious and nobody even when one knows that there is a hidden message.
3- The embedded message is extracted directly from the transmitted AVI file, so there is no need for the original cover file in the extraction process.
4- From the experimental result, the proposed hiding system based on Haar Wavelet Transform is more robust to channel noise and attacks than LSB insertion method.
5- There is a trade off between the amount of the information that would to be hide and the robustness of the system, it is impossible to obtain both. If the information to be hidden has big size then this will reduce the robustness of the system.
6- Even if the AVI cover file is small, it is obvious that the video stream consists of a large number of frames so one can hide a reasonable large amount of information by separating it on the video and audio parts.

References

[1] T. Morkel, J.H.P. Eloff and M.S. Olivier, "**An Overview of Image Steganography**", in Proceedings of the Fifth Annual Information Security South Africa Conference (ISSA2005), Sandton, South Africa, June/July 2005.

[2] M. Owens, "**A Discussion of Covert Channels and Steganography**", as part of the Information Security Reading Room. SANS Institute 2002.

[3] A. H. Ouda and M. R. El-Sakka, "**A Step Towards Practical Steganography Systems**", Computer Science Department, University of Western Ontario, London, Ontario, Canada, ICIAR 2005, LNCS 3656, pp. 1158 – 1166, 2005.

[4] A. Setiawan, D. Adiutama, J. Liman, A. Luther and R. Buyya," **Grid Crypt: High Performance Symmetric Key Cryptography using Enterprise Grids**", Grid Computing and Distributed Systems Laboratory**,** Dept. of Computer Science and Software Engineering**,** The University of Melbourne, Australia,2004.

[5] N. Sharma, J. S. Bhatia, N. Gupta, "**An Encrypto-Stego Technique Based Secure Data Transmission System**",The Infosec Writers Text Library (RSS) , 2005.

[6] Ashwaq T. Hashim "**Type-3 Feistel Network of The 128-bits Block Size Improved Blowfish Cryptographic Encryption**", Eng & Tech, Journal, Vol. 27, No. 2, 2009.

[7] Susan S.Ghazoul, **"Development of Information Hiding System Based on AVI Format"**, M.Sc. thesis**,** Control and Systems Eng. Dept, University of Technology, April 2007.

[8] Venkatraman. S, A. Abraham and M. Paprzycki, "**Significance of Steganography on Data Security**", Dept. of Computer Science & Engineering, University of Madras, INDIA, Dept. of Computer Science, Oklahoma State University, USA, Proceedings of the International Conference on Information Technology, 2004IEEE.

[9] X F Ma and T Jiang, "**The Research on Wavelet Audio Watermark Based on Independent Component Analysis**", International Symposium on Instrumentation Science and Technology. Journal of Physics: Conference Series 48(2006) 442-446.

Table (1) Result for hiding message samples in Globe.avi (video stream)

Secret file	Method Type	Number of frames used	Cover-stego_cover			Execution time(sec)	
			MSE	PSNR	Corr	Hiding	Extraction
PIC1 13.3kb	LSB	2	0.295	53.432	0.99999	11.366	2.563
	HWT	6	0.467	51.437	0.99998	9.038	5.574
PIC2 90KB	LSB	10	0.471	51.400	0.99999	9.759	4.927
	HWT	38	0.497	51.167	0.99999	57.874	29.797
PIC3 230.4KB	LSB	25	0.503	51.115	0.99998	14.875	6.875
	HWT	97	0.522	50.954	0. 99998	147.72	79.921
TEXT1 12.6KB	LSB	2	0.330	52.945	0.99999	1.216	3.06
	HWT	6	0.422	51.877	0.99999	10.891	6.278
TEXT2 25.6KB	LSB	3	0.450	51.598	0.99999	1.938	1.97
	HWT	11	0.482	51.300	0.99999	19.367	10.967
Audio1 65KB	LSB	7	0.491	51.219	0.99999	4.649	3.73
	HWT	28	0.494	51.193	0.99998	5.288	24.982
Audio2 131 KB	LSB	14	0.488	51.246	0.99999	9.215	5
	HWT	55	0.496	51.175	0.99998	90.823	47.527

