Audio Steganography Coding Using the Discrete Wavelet Transforms

Siwar Rekik *Siwar.Rekik@etudiant.univ-brest.fr*
Faculty of Engineering
Université de Bretagne Occidentale
Brest, 29200, France

Driss Guerchi *Driss@cud.ac.ae*
Faculty of Engineering
Canadian University of Dubai
Dubai, 117781, UAE

Habib Hamam *habib.hamam@umoncton.ca*
Faculty of Engineering
University of Moncton
Moncton, E1A 3E9, Canada

Sid-Ahmed Selouani *sid-ahmed.selouani@umcs.ca*
Faculty of Engineering
Université de Moncton
Shippagan, E8S 1P6, Canada

Abstract

The performance of audio steganography compression system using discrete wavelet transform (DWT) is investigated. Audio steganography coding is the technology of transforming stego-speech into efficiently encoded version that can be decoded in the receiver side to produce a close representation of the initial signal (non compressed). Experimental results prove the efficiency of the used compression technique since the compressed stego-speech are perceptually intelligible and indistinguishable from the equivalent initial signal, while being able to recover the initial stego-speech with slight degradation in the quality .

Keywords: Speech Compression, Steganography, Information Hiding, Discrete Wavelet Transform (DWT).

1. INTRODUCTION

One of the concerns in the area of information security is the concept of audio steganography coding. Today’s reality is still showing that communication between two parties over long distances has always been subject of interception. Providing secure communication has driven researchers to develop several cryptographic and steganographic schemes. While cryptography consists in making the signal look garbled to unauthorized people, steganography consists in secret communication by camouflaging the secret signal in another signal (named the cover signal), to avoid suspicion. Steganography is the art of hiding information in order to covert communication from eavesdroppers. To provide secure channels for communicating entities steganography is the set of techniques striving to hide the presence of secret information from a third party. When compared to encryption techniques, steganography has the advantage of arousing less suspicion. The word steganography is derived from two Greek words: Stego (means cover) and graphy (means writing). The two combined words constitute steganography, which means covert writing is the art of hiding written communications. Steganography dates back to ancient times. Several steganography were used to send message secretly during wars through the territories of enemies. The use of steganography dates back to ancient time where it was used by romans and ancient Egyptians [1]. Several steganography techniques where used to send messages secretly during wars through the territories of enemies. One technique according to Greek historian Herodotus, was to shave the head of a slave, tattoo the message on the slave's scalp, and send him after his hair grew back. Another technique was to write the secret message underneath the wax of a writing tablet. A third one is to use invisible ink to write secret messages within covert letters [2].

The relative necessities with secure channels for communication and the unlimited amount of bandwidth led us propose an audio steganography codec [3]. Therefore, there is a need to code and compress audio signals to more compact form before being transmitted. Hence researchers were incited to work on this burning field to develop schemes ensuring audio coding to reduce the coded bit rate [4]. In recent years, several applications for audio coding and compression gained ground in domains such as satellite communications, digital broadcasting, teleconferencing systems and voice mail systems.

The main task of high quality audio coding systems is to compress the signal in a way that compressed stego-signal is perceptually indistinguishable from the initial one. The stego-signal is referred to as the signal containing both cover signal and embedded information (secret information). Recently, audio compression techniques using Wavelet Transform (WT) have accommodated more attention, due to their encouraging compression ratio, Signal to noise ratio (SNR), and flexibility in representing speech signals [5]. The main issues related to the development of audio steganography codec using Discrete Wavelet Transform (DWT) are choosing optimal an wavelet transform for stego signals, decomposition level in the DWT and threshold criteria of coefficient truncation which is the basis to provide compression ratio for audio with appropriate peak signal to noise ratio (PSNR).

This paper is organized as follows: section 2, provide a brief overview of the related work on discrete wavelet transform based audio speech compression. After presenting the objectives in Section 3, we describe discrete wavelet transforms including speech decomposition and speech reconstruction in Section 4. Section 5 will describe the audio steganography compression based approach. The general step to compress a stego-speech signal is also included in this section. Then a description of the database, the parameters of our experiments, the evaluation and discussion of the results of our proposed audio steganography compression approach are presented in Section 6. Finally, we conclude and suggest directions for further research in Section 7.

2. RELATED WORK

Speech compression is the technology of removing the redundancy between neighboring samples of a speech signal and/or between the next cycles. Furthermore compression converts human speech into an efficiently encoded illustration that can later be decoded to produce a close approximation of the original signal (in this paper the original signal is referred to as the original stego-signal before encoding). Compression techniques can be classified into two main categories that can be used to reduce the coded bit rate: lossless and lossy. The first techniques take advantage from the statistical redundancy in the audio signal, in these methods the original audio signal can be completely recovered from the encoded signal. In the second method, the original and reconstructed audio signal are not completely identical, this technique separate necessary information and perceptual irrelevant signal than can be removed later.

D. Sinha et al., [6] have presented a new approach for audio compression using discrete wavelet transform. This technique is based on the dynamic dictionary approach with selection of the best adaptive wavelet choice and optimal coefficients quantization procedures. The proposed technique takes advantages from the masking effect persisting in the human hearing through the optimal wavelet transform selection and transform bit allocation measures. A permanent distortion level is used in order to reduce the required quantity of bits representing each frame of audio signal. The used dynamic dictionary significantly decreases statistical redundancies in the audio source.The experimental results of the proposed method able to accomplish a transparent coding of monophonic compact disk (CD), sampled at 44.1 kHz at bit rates of 64-70 kilobits per second (kb/s). The combined adaptive wavelet selection and dynamic dictionary coding procedures realize approximately transparent coding of monophonic CD quality signals at bit rates of 48-66 kb/s.

