Military Technical College
Kobry El-Kobba,
Cairo, Egypt

11-th International Conference on Aerospace Sciences & Aviation Technology

SECURE STORAGE FOR VOICE, IMAGE AND TEXT DATA USING STEGANOGRAPHY PARADIGM

Mahmoud E. Gadallah* and Abbass. S. Abbass**

ABSTRACT

In this paper, an algorithm for data hiding is introduced. The purpose of this work is to secure the storage of information such as voice, images and text. Also, this technique has been tested to hide a mixture of text with images. The proposed algorithm is based on using the wavelet transform with post processing to increase the complexity of the hiding. The hiding algorithm reported in this paper has shown promising results from the point of view of the embedding capacity and the quality of the cover data (speech or images) as well as the quality of the recovered data.

KEY WORDS:

Data storage - Secure Communications – Wavelet Transform – Embedding – Cover Image – Secret Image – Encryption – Cryptography – Steganography – Stego image.

I. INTRODUCTION

With the advance of the information technology and its applications in the different fields, data securing becomes an essential issue that must be considered when the data is either stored or transferred via communication media. It becomes very difficult to ensure data protection by just preventing its transfer using storage media because of the appearance of the very large storage capacities with very small sizes. Thus, the best way to overcome this problem is by changing the nature of the data to become unreadable (using encryption) or readable but with different meaning (using hiding or embedding).

Encryption uses powerful mathematics to map plaintext into an unreadable cipher text that can be stored or sent over a channel to the recipient. When a message is

* Prof. Dr., Armed Forces.
** Department of Computer Science, Modern Academy in Maadi.

encrypted, a secret key is used. To decrypt a message, the secret key is used to reverse the process. For an ease dropper to defeat the system, he or she must acquire the secret key. Typically it is assumed that this must be done be searching over the entire key space, a so called "brute force" attack. As this is a very time consuming endeavor, the encrypted message is considered safe.

The second method for secure data storage or communication is called Steganography which offers data protection by hiding the important data into another data that is called cover. The practice of Steganography cannot be seen as a replacement to Cryptography, indeed many of the steganographic software packages, Cryptography is used in addition to Steganography. Techniques for information hiding have become increasingly more sophisticated and widespread [1]. **Least significant bit substitution (LSB)** is the most common form of data embedding. Digital watermarks technology began humbly around 1993 with the exploration of this technique [2]. LSB works by breaking the covert message into individual bits and replacing the LSBs of the pixels are altered. Audio algorithms change the LSBs of samples [3]. **Echo Embedment** is another approach for hiding of information in a discrete signal by introducing an echo in the cover [4]. In [5], some audio data embedding approaches are presented. Using **a phase-coding approach**, data are embedded by modifying the phase values of Fourier transform coefficients of audio segments. **Another one** is based on replacing the Fourier transform coefficients over the middle frequency bands, 2.4-6.4 kHz, with spectral components from hidden data. The middle frequency band was selected so that the data remain outside of the more sensitive low-frequency range.

One of the well-known data embedding methodology is **spread spectrum.** This type attempts to insert a signal throughout the spectrum of a broadband noise carrier [2]. There exist a set of techniques that work in the frequency domain which are based on the spread spectrum concepts. One of such techniques is a technique that embeds a spread spectrum signal into the Fourier coefficients of an image carrier. Frequency-based spread spectrum methods appear to be more robust than their spatial counterparts [6]. Another set of techniques depends on masking phenomenon ([2], [5], and [6]). **Frequency Masking** is well-known technique that uses the masking as a base for the embedding process [1]. Jonathan Foote and John Adcock proposed a method that is called **Time Base Modulation**. This method is based on subtly and inaudibly compressing or expanding time regions of an audio file. By comparing the altered file with a reference copy, compressed and expanded regions can be detected. This method was used in watermarking [7]. A simple technique uses the **DC level** of the audio signal as a way to embed the bits of information such as product id [8]. Also, this method was used in watermarking.

With the introduction of the **wavelets,** new techniques appeared that depend on it. One of them uses the nature of the wavelet coefficients, by finding the coefficients whose values are below a specified threshold and replacing them with the bits of data to be hidden [9]. Embedding data in a transformed content is not restricted to the obvious transforms that are widely used for compression as Discrete Cosine Transform (DCT), wavelet and fractal transforms [10].