Table (2) Result for hiding message samples in Globe.avi (audio stream)

Secret file	Method used	Cover-stego_cover			Execution time(sec)	
		MSE	PSNR	Correlation	Hiding	Extraction
PIC1 13.3KB	LSB	0.0289	63.521	0.99966	22.687	3.563
	HWT	0.118	57.411	0.99587	29.782	3.350
PIC2 90KB	LSB	0.192	55.297	0.99799	65.136	16.219
	HWT	The cover size is not enough				
PIC3 329KB	LSB	The cover size is not enough				
	HWT	The cover size is not enough				
TEXT1 3.08KB	LSB	0.005	70.349	0.99993	18.975	1.960
	HWT	0.03	63.359	0.099588	21.185	1.677
TEXT2 10.4KB	LSB	0.022	64.706	0.99974	23.720	3.107
	HWT	0.095	58.353	0.99582	22.715	3.029
TEXT3 20KB	LSB	0.042	61.898	0.99952	30.549	4.626
	HWT	0.174	55.725	0.99560	30.245	4.636
AUDIO1 65KB	LSB	0.142	56.607	0.99847	53.248	12.265
	HWT	The cover size is not enough				
AUDIO2 71KB	LSB	0.154	56.255	0.99833	53.587	13.158
	HWT	The cover size is not enough				
AUDIO3 80KB	LSB	0.171	55.800	0.99825	67.103	16.230
	HWT	The cover size is not enough				