P. Srinivasan et al., [7] have proposed a new high quality audio compression approach using an adaptive wavelet packet to achieve perceptual transparent compression of high-quality (44.1KHz) audio signals at about 45 kb/s. The adapted filter bank structure available at the decoder can achieve according to psychoacoustic criteria and computational complexity. The proposed technique takes advantage from the availability of a computational power in order to realize real time coding/decoding. The bit allocation method is an adapted zero-tree scheme taking input from the psychoacoustic model. The quantity performance measurement of the proposed technique named sub-band perceptual rate, is adapted by the filter bank structure to approach the perceptual entropy (PE) as closely as possible. This technique is able to accomplish a good quality of possible reconstruction considering the size of the bit stream existing at the decoder side, amenable to progressive transmission. Thus this technique presents a new scheme to approve the results in wavelet packets and perceptual coding in order to make a well matched algorithm to high-quality audio dedicated for internet and storage applications.

The wavelet analysis process is to implement a wavelet prototype function, known as an analyzing wavelet or mother wavelet. Coefficients in a linear combination of the wavelet function can be used in order to represent the development of the original signal in terms of a wavelet, data operations can be performed with the appropriate wavelet coefficients [8]. Choose the best wavelets adapted to represent your data, also truncate the coefficients below a threshold; your data is sparsely signified. Because of this sparse coding the wavelets is considered as an exceptional technique in the field of data compression. A general overview of the discreet wavelet transform is given in this section.

G.Amara et al.,[9] have presented a general overview of wavelets that cut up information into different frequencies in order to perform a study of these different component with resolution corresponding to its scale. They introduced wavelets to the concerned industrial person outside the digital signal processing domain. A detailed history description of wavelets have been presented starting with the Fourier method, a comparison between these two methods was investigated. The main goal of this investigation was to state the properties and the special aspects of wavelets, and finally to list some interesting applications such as image compression, musical tones, and de-noising noisy data.

G.Tzanetakis et al.,[10] have given a brief description of their study that consist on analyzing the temporal and spectral properties of non stationary signals such as audio using the Discrete Wavelet Transform. They also give a detailed description of some applications using Discrete Wavelet Transform and the difficulties of extracting data from non-speech audio. A detailed automatic classification of different types of audio using Discrete Wavelet Transform is described also a comparison with other traditional feature extractor proposed in the literature was given.

M. L. Hilton et al,. [11] Have proposed an adaptive data selecting scheme for the threshold for wavelet contraction based noise removal. The proposed method involves a statistical test of theory based on a two dimensional cumulative sum of wavelet coefficients, that takes into consideration the coefficients magnitude and their positions.

G. Kronquist et al., [12], have presented a thresholding method on the discrete wavelet coefficients since the background noise has to be removed from the speech signal.

A soft thresholding have been used and particularly adopted to speech signal in order to reduce the coefficients. The training sequence is used to determine the noise levels which are adaptively changed. Their proposed technique of thresholding for denoising speech signal improve that the signals does not change the characteristics of background noise, only its amplitude is decreased.

3. OBJECTIVES

Our objective is to develop a high performance compression speech steganography system. The design of such system mainly consists in the optimization of the following attributes:

- The compression ratio, used to quantify the reduction in speech-representation size produced by a compression algorithm. It is defined as the ratio of the size of the compressed signal to that of the initial signal.
- The impact of the compression process on the initial stego-speech (stego-signal) quality. We intend to produce a compressed stego-signal that is perceptually indistinguishable from the initial signal.
- Lossless compression, our aim is to allow the reconstruction of the initial stego-speech from the compressed signal after going through the compression-decompression process.
- The accuracy with which the compressed signal can be recovered at the receiver. Efficient techniques are to be developed to minimize the impact of compression on the stego-signal.

4. SPEECH DISCRETE WAVELET TRANSFORMS

The wavelet transform transforms the signal from the time domain to the wavelet domain. This new domain contains more complicated basis functions called wavelets, mother wavelets or analyzing wavelets [13]. The fundamental idea behind wavelets is to analyze the behavior of the signal with respect to scale. Any signal can then be represented by translated and scaled versions of the mother wavelet. Wavelet analysis is capable of enlightening aspects of data that other signal analysis techniques are unable to perform, aspects like trends, discontinuities in higher derivatives, breakdown points and self-similarity.

The basic idea of DWT for one-dimensional signals is shortly described. The wavelet analysis enables splitting a signal in two parts, usually the high frequencies and the low frequencies part. This process is called decomposition. The edge components of the signal are largely limited to the high frequencies part. The signal goes through series of high pass filters to analyze the high frequencies, and goes through series of low pass filters to analyze the low frequencies. Filters of different cutoff frequencies are used to analyze the signal at different resolutions [14,15].
The DWT involves choosing scales and positions based on powers of two, so called dyadic scales and positions. The mother wavelet is rescaled by powers of two and transformed by integers. Specifically, a function $f(t) \in L^2(R)$ (defines space of square integrable functions) can be represented as:

$$f(t) = \sum_{j=1}^{L} \sum_{k=-\infty}^{\infty} d(j,k)\psi(2^{-j}t-k) + \sum_{k=-\infty}^{\infty} a(L,k)\phi(2^{-L}t-k) \qquad (1)$$