The proposed work is an extension to that reported in [11], and [12] in which authors have introduced methods to secure speech messages by embedding it into another one using the Discrete Wavelet Transform (DWT) In addition to applying the mentioned methodology to images and text data, which yield a generalized framework The wavelet transform is used as a basis for the embedding and detection processes because it facilitates separating the important components of the signal (or image) and utilizing the redundant components to replace it with the data which we want to embed. Three algorithms are proposed depending on the above approach to increase the amount of the embedded data (hiding capacity) with respect to the size of the cover (speech or image) without introducing perceptual degradation into the resultant stego signals or images. The paper is organized as follows: In section (II) the wavelet transform is briefly motioned, then the proposed technique is presented in section (III) . Section (IV) introduces the application of the proposed algorithm and results of these tests. Finally, section (V) gives the conclusions .

II. WAVELET TRANSFORM (WT)

The theory on wavelet transform, which originate as a branch of applied mathematics in the 1980's, was first introduced into the signal processing field by French mathematicians I.Daubechies and S. Mallat. Today, intertwined with multi-resolution and filter bank theory, Wavelet analysis plays an important role in time-frequency analysis [13]. Indeed, in their brief history within the signal-processing field, Wavelets have already proven to be an indispensable in addition to the analyst's collection of tools and continue to enjoy a burgeoning popularity today [14]. A wavelet is a waveform of limited duration with an average value of zero. One-dimensional wavelet analysis decomposes a signal into basis functions, which are shifted and scaled versions of a *mother* wavelet. Wavelet coefficients are generated and are a measure of the similarity between the basis function and signal being analyzed. To scale a wavelet is to compress or extend it along the time axis. A compressed wavelet will produce higher wavelet coefficients when evaluated against high frequency portions of the signal. Therefore, compressed wavelets are said to capture the high frequency events in a signal. A smaller scale factor results in a compressed wavelet because scale and frequency are inversely proportional [15]. There are different types of wavelet transforms, including the Continuous Wavelet Transform (CWT) and the Discrete Wavelet Transform (DWT). The CWT is used for signals that are continuous in time and the DWT is used when a signal is being sampled, such as during digital signal processing or digital image processing. The continuous and discrete wavelet transforms are given in (1) and (2), respectively [16]:

$$(T^{wav}f)(a,b) = |a|^{-1/2} \int f(t)\psi\left(\frac{t-b}{a}\right) dt \qquad (1)$$

$$T^{wav}_{m,n}(f) = a_0^{-m/2} \int f(t)\psi\left(a_0^{-m}t - nb_0\right) dt. \qquad (2)$$

The DWT has a scaling function and a wavelet function associated with it. The scaling function can be implemented using a low pass filter and is used to create the scaling

coefficients that represent the signal approximation. The wavelet function can be implemented as a high pass filter and is used to create the wavelet coefficients that represent the signal details. If the DWT is used by scaling and shifting by powers of two (dyadic), the signal will be well represented and the decomposition will be efficient and easy to compute. In order to apply the DWT to images, combinations of the filters (combinations of the scaling function and the wavelet function) are used first along the rows and then along the columns to produce unique sub bands. The (LL) sub band is produced by low pass filtering along the rows and columns and is commonly referred to as a course approximation of the image because the edges tend to smooth out. The LH sub band is produced by low pass filtering along the rows and high pass filtering along the columns, thus capturing the horizontal edges. The HL sub band is produced by high pass filtering along the rows and low pass filtering along the columns, thus capturing the vertical edges. The HH sub band is produced by high pass filtering along the rows and columns, thus capturing the diagonal edges. The LH and HL sub bands are considered the band pass sub bands and the LH, HL, and HH sub bands together are called the detail sub bands [15].

III. THE PROPOSED DATA HIDING ALGORITHM

The basis of the introduced algorithm is to apply the wavelet transform to decompose the cover (speech or image) into their sub band units. These wavelet coefficients have high energy with the transformed signal's energy concentrated within it and other coefficients that has a low energy with low contribution to the signal energy. In wavelet terms this is corresponding to approximations and details coefficients, respectively. By perfect processing for the details coefficients, a considerable amount of space can be obtained to hide information without a noticeable degradation for this signal or image. Then the resultant signal or image will be a stego version of the original one, which is considered as a cover.

It must be noted that the wavelet coefficients are corresponding to diagonal, horizontal, and vertical details in the case of image or the details of the five-level wavelet transform in the case of speech. The gain of this process is the large capacity which can be used to store the data. For example, if we have a gray level image which is wavelet transformed with only one level of decomposition, then if the secret data is embed into the horizontal, vertical, and diagonal details, so we will have 75 % of the size of coefficients can be used to embed secret data. Figure 1 illustrates the basic idea of the proposed technique.