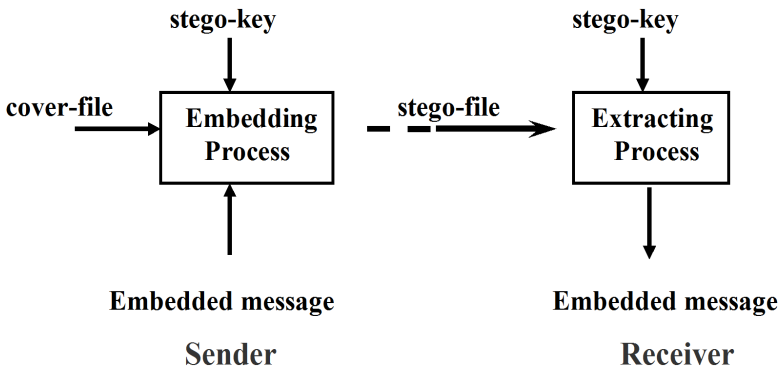


Figure (1) Framework of the embedding process

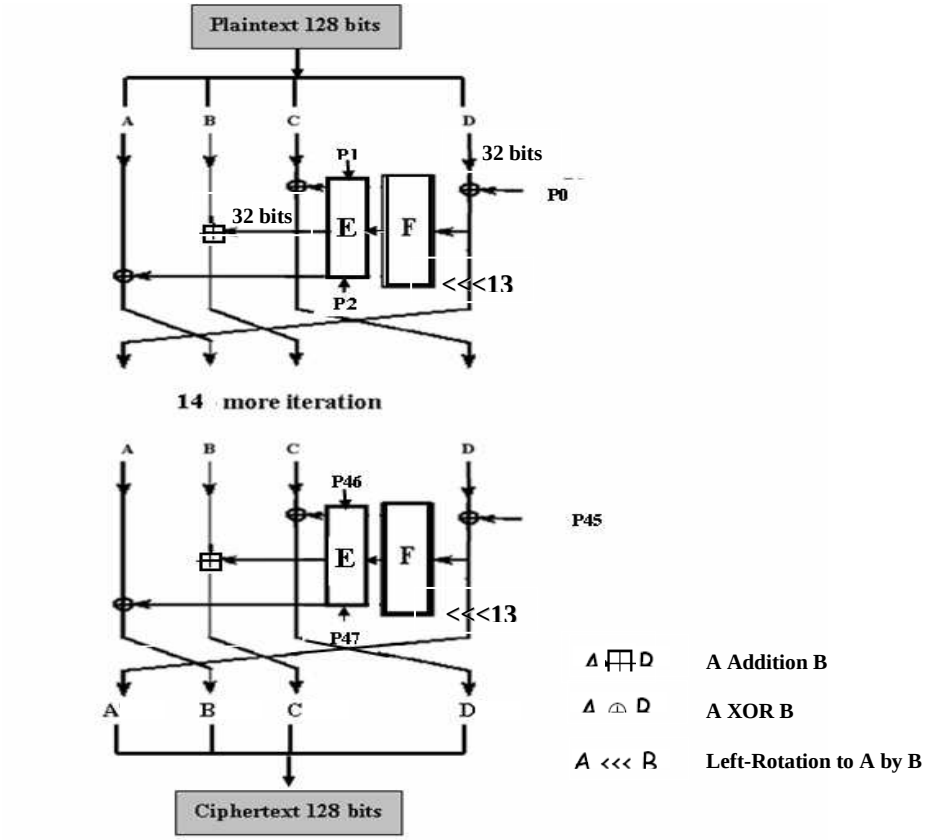


Figure (2) Type-3 Feistel Network of The 128-bits block sizeImproved Blowfish Encryption Algorithm