The function $\psi(t)$ is known as the mother wavelet, while $\phi(t)$ is known as the scaling function. The set of function $\left\{\sqrt{2^{-L}}\phi(2^{-L}t-k), \sqrt{2^{-j}}\psi\left(2^{-j}t-k\right) \middle| j \le L, j,k,L \in Z\right\}$,Where Z is the set of integers is an orthonormal basis for $L^2(R)$ .The numbers $a(L,k)$ are known as the approximation coefficients at scale L, while $d(j,k)$ are identified as the detail coefficients at scale j. The approximation and detail coefficients can be expressed consecutively as:

$$a\left(L,k\right) = \frac{1}{\sqrt{2^L}} \int_{-\infty}^{\infty} f\left(t\right)\phi(2^{-L}t-k)dt \qquad (2)$$

$$d(j,k) = \frac{1}{\sqrt{2^j}} \int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)dt \qquad (3)$$

To better understand the above coefficients let's consider a projection $f_l(t)$ of the function $f(t)$ that provides the best approximation (in the sense of minimum error energy) to $f(t)$ at a scale l. This projection can be constructed from the coefficients $a(L,k)$, using the equation:

$$f_l(t) = \sum_{k=-\infty}^{\infty} a(l,k)\phi\left(2^{-l}t - k\right) \qquad (4)$$

As the scale L decreases, the approximation becomes finer, converging to $f(t)$ as $l \to 0$. The difference between the approximation at scale $l+1$ and that at l, $f_{l+1}(t) - f_l(t)$, is totally defined by the coefficients $d(j,k)$ using the equation of decomposition and can mathematically be expressed as follows:

$$f_{l+1}(t) - f_l(t) = \sum_{k=-\infty}^{\infty} d(l,k)\psi(2^{-l}t - k) \qquad (5)$$

Using these relations, given $a(L,k)$ and $\{d(j,k) \mid j \le L\}$, are useful for building the approximation at any scale. Hence, the wavelet transform breaks the signal up into a coarse approximation $f_L(t)$ (given $a(L,k)$)and a number of layers of detail $\{f_{j+1} - f_j(t) \mid j < L\}$ (given by $\{d(j,k) \mid j \le L\}$). As one layer of details is added, the approximation at the next higher scale is achieved.

4.1 Signal decomposition

Starting with a discrete input stego-speech signal, the primary steps of the DWT algorithm consists in decomposing the signal into sets of coefficients. These are the approximation coefficients cA1 (low frequency information, Figure 1) and the detail coefficients cD1 (high frequency information). In order to obtain the coefficient vectors, the signal s goes through the low-pass filter Lo_D (mathematically stated, a convolution operation is performed) and through the high-pass filter Hi_D for details. A down sampling by a factor of 2 or a dyadic decimation is then applied to obtain the approximation coefficients [16]. The filtering operation process of the DWT is shown in Figure 1.

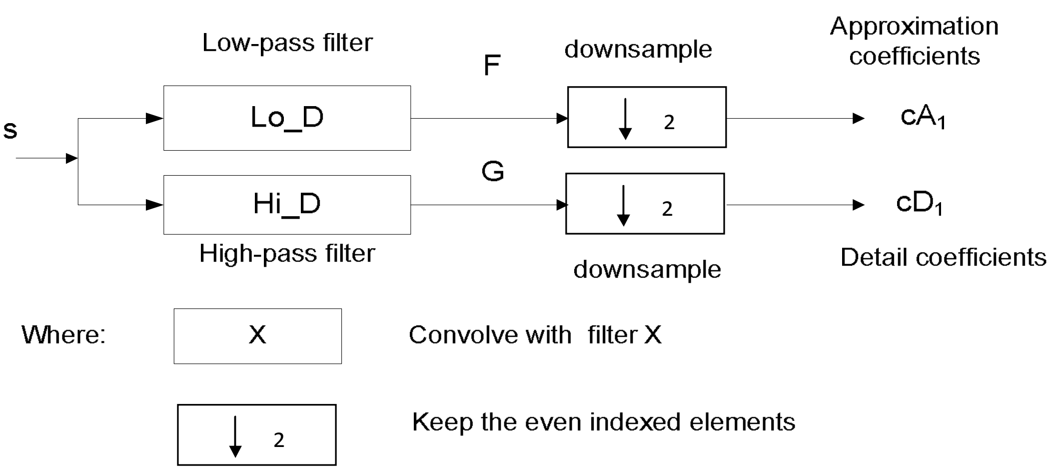

FIGURE 1: Filtering Operation of the DWT

Mathematically the two-channel filtering of the discrete signal can be represented by:

$$cA_1 = \sum_k c_k s_{2i-k} \quad (6), \quad cD_1 = \sum_k g_k s_{2i-k} \quad (7)$$

These equations implement a convolution with a down sampling by a factor 2, then transfer the forward discrete wavelet transform. If the length of the initial stego-signal s is equal to n, and if the length of all filter is equivalent to 2N, then the equivalent lengths of the coefficients cA1 and cD1 are calculated by:

$$floor(\frac{n-1}{2})+N \qquad (8)$$

This shows that the total length of the wavelet coefficients vector is always slightly greater than the length of the initial signal due to the filtering process used. Wavelet decomposition tree can be constructed by following an iterative decomposition process with successive approximations [17]. Thus, the input stego-signal is broken down in several subordinate resolution components. Figure 2 shows the decomposition in approximation and details of signal s in 3 levels.

