Indian Journal of Science and Technology, *Vol 7(3), 323–334, March 2014*
ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645

Lossless Audio Steganography based on Lifting Wavelet Transform and Dynamic Stego Key

Haider Ismael Shahadi[1,2*], Razali Jidin[2] and Wong Hung Way[2]

[1]Electrical Engineering Department, University of Babylon, Hila, Babil, Iraq; haider_almayaly@yahoo.com
[2]Centre for Automation & Embedded Computing System, Tenaga National University (UNITEN), Kajang, Malaysia; Razali@uniten.edu.my, Hungww@uniten.edu.my

Abstract

This paper presents a new lossless audio steganography approach based on Integer-to-Integer Lifting Wavelet Transform (Int2Int LWT) and Least Significant Bits (LSBs) substitution. In order to increase the security level a simple encryption with adaptive key has been proposed. The experimental results show that this approach has excellent transparency (above 45 dB of signal to noise ratio) with fixed high embedding capacity (25% from the audio cover signal size) and full recovery for the hidden secret message. Furthermore, robustness tests show immunity of the method against additive noise and perceptual statistical analysis. The proposed hiding and recovery procedures are simple and symmetry; therefore it can be easily used for real-time covert communication.

Keywords: Audio Steganography, Embedding Capacity, Full Recovery, Imperceptibility, Security Lifting Wavelet Transform (LWT)

1. Introduction

Digital steganography plays recently a major role for security beside cryptography. Steganography can provide higher level of security than cryptography due to it hides secret information in a cover medium without attracting attention[1,2]. While encrypted message may arouse the suspicion of an eavesdropper. Nevertheless, both the technologies can be combined for a higher level of information protection[3].

There are several features should be available in steganography[3,4]. Firstly, perceptual quality or imperceptibility that means a resulted signal from the steganography process which is named stego signal should be closes to the cover signal. In other word, an eavesdropper cannot imperceptible the difference between the cover and stego signal. Secondly, embedding capacity that refers to the amount of information that can be inserted into the cover signal without losing imperceptibility and it should be high. Thirdly, robustness, that measures to the immunity of the hidden data against additive noise and attacking. Fourthly, security that indicates embedded messages should be secure and not exposed by any evidences about their existence in the cover signal. Fifthly, computational complexity that refers to an algorithm complexity and its processing time, the good algorithm should be not complex. There is a trade-off among the above features. For instance, increasing an embedding capacity in a steganography system leads to perceptual quality degradation of the stego signal and decreasing of robustness.

During the last decade numerous audio steganography schemes have been proposed, however, many of them have absence for one or more feature. In literature, we will review only the methods that verify high embedding capacity with good imperceptibility, these method usually use Least Significant Bits (LSBs) substitution either in time or transform domain to satisfy the above two purposes (high embedding capacity with good imperceptibility). The LSBs methods introduced in the time domain have low time processing and low complexity[4],

*Author for correspondence

However, the direct LSBs methods are sensitive to additive noise; thus, authors[5] have attempted to increase the robustness of LSB methods by altering the depth of the inserted data in higher layer (bit position) in the cover sample. The technique[5] has less perceptual quality than conventional LSBs method.

Several audio steganography techniques hide secret data in the LSBs of the transform domain coefficients such as Discrete Cosine Transform (DCT)[6], Discrete Furrier Transform (DFT)[7], and Discrete Wavelet Transform (DWT)[7–9]. All these techniques provide high embedding capacity with good perceptual quality. The main disadvantage in these techniques is in the robustness. Since the above techniques require data type conversion (integer to floating conversion and vice versa) before and after the data embedding, thus, errors can occur in the recovered messages because of the losing that may happened in rounding processes. To reduce or eliminate this type of errors in the retrieved hidden data, algorithms that adopt an integer transform domain have been developed[10,11]. The algorithms[10,11] can satisfy embedding capacity about 25% from the audio cover size with good imperceptibility, however the algorithms are complex somewhere. The algorithms[12,13] are simples and satisfy full recovery in normal case (without noise), however the embedding capacity is less than 20% from the audio cover size with critical imperceptibility for embedding capacities between 15–20% from the file size.

In general, audio steganography that based on integer to integer (Int2Int) Lifting Wavelet Transform (LWT) can satisfy full recovery for the embedded secret messages in the receiver side. The LWT superior on convolution based DWT not only in its lossless transformation[14–16], but also it is less complexity, faster by about twice, and it require less resources to be implemented onto the hardware, either Digital Signal Processing (DSP) chips or Field Programmable Logic Arrays (FPGA)[17]. Moreover, it does not require an auxiliary memory to execute the inverse transform, and also less memory size is needed to store its integer coefficients, compared to the conventional DWT, which demands higher storage capacity to hold its floating numbers coefficients.

In this paper, we propose simple and lossless audio steganography scheme that employ Int2Int LWT and LSBs substitution. The proposed scheme generate adaptive steganography key (stego-key) that is used for encryption the embedding data. This scheme has fixed high embedding capacity (25% from the audio cover size) with excellent perceptual quality. Furthermore, it is full recovery and robust against additive noise and perceptual statistical analysis.

The rest of the paper is organized as follows: section 2 describes the proposed audio steganography scheme. Section 3 presents the scheme evaluation and some experimental results that compared with some of the related works, and ultimately the section 4 concludes this paper.