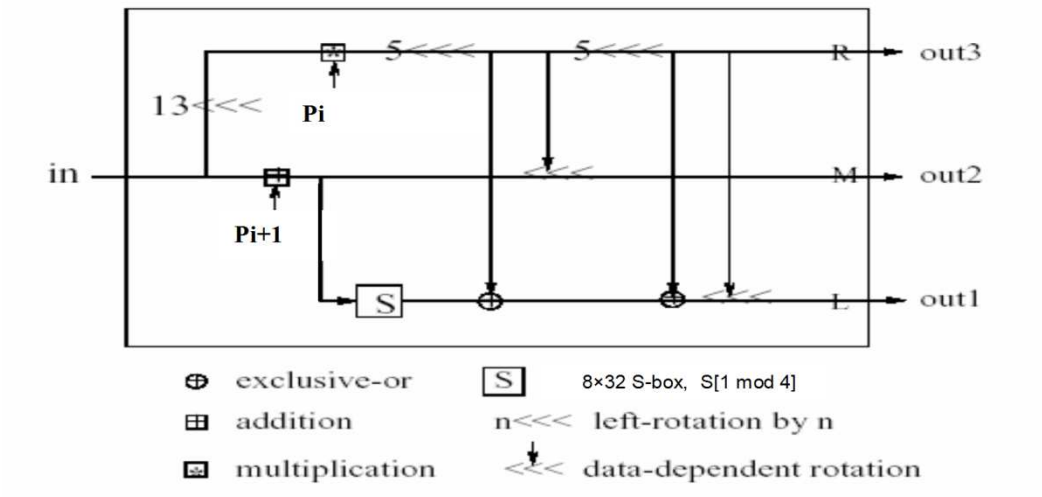


Figure (3) The E -function

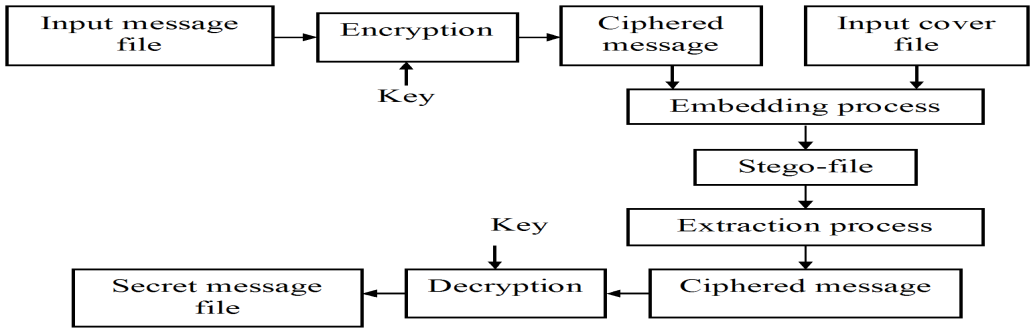


Figure (4) Overall System Model Scheme

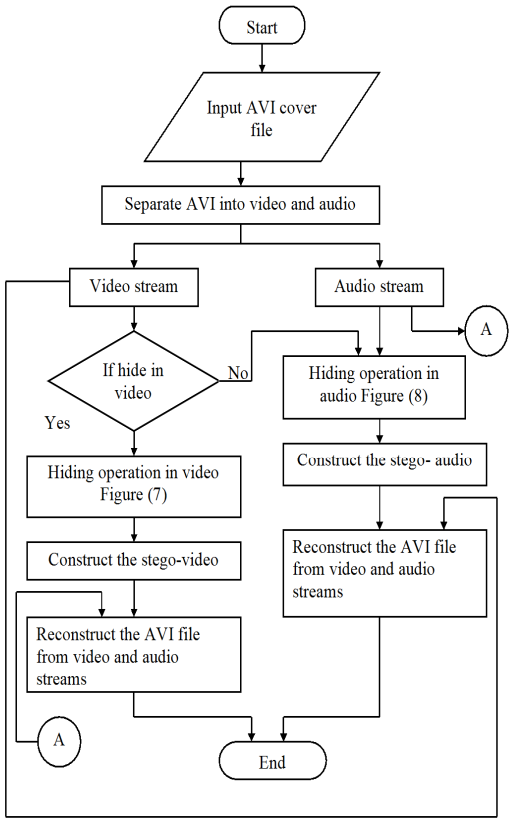


Figure (6) Flowchart of the proposed system (hiding operation)

Figure (5) flowchart of the proposed system (hiding operation)