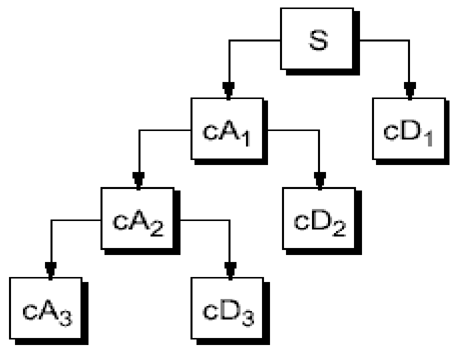

FIGURE2: Decomposition of DWT coefficients

4.2 Signal Reconstruction

The initial signal can be reconstructed using the Inverse Discrete Wavelet Transform (IDWT), following the above procedure of Figure 2 in the reverse order. As pointed out is shown in Figure 3, the synthesis starts with the approximation and detail coefficients cA_j and cD_j, and then reconstructs $cA_j - 1$ by up sampling and filtering with the reconstruction filters [18, 19].

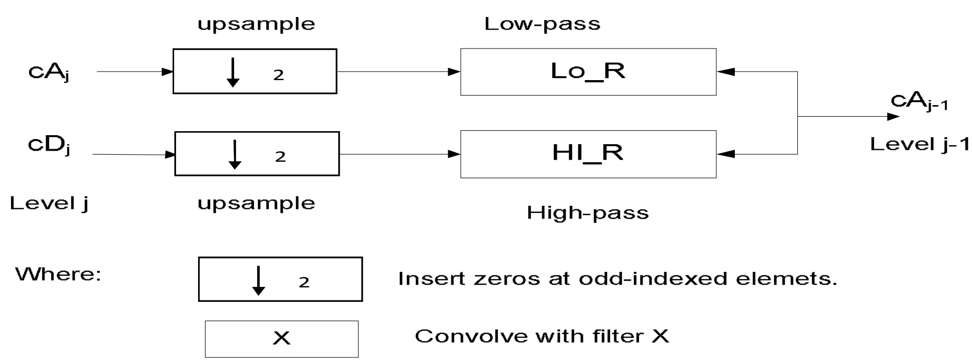

FIGURE3: Wavelets Reconstruction

The effect of aliasing created in the wavelet decomposition stage is revoked with the effect of the reconstruction filters. This process forms a system known as quadrature mirror filters (QMF) composed with the reconstruction filters (Lo_R and Hi_R) and with the low and high pass decomposition filters. For a multilevel analysis, approximations at finer resolutions and a synthesized initial signal can be produced during the reconstruction procedure.

5. AUDIO STEGANOGRAPHY COMPRESSION BASED APPROACH

The idea behind stego-signal compression using wavelets is principally related to the relative scarceness of the wavelet domain representation for the signal. Wavelets concentrate speech information (energy and perception) into a few neighboring coefficients [20].Therefore; several coefficients will either be zero or have insignificant magnitudes, this effect resulting from taking the wavelet transform of a signal [21]. Data compression is then achieved by treating small valued coefficients as insignificant data and thus removing them the process of compressing a speech signal using wavelets involves a number of different stages, each of which are discussed below. Figure 4 shows the block diagram of the different steps involved in the compression of the stego-speech signal by using discrete wavelet transform, using Matlab version 9. In designing a wavelet based stego-speech coder, the major issues covered in this section are:

1. Hiding an audio speech signal in a cover signal
2. Setting up an audio steganography database using a different hiding technique.
3. Choosing optimal wavelets for stego-speech,
4. Selecting decomposition level in wavelet transforms,
5. Choosing Threshold criteria for the truncation of coefficients,
6. Efficiently representing zero valued coefficients,
7. Quantizing and digitally encoding the coefficients,
8. Signal reconstruction

The performance of the wavelet compression method in coding stego-speech signals and the quality of the reconstructed signals is also evaluated.

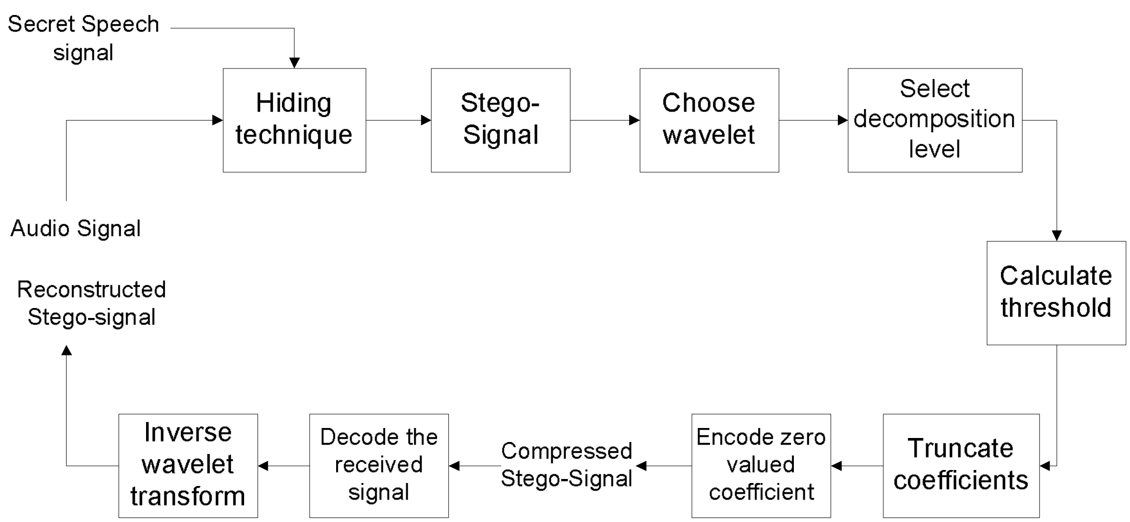

FIGURE 4: Block diagram of the based stego-speech encoder/decoder

5.1. Audio Steganography Database

Steganography provides secrecy by making secret information invisible to opponents [22]. Steganography hide the data by embedding it in another data medium called cover. In this paper we intend to apply our compression techniques on speech by generating speech-in-speech files. In order to generate our stego-speech database, we use three different steganographic techniques. The first one was our speech steganography technique using discrete wavelet and fast fourier transform developed in [23]. The two other methods are Steganos[24] and StegHide [25]. These tools were chosen on the basis of being popular methods and also with readily available software.