The algorithm consists of two stage, the first one is the embedding stage where the secret data are embedded into the cover one. The input data can be voice message, an image, a text or a mixture of images and text. The wavelet transform of the cover image is calculated, and using different keys, the secret data can be embedded into the wavelet coefficients of the cover image. Finally, the coefficients of the cover are transformed back to yield the stego image. In the second stage, namely the detection stage, the stego image is transformed to the wavelet domain and the embedded data

are extracted by an inverse operations to those used in the embedding process. It must be mentioned that the detection stage can extract the secret data if and only if it has the same keys that have been used in the embedding stage.

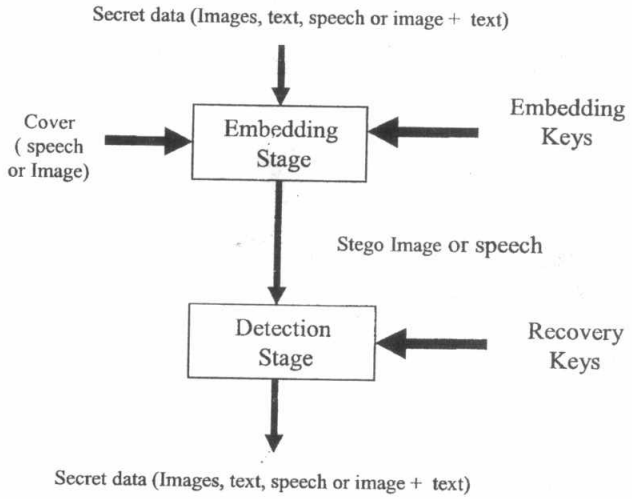

Fig. 1 The block diagram of the proposed algorithm

IV. RESULTS AND DISCUSSIONS

The embedding capacity depends mainly on the levels of decomposition of the WT. In testing this algorithm, it has been found that in the case of embedding a hybrid of images and text into a cover image, the cover image couldn't be decomposed to more than one level. Otherwise, a noticeably degraded stego image is obtained. Thus the hybrid data have been embed into the details of the first level (horizontal, vertical, and diagonal) which means that the size of the embedded data can reach up to 75 % of the size of the covert image. In the case of embedding speech into another speech signals, the algorithm has been tested for different levels of WT decomposition. Up to 5^{th} level have been tested to evaluate subjectively and objectively the degradation of the quality of the covert signal. By embedding speech into the details of the different levels, we reached up to 86.4% of the size of the covert speech signal. Of course we can embed text only in both of the mentioned techniques. It must be noted that the task of embedding a large capacity of the data like text data and gray level image into a single gray level image or embedding a speech data into another speech signal in the wavelet coefficients is not a trivial task, so the wavelet coefficients must be shaped to accept this large capacity.

256 * 256 gray level images were used as covert images, 128*128 gray level images were used as secret images in addition to MS word text data. Figure 2 shows the result

of embedding a hybrid data (2 gray level images in addition to 6-pages MS word text) into this covert image. Figure 3 shows the result of the detection stage where the two embedded images and the embedded text are extracted.