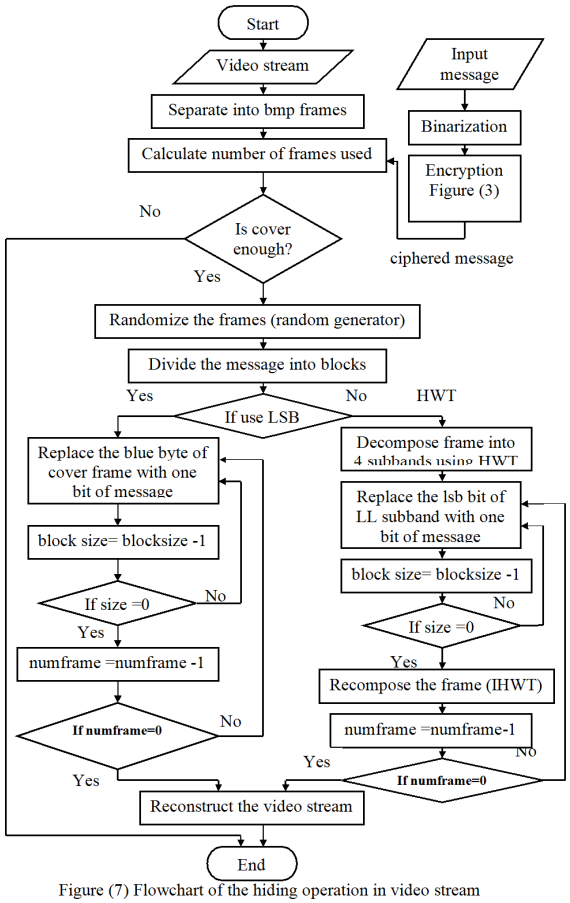


Figure (7) Flowchart of the hiding operation in video stream

Figure (6) flowchart of the hiding operation in video stream

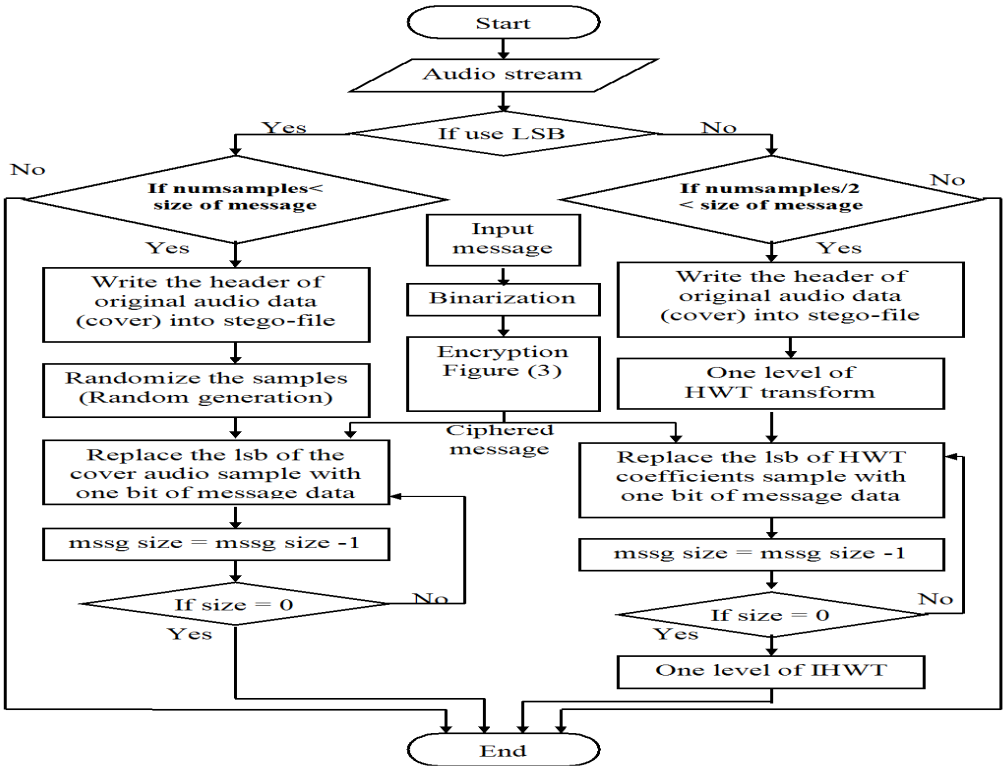


Figure (7) flowchart of the hiding operation in audio stream

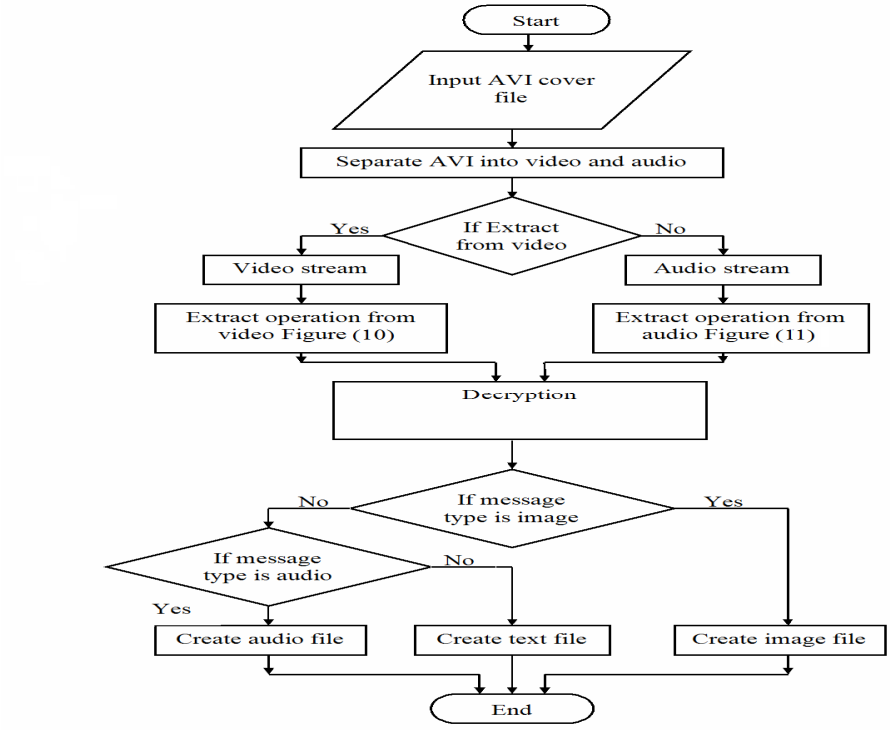


Figure (8) Flowchart of the proposed system (Extraction Process)