5.2 Choice Of The Mother-Wavelet

The choice of the appropriate mother-wavelet is of great importance for designing high quality speech coders. The choice of the optimal wavelet function is based on several criteria. Since the main objective is to maximize the signal to noise ratio (SNR). In general the amount of energy a wavelet basis function can concentrate into the level 1 approximation is one of the criteria that can be used to select the optimum wavelet. By giving a better SNR ratio, daubechies wavelets provide a good compression property for wavelet coefficients.

5.3 Wavelet Decomposition

For a given stego-signal the wavelet starts by decomposing a signal at different level that can reach up to L = 2K levels, where K is the length of the discrete signal. Thus we can perform the transform at any of these levels. The type of signal being analyzed usually affects the choice of the decomposition level. In order to represent the accurately signal components the selection of the appropriate number of approximation and details coefficients is extremely important in the compression procedure. For the processing of speech signals decomposition up to scale 5 is sufficient, with no further advantage gained in processing beyond scale 5 [26]. In the paper, comparisons were made with level 4 and 6 decomposition after performing levels 5 decomposition.

5.4 Threshold Calculation

After applying DWT the obtained coefficients on the frame concentrate energy in few neighbors. Thus we can truncate all coefficients with “low” energy and preserve those holding the high energy value. Two different techniques can be used for calculating thresholds. The first, recognized as Global Thresholding consist of taking the wavelet expansion of the signal and preserving the largest absolute value coefficients. In this method we can set a global threshold manually, thus just a single parameter needs to be selected. The coefficient values below this value should be set to zero, to achieve compression. The second technique known as By Level Thresholding consists of applying visually determined level dependent thresholds to each decomposition level in the wavelet transform [27].

The following figure5 shows the setting of global threshold for a typical stego-speech signal. In this figure, the X-axis represents the coefficient values. The black (dark) vertical line moves to right or left, thereby changing the threshold. The intersection of this line with green line indicates the percentage of zero coefficients below this threshold. Its intersection with the red line indicates the percentage of signal energy retained after truncating these coefficients to zero.

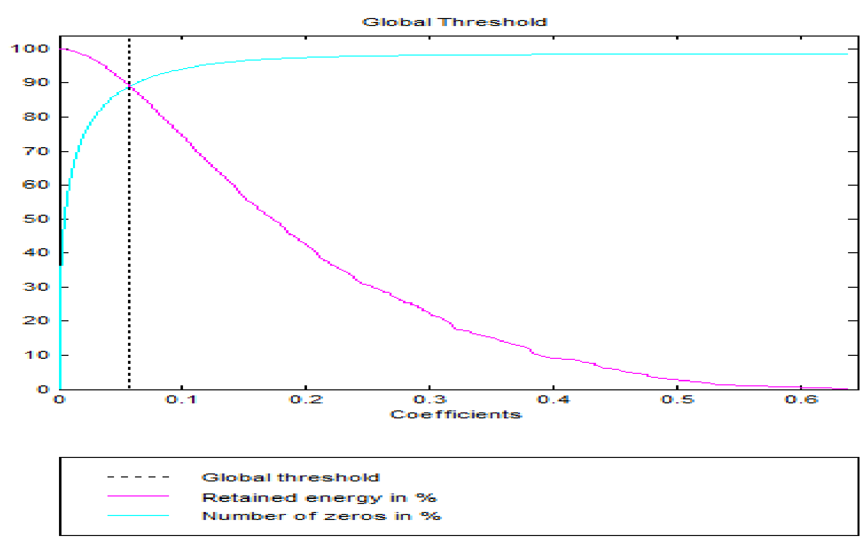

FIGURE 5: Setting of global threshold for a typical stego-speech signal

5.5 Encoding Coefficients

Signal compression is achieved by first truncating small-valued coefficients and then encoding them. Storing the coefficients along with their corresponding positions in the wavelet transform vector, in order to represent the high-magnitude coefficients [28]. For a stego-speech signal of frame size F, taking the DWT generates a frame of size T, slightly larger than F. If only the largest L coefficients are retained, then the compression ratio C is given by:

$$C = \frac{F}{2L} \qquad (9)$$

Another approach to compression is to encode consecutive zero valued coefficients [29], with two bytes. One byte to present a series of zeros in the wavelet transforms vector and the second byte representing the number of consecutive zeros. The transform vector has to be compressed, after zeroing wavelet coefficients with negligible values based on either calculating threshold values or simply selecting a truncation percentage.

In this proposed audio steganography coding, consecutive zero valued coefficients are encoded with two bytes. One byte is used to identify a starting string of zeros and the second byte keeps track of the number of successive zeros. Due to the scarcity of the wavelet representation of the stego-speech signal, this encoding technique leads to a higher compression ratio than storing the non-zero coefficients along with their respective positions in the wavelet transform vector, as suggested in the literature review [30].