2. The Proposed Audio Steganography Scheme

In this section, we present a lossless and secure audio steganography scheme that has high capacity with excellent perceptual quality. Although several audio steganography approaches have been developed in the literature, only a few of them characterize by high capacity and imperceptibility beside robustness and security. The steps of the hiding and recovery algorithms are shown in the below sub-sections:

2.1 Hiding Algorithm

An input cover audio signal with resolution of 16 bits/sample is framed to small frames without overlap; each frame has 4 audio samples. Also, an input message signal with resolution of 8 bits/sample is framed into 2 samples/frame. The message signal can be any binary data; in the tests of the approach, images and audio signals have been used as messages. Figure 1 shows the steps of the hiding algorithm, the processing of these steps as follows:

1. Analyze each cover audio frame (4-samples) by using 2-levels of Haar Int2Int LWT. The analysis results are three sub-bands: first detail sub-band has 2 coefficients, second detail sub-band has 1 coefficient, and finally, smooth sub-band has 1 coefficient. The first and second detail sub-band coefficients are used for message data embedding in the embedding process while the smooth coefficient is used for the generation of the dynamic stgo key (S_K) by S_K generator process.
2. In S_K generator, the smooth coefficient is converted to binary with resolution 16 bits/coefficient, then copy only the 12 Most Significant Bits (MSBs) from the coefficient and entered to rotating shift register. Based on the key that is given by Pseudo Random Number Generator (PRNG), the shift register is left rotated. Then, choose only 8 bits from the rotated 12 bits based on Private Position Key (P_K) that is entered by the

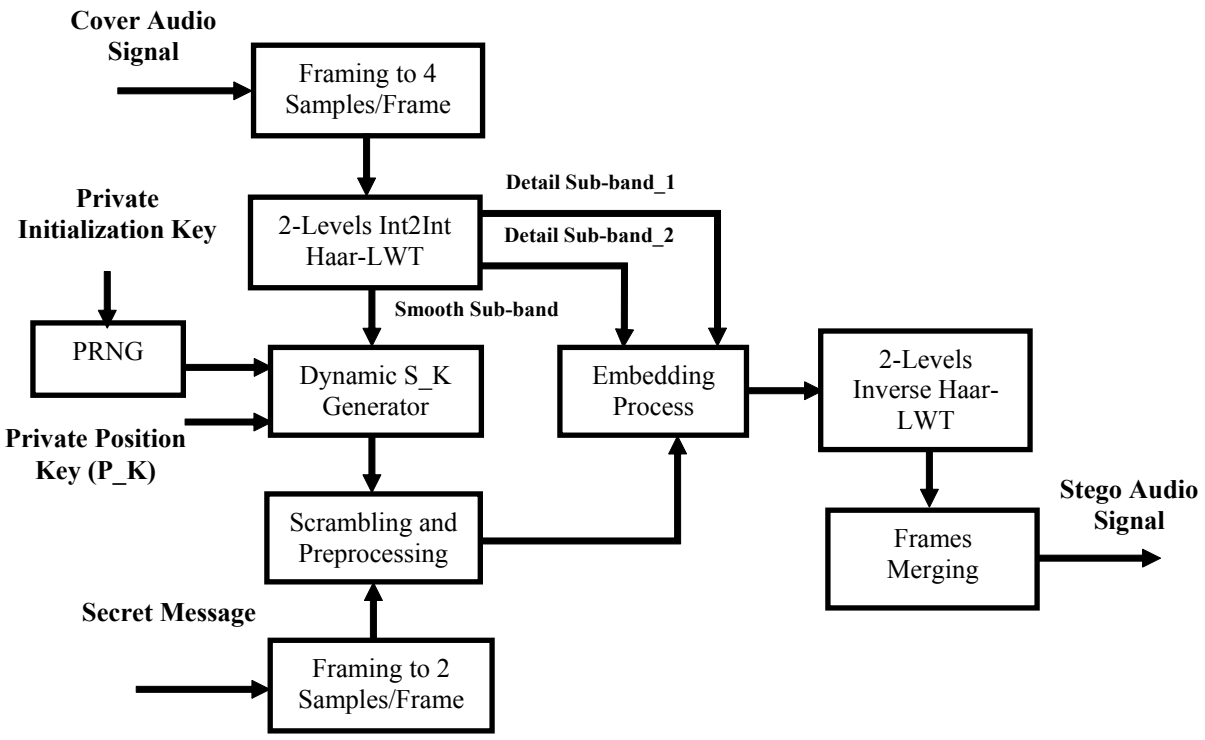


Figure 1. The general block diagram of the proposed Hiding Algorithm.

user of the system. Now the S_K is ready to be used for scrambling the message frame. In S_K generator we use only the 12 MSBs to choose the encryption key in order to increase the S_K robustness against additive noise because higher depth bits has higher power and as a result has higher resistance to additive noise. In this work, two secret keys are used to be entered by the user of the system: the first one is used to initialize the PRNG and the second one is used to select the 8- bits of S_K. These two keys should be shared between transmitter and receiver as secret keys.

3. Subsequently, each sample of the message frame is converted to 8 binary bits and scrambled with binary data of the generated S_K by performing Exclusive OR (XOR) between them. Then, combine the 2 scrambled message samples bits in a one vector that has 16-bits defined as *SM*.
4. In the embedding process, the first 12-bits from the *SM* is inserted in the first 6 LSBs after the Starting Depth (*SD*) from the first and second coefficients of the first detail sub-band, respectively. The reminder 4-bits from the *SM* are inserted in the first 4 LSBs after the *SD* from the coefficients of the second detail sub-band. Now all the