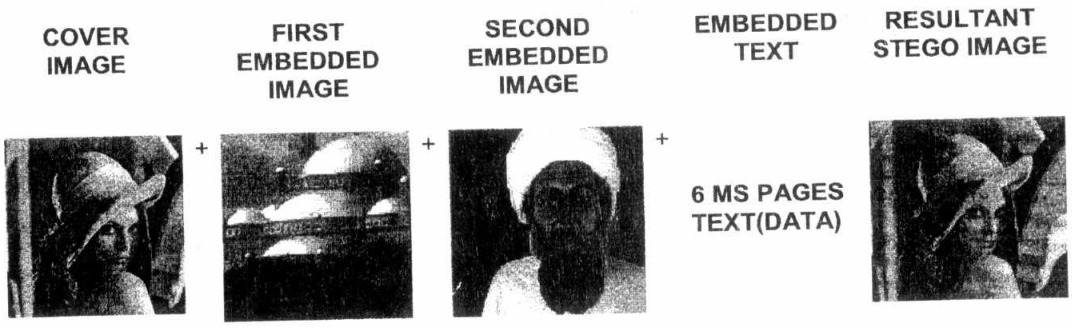

Fig. 2. Embedding 2 images and text data into single image

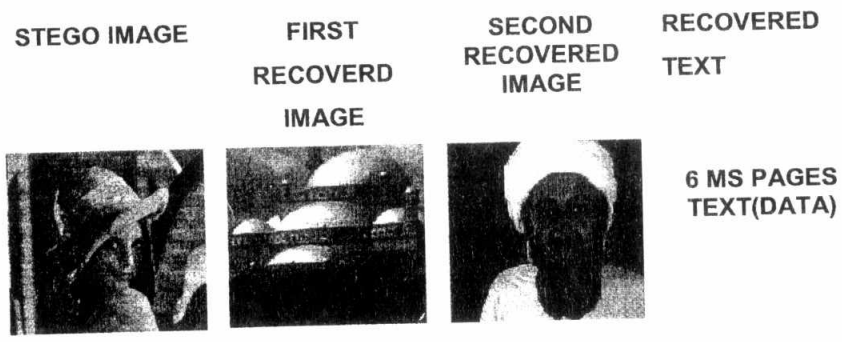

Fig. 3. The extraction of the embedded images and text

Figure 4 and Figure 5 show the result of applying the algorithm to hide a voice message into another covert one.

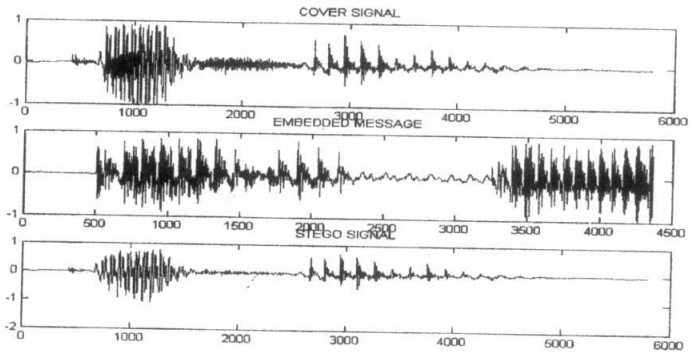

Fig 4 Embedding speech signal into another one

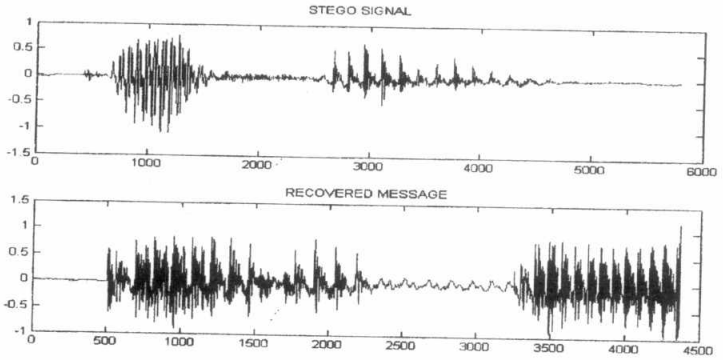

Fig. 5 The extraction of the embedded speech signal

Subjectively, stego image is indistinguishable from the covert one and the stego speech when played back is indistinguishable from the covert speech. Also, there are no noticeable differences between the extracted data and the original secret data (speech, image or text). To ascertain these subjective observations, objective tests are performed to evaluate the performance of the embedding and the detection stages. The proposed algorithm has been evaluated objectively by measuring the root mean square error (RMSE) between the covert data and the stego one as well as between the extracted data and the secret one. The RMSE is defined as:

$$RMSE = \sqrt{\sum\sum \frac{(X-Y)^2}{m*n}} \qquad (3)$$

Where X and Y are the matrices of the two images or the arrays of the two signals and m, n are their dimensions The RMSE has been measured for different embedding capacities. The embedding capacity (EC) is defined as:

$$EC = \frac{ED}{CD} * 100 \quad (4)$$

Where ED is the size of embedded data in bytes
And CD is the size of covert data in bytes

In Figure 6, the RMSE between the covert image and the stego one is plotted against the ED. Figure 7 and 8 show the RMSE of the embedded and extracted images. It must be noted that the error is in the order of 10^{-11}. Finally, in Figure 9, the RMSE between the extracted text and the embedded one is shown.