6. EVALUATION

6.1 Experimental Setup

To evaluate the performance of the proposed audio steganography coding, we conducted several computer simulations using NOIZEUS database [31,32,33]. This corpus contains thirty sentences from the IEEE sentence database, recorded in a sound-proof booth using Tucker Davis Technologies (TDT) recording equipment. The sentences are produced by three male and female speakers. The thirty sentences: 15 male and 15 female include all phonemes in the American English language. The sentences were originally sampled at 25 kHz and down-sampled to 8 kHz. The length of the speech file varies between 0.02 ms to 0.03 ms. In the comparative evaluation, we conducted three sets of tests, before starting the compression procedure. In the first set of simulations, we embedded each of the 15 male speech files in the remaining 14 male speech files using DWT-FFT technique [23]. In the second set of tests, we hide the same gender speech files using StegHide software [24]. In the third set of tests, we hide the same gender speech files using Steganos software [25]. Every set is iterated with female speech files. Each set is iterated for 2 different wavelet families (Haar, Daubechies).Db4,Db6,Db8,Db10, Haar. The selection of these appropriate mother-wavelets is based on the amount of energy a wavelet basis function can concentrate into the level 1-approximation coefficients.

All the stego-speech signals were decomposed to scale 5 and Global Thresholding was applied.
The entire signal was decomposed at once without framing. A summary of the performance is given below for the different used wavelets.

6.2 Evaluation Outcomes

One of the performance measures of any stego-speech coding system is the comparison between the original (stego-speech) and the compressed signals. In this work, we used subjective and objective performance measures. In the subjective measures, we conducted several informal listening comparative tests. In these simulations, we played in a random order the original signal and the compressed signal to several listeners. Each listener had to identify the better quality speech file among the original and the compressed signals. The majority of listeners couldn't distinguish between the two speech files.

As an objective measure, a number of quantitative parameters is used to assess the performance of the wavelet based stego-speech coder, in terms of both reconstructed signal quality after decoding and compression scores. The following parameters are compared (Signal to Noise Ratio (SNR), Normalized Root Mean Square Error (NRMSE), and Retained Signal Energy).

The results obtained for the above quantities are calculated using the following formulas:

- Signal to Noise Ratio

$$SNR = 10\log10\left(\frac{\sigma_x^2}{\sigma_e^2}\right) \qquad (10)$$

σ_x^2 is the mean square of the speech signal and σ_e^2 is is the mean square difference between the original and reconstructed signals.

- Peak Signal to Noise Ratio

$$PSNR = 10\log_{10}\frac{NX^2}{\left\|x-r\right\|^2} \qquad (11)$$

N is the length of the reconstructed signal, X is the maximum absolute square value of the signal x and $\left\|x-r\right\|^2$ is the energy of the difference between the original and reconstructed signals.

- Normalized Root Mean Square Error

$$NRSME = \sqrt{\frac{(x(n)-r(n))^2}{(x(n)-\mu_x(n))^2}} \qquad (12)$$

$x(n)$ is the speech signal, $r(n)$ is the reconstructed signal, and $\mu_x(n)$ is the mean of the speech signal.

- Retained Signal Energy

$$RSE = \frac{100*\left\|x(n)\right\|^2}{\left\|r(n)\right\|^2} \qquad (13)$$

$\left\|x(n)\right\|$ is the norm of the original signal and $\left\|r(n)\right\|$ is the norm of the reconstructed signal. For one-dimensional orthogonal wavelets the retained energy is equal to the L2-norm recovery performance.

- Compression Ratio

$$C = \frac{Length(x(n))}{Lenght(cWC)} \qquad (14)$$

cWC is the length of the compressed wavelet transform vector. The original stego-speech signal which was used to obtain the performance measure.

In table 1, we present the average of the used quantitative parameters for each of the three different sets of tests for level 4 decomposition. In table 2, we present the average of the same sets of tests for level 5 decomposition. For level 6 decomposition the result is presenting is table 3.

Level 4	**Male speaker**					**Female speaker**				
Wavelet	**C**	**SNR**	**PSNR**	**NRMSE**	**RSE (%)**	**C**	**SNR**	**PSNR**	**NRMSE**	**RSE (%)**
Db4	2.95	18.86	83.65	0.86	95.88	2.82	18.76	86.64	0.57	94.98
Db6	2.21	18.63	84.53	0.59	96.87	2.61	18.73	79.43	0.58	97.98
Db8	3.09	18.58	79.68	0.68	97.72	3.01	17.58	76.54	0.75	96.87
Db10	3.19	17.91	74.83	0.59	96.75	3.09	17.61	74.63	0.77	96.78
Haar	3.18	17.55	73.24	0.89	94.75	3.05	18.45	83.24	0.83	96.84

TABLE 1: A male and female Stego-speech decomposed at level 4.

Level 5	**Male speaker**					**Female speaker**				
Wavelet	**C**	**SNR**	**PSNR**	**NRMSE**	**RSE (%)**	**C**	**SNR**	**PSNR**	**NRMSE**	**RSE (%)**
Db4	3.15	19.86	87.65	0.57	96.88	3.08	19.76	87.65	0.67	95.88
Db6	3.21	19.93	89.45	0.37	98.98	3.11	19.83	89.45	0.39	96.98
Db8	3.19	19.68	86.64	0.58	97.92	3.01	19.58	86.64	0.58	96.87
Db10	3.09	19.61	84.73	0.47	96.69	3.09	19.51	84.73	0.47	96.78
Haar	3.15	19.75	83.24	0.78	97.85	3.05	19.65	83.24	0.77	96.84

TABLE 2: A male and female Stego-speech decomposed at level 5.