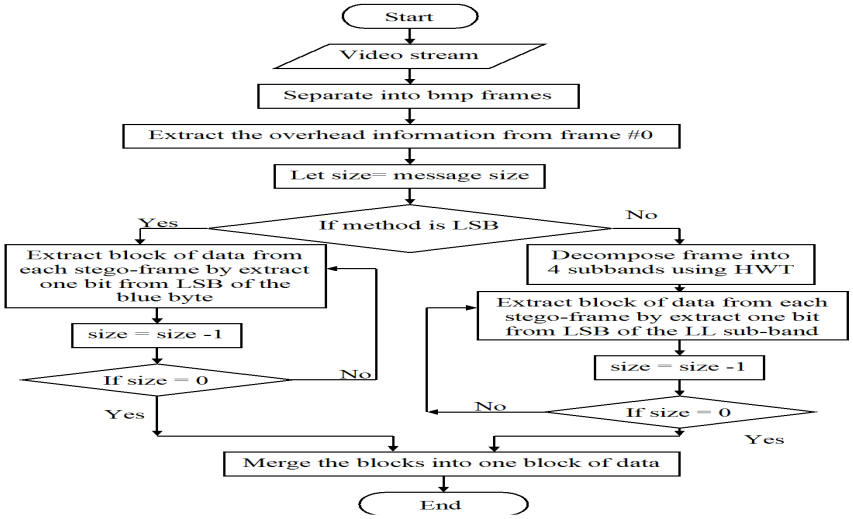


Figure (9) Flowchart of the Extraction process from video stream

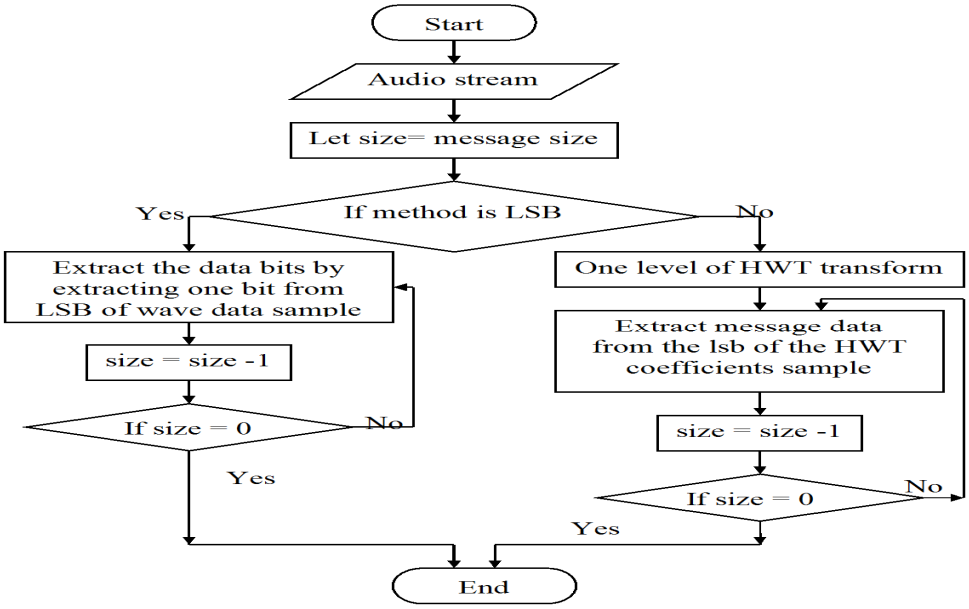
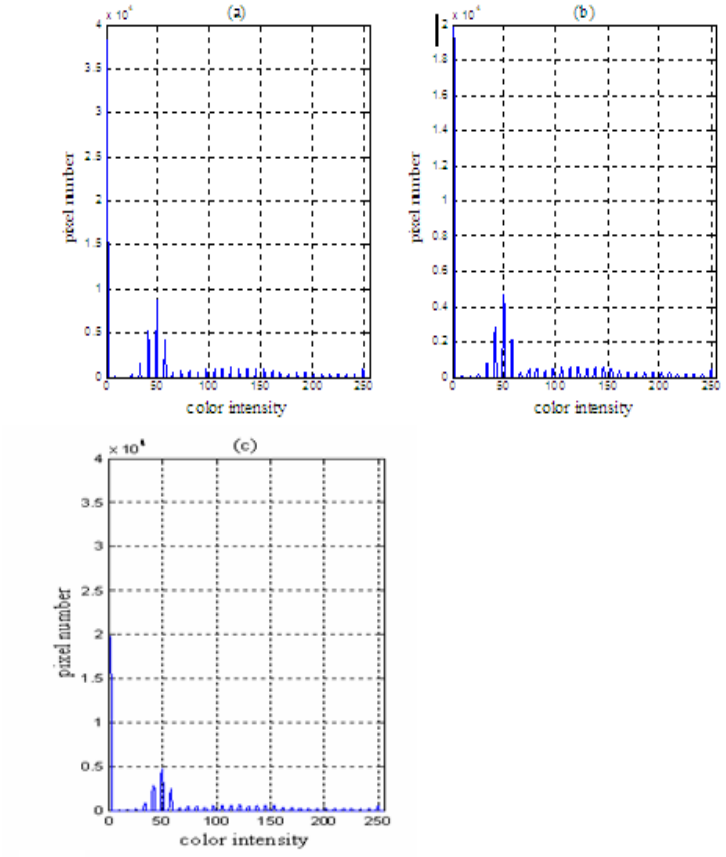