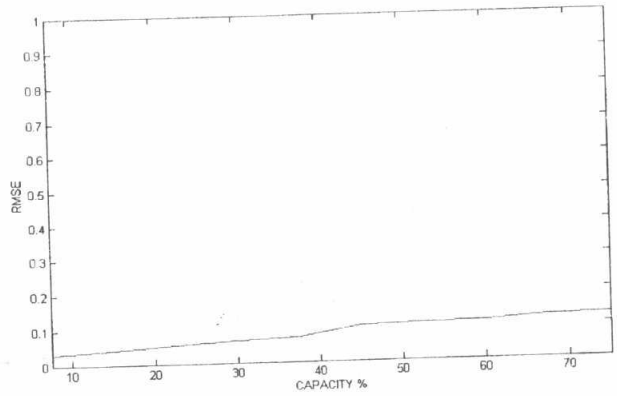

Fig.6. The RMSE between the covert and stego images

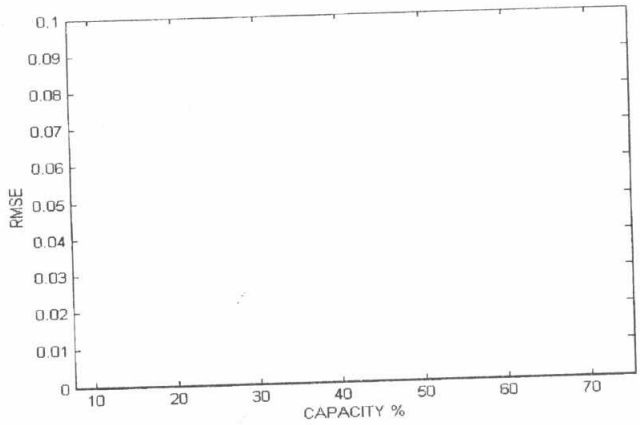

Fig. 7. The RMSE between the embedded and extracted images (the first image)

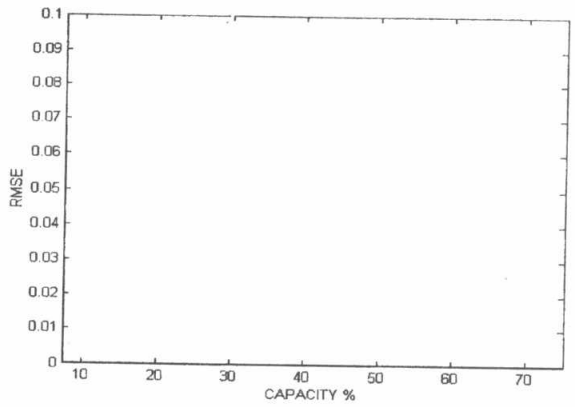

Fig. 8. The RMSE curve between the embedded and extracted image (the second image)

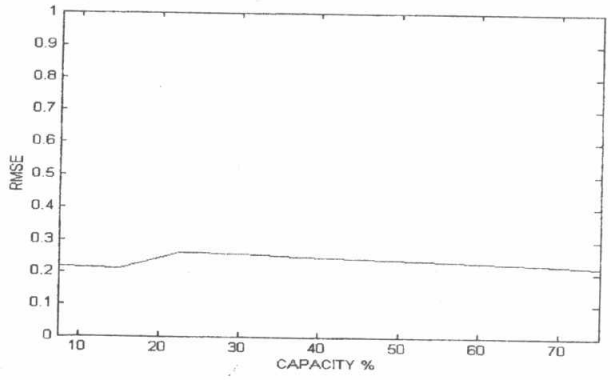

Fig. **9.** The RMSE between the embedded and extracted texts

In Figure 10, the RMSE between the covert speech signal and the stego speech against the EC is shown. The RMSE between the extracted speech and the embedded message is shown in Figure 11.