Level 6	**Male speaker**					**Female speaker**				
Wavelet	**C**	**SNR**	**PSNR**	**NRMSE**	**RSE (%)**	**C**	**SNR**	**PSNR**	**NRMSE**	**RSE (%)**
Db4	2.45	17.86	82.65	0.93	94.88	2.92	17.76	85.64	0.97	93.98
Db6	2.01	17.63	83.53	0.69	95.87	2.51	17.73	78.43	0.78	96.98
Db8	2.09	17.58	78.68	0.78	96.72	2.91	17.68	75.54	0.85	95.87
Db10	2.19	16.91	73.83	0.79	95.75	3.09	17.11	73.63	0.97	96.38
Haar	2.88	17.55	72.24	0.99	93.75	2.45	18.15	82.24	0.73	96.14

TABLE 3: A male and female Stego-speech decomposed at level 6.

It is observed that no advantage is gained in going beyond scale 5 and usually processing at a lower scale leads to a better compression ratio. Therefore, for processing speech signals choosing the right decomposition level in the DWT is important. Male voices have relatively more approximate coefficients than female voices for all levels decomposition. Based on the energy retained in the first N/2 coefficients criterion, Daubechies 6 preserves perceptual information better than all the other wavelets tested. The Db6 wavelet also provides the highest SNR, PSNR, compression ratio, and lowest NRMSE, as shown in table 2.

Figure 6 show the original stego-signal in red color and the compressed stego-signal in black color using Db6 wavelet and 5 level decomposition.

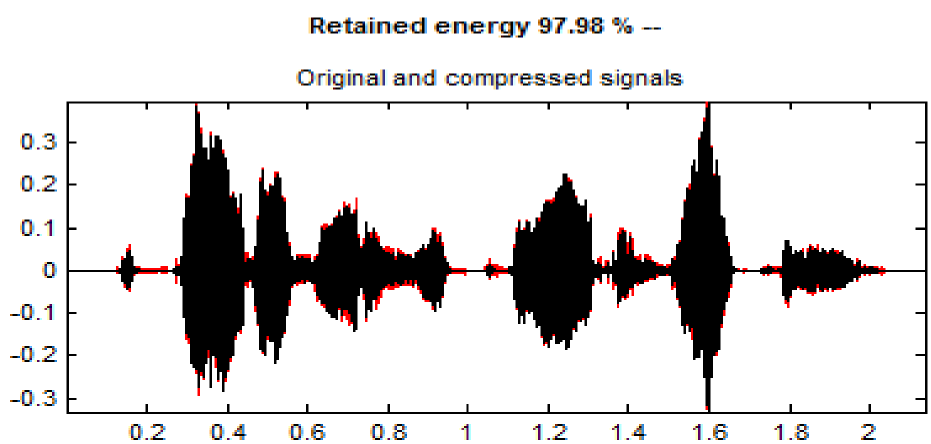

FIGURE 6: Original and compressed stego-signal

The performance of the stego-audio codec is evaluated by taking into consideration diverse parameters such as steganography system, decomposition levels, optimal wavelets and threshold value for wavelet coefficients in order to obtain low bit rate signal. In the simulation results show we can observed that the optimum number of wavelet decomposition level is 5 by using the Daubechies 6 wavelet. More specifically, the results demonstrates that the Daubechies wavelets best suits the compression of stego-speech signals due to their high SNR values of 19.93 dB and low NRMSE value of 0.39 compared with the other family wavelets. It is well recognized that the SNR cannot faithfully indicate speech quality. For evaluation of performance of wavelets for speech enhancement one more criteria used is subjective test. As subjective measures we have conducted informal listening tests. In these simulations, the evaluators have listened randomly to both the original stego-speech and the compressed signal and give their opinion for which one has the better quality. Most of the listeners couldn't distinguish between original stego-speech a compressed speech.

7. CONCLUSIONS

This work has proven through informal tests that our audio steganography compression technique is robust. The proposed coding method produces a compressed stego-speech files that are indistinguishable from their equivalent stego- speech files. Therefore this method does not depend on the type of steganographic technique. The results show that the use of wavelet transform achieves high compression ratios with acceptable SNR. The proposed wavelet based steganography compression system reaches an SNR of 19.93 dB at a compression ratio of 3.21 by using the Daubechies 6 wavelet. Performance was measured by using the Haar and Daubechies wavelet families. These parameters values are very significant in design efficient wavelet based stego-speech compression software for mobile applications and multimedia. Furthermore the compression ratios can be easily varied when using the wavelets technique, while most other compression method has a fixed compression ratio.

In future work, we will extend this work to applications involving Voice over IP (VoIP) audio steganography coding and to using other speech coding technique such as excited linear predictive coding (CELP) for speech.

8. REFERENCES

[1] Petitcolas, Fabian A.P.; Katzenbeisser, Stefan. "Information Hiding Techniques for Steganography and Digital Watermarking". Artech House Publishers. ISBN 1-58053-035-4. (2000).

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

[3] S. I. Gel'fand and M. S. Pinsker, "Coding for channel with random parameters," Prob. of Control and Inform. Theory, vol. 9, no. 1, pp. 19–31, 1980.

[4] N. F. Johnson and S. Katzenbeisser. "Information hiding - a survey", Artech House, Norwood, MA, 2000.

[5] Oppenheim, Alan V.; Schafer, Ronald W.; Buck, John A. (1999). Discrete-time signal processing. Upper Saddle River, N.J.: Prentice Hall. pp. 468–471. ISBN 0-13-754920-2.