Figure (10) Flowchart of the extraction process from audio stream

Figure (11) GLOBE. Avi video stream

*



Figur[12] a) Blue channel histogram of original frame#5 of GLOBE.avi file. b) Blue channel histogram of stego-frame#5 after embedding pic3 using LSB. c) Blue channel histogram of stego-frame#5 after embedding pic3 using HWT.

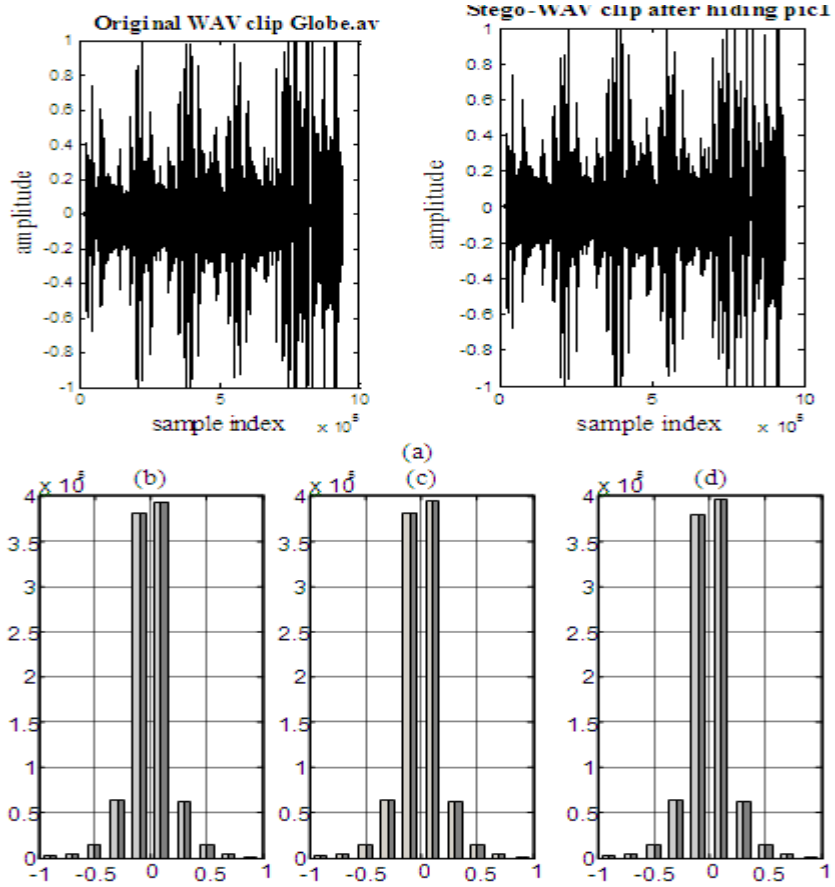


Figure 13 a) Comparison of original audio clip Globe.avi and stego after hiding pic1 in LSB method stream in time domain. b) Histogram of original clip Globe.avi. c) Histogram of stego-WAVE of Globe.avi after embedding pic1 using LSB. d) Histogram of stego-WAVE of Globe.avi after embedding pic1 using HWT.