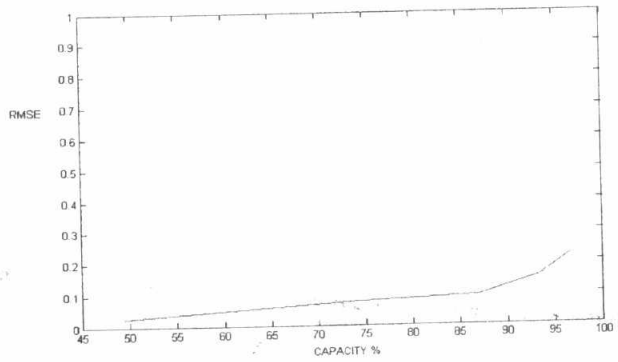

Fig. 10. The RMSE between the covert and stego speech signals

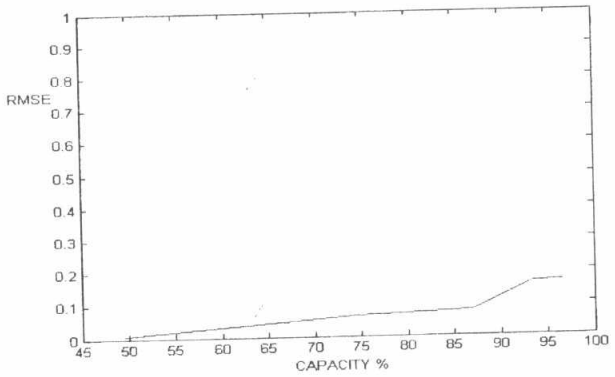

Fig. 11. The RMSE between the embedded and extracted speech signals

V. CONCLUSIONS

In this paper, a proposed technique for data hiding is introduced. The work depends on exploring the redundancy in digital data using the wavelet transform. The algorithm has been evaluated subjectively and objectively. The dependency of the quality of the stego image or speech as well as the extracted data and the embedding capacity has been studied. The proposed algorithm has shown that up to 86% from the size of the covert speech message can be used for embedding a voice message and up to 75 % in the case of embedding a mixture of text and gray level images into a single gray level

image. The obtained results encourage continuing in trying to develop the algorithm in order to increase the storage capacity, adding more security layers and complexities, and exploring new transform other than wavelet transform.

VI. REFERENCES

[1] Hany Farid, "Detecting hidden messages using higher-order statistical models", International Conference on Image Processing, Rochester, (2002).

[2] Christoph Busch, Wolfgang Funk, and Stephen Wolthusen, "Digital Watermarking: From Concepts to Real-Time Video Applications", IEEE Computer Graphics and Applications, vol. 19 no. 1 pp. 25–35, Jan./Feb., (1999).

[3] Erich J.Smythe, "Data embedding for information assurance", state-of-the-art-report, Information Assurance Technology Analysis Center, (1999).

[4] R A Isbell, "Steganography hidden menace or hidden saviour", LIRIC Associates Ltd, (2002).

[5] M.D. Swanson, M. Kobayashi, and A.H. Tewfik, "Multimedia data-embedding and watermarking technologies", Proceedings of the IEEE, vol. 86, no. 6, June (1998).

[6] Jiri Fridrich,"Applications of data hiding in digital images", ISPACS'98, (1998).

[7] Jonathan Foote and John Adcock, "Time base modulation: a new approach to watermarking audio and images",Proc,ICME (2002).

[8] Umut Uludag and Levent M. Arslan, "Audio watermarking using dc level shifting", EE 683.01 Advanced Topics in Speech Processing project report, (2001).

[9] Han-Yang Lo, Sanjeev Topiwala, and Joyce Wang," Wavelet based steganography and watermarking", Cornell University, CS 631 (1998).

[10] R. J. Anderson and F. A. P. Petitcolas,"On The Limits of Steganography", IEEE Journal of Selected Areas in Communications, vol. 16 no. 4 pp. 474–481, Special issue on copyright & privacy protection, May (1998).

[11] Mahmoud E. Gadallah and A. S. Abbas, "Secured Communication Technique for Voice Messages", 10 th International conference for Aerospace and Avionics Technology, Cairo, MTC, (2003).

[12] Mahmoud E. Gadallah and A. S. Abbas, "Data Securing using a Hybrid Cryptography/Data embedding Technique", 1st URSI-Egypt Workshop On Signal Processing", Cairo, AASTMT, (2004).

[13] Xiaolong Yuan, "Auditory Model-based Bionic Wavelet Transform For Speech Enhancement", M.SC Thesis, (2003).

[14] Michel Misiti, Yves Misiti, Georges Oppenheim, and Jean-Michel Poggi,"Wavelet Toolbox User's Guide", MathWorks, Inc., (2000).

[15] Jacob T. Jackson, Gregg H. Gunsch, Roger L. Claypoole, and Jr., Gary B. Lamont, "Blind Steganography Detection Using a Computational Immune System Approach: A Proposal", (2002).

[16] Yiwei Wang, John F. Doherty, and Robert E. Van Dyck, "A Wavelet-Based Watermarking Algorithm for Ownership Verification of Digital Images", IEEE transactions on image processing, vol. 11, no. 2, February (2002).