[6] D. Sinha and A. H. Tewfik, "Low bit rate transparent audio compression using adapted wavelets," IEEE Transactions on Signal Processing, vol. 41, no. 12, pp. 3463–3479, 1993.

[7] P. Srinivasan and L. H. Jamieson, "High Quality Audio Compression using an Adaptive Wavelet Packet decomposition and Psychoacoustic Modelling," *IEEE Transactions on Signal Processing* , Vol. 46, No. 4, April 1998, pp. 1085-1093.

[8] 1/C Chris Eubanks**, "**Haar Wavelets, Image Compression, and Multi-Resolution Analysis**",** Initial Report for Capstone Paper**,** SM422 April 4, 2007**.**

[9] G.Amara "An introduction to Wavelets", IEEE Computational Science and Engineering, 1995, vol. 2, num. 2.

[10] G.Tzanetakis, G. Essl, P. Cook," Audio Analysis using the Discrete Wavelet Transform", conference on Signal Processing, 2001.

[11] M.L. Hilton, R.T, Ogden, "Data analytic wavelet threshold selection in 2-D signal denoising", Signal Processing, IEEE Transactions on, Issue 2, pages 496-500, August 2002.

[12] G. Kronquist and H. Storm, "Target Detection with Local Discriminant Bases and Wavelets," Proc. of SPIE, v. 3710, pp 675-683, 1999.

[13] M. Misiti,Y.Misiti, G.Oppenheim and J-M Poggi,"Wavelet Toolbox™ 4 User's Guide", The Math Works Inc., 2000.

[14] B. Lin, B. Nguyen and E.T. Olsen, "Orthogonal Wavelets and Signal Processing", Signal Processing Methods for Audio, Images and Telecommunications,P.M.Clarkson and H.Stark, ed., Academic Press, London, 1995, pp. 1-70.

[15] S. Mallat, "A Wavelet Tour of Signal Processing", Academic Press, San Diego, Calif., 1998.

[16] J.I. Agbinya, "Discrete Wavelet Transform Techniques in Speech Processing",IEEE Tencon Digital Signal Processing Applications Proceedings, IEEE, New York, NY, 1996, pp 514-519.

[17] H.Shahina and T. Takara, "Performance Comparison of DaubechiesWavelet Family for Bangla Vowel Synthesis", JOURNAL OF COMPUTING, VOLUME 3, ISSUE 9, SEPTEMBER 2011, ISSN 2151-9617.

[18] Y. Nievergelt, "Wavelets made easy", Birkhäuser, Boston, 1999.

[19] J. Ooi and V. Viswanathan, "Applications of Wavelets to Speech Processing", Modern Methods of Speech Processing, R.P. Ramachandran and R. Mammone, ed., Kluwer Academic Publishers, Boston, 1995, pp. 449-464.

[20] Cai, T. and Silverman, B. W. "Incorporating information on neighboring coefficients into wavelet estimation", (1999).

[21] Coifman, R.R. and Donoho, D.L. (1995). Translation invariant denoising. In A. Antoniadis and G. Oppenheim (eds),Wavelets and Statistics, Lecture Notes in Statistics 103. NewYork: Springer-Verlag, pp. 125–150.

[22] M.M. Amin, M. Salleh, S. Ibrahim, et al., "Information Hiding Using Steganography", *4th* National Conference On Telecommunication Technology Proceedings *(NCTT2003),* Shah Alam, Malaysia, pp. 21-25, January 14-15, 2003.

[23] R.Siwar, G.Driss, S.Sid-Ahmed, H. Habib "Speech Steganography using Wavelet and Fourier Transforms" EURASIP Journal on Audio, Speech, and Music Processing, Accepted.

[24] Steganos, http://www.steganos.com.

[25] http://steghide.sourceforge.net/

[26] J.I. Agbinya, .Discrete Wavelet Transform Techniques in Speech Processing,.IEEE Tencon Digital Signal Processing Applications Proceedings, IEEE, New York, NY, 1996, pp 514-519.

[27] A.Jayaraman ,P.Indumathy ,"A New Threshold Calculation Approach in the Performance Enhancement of Spread Spectrum System Using Double Density Discrete Wavelet Filter" Communications in Computer and Information Science, 2010, Volume 101, Part 3, 654-659, DOI: 10.1007/978-3-642-15766-0_115.

[28] E.B. Fgee, W.J. Phillips, W. Robertson, .Comparing Audio Compression using Wavelets with other Audio Compression Schemes,. IEEE Canadian Conference on Electrical and Computer Engineering, IEEE, Edmonton, Canada, 1999, pp. 698-701.

[29] W. Kinsner and A. Langi, “Speech and Image Signal Compression with Wavelets”,. IEEE Wescanex Conference Proceedings, IEEE, New York, NY, 1993, pp. 368-375.

[30] B. Lin, B. Nguyen and E.T. Olsen, “Orthogonal Wavelets and Signal Processing”,.Signal Processing Methods for Audio, Images and Telecommunications, P.M.Clarkson and H.Stark, ed., Academic Press, London, 1995, pp. 1-70.

[31] Hu, Y. and Loizou, P. “Subjective evaluation and comparison of speech enhancement algorithms,” Speech Communication, 2007, 49, 588-601.

[32] Hu, Y. and Loizou, P. “Evaluation of objective quality measures for speech enhancement,” IEEE Transactions on Speech and Audio Processing, 2008, 16(1), 229-238.

[33] Ma, J., Hu, Y. and Loizou, P. "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions", Journal of the Acoustical Society of America, 2009,125(5), 3387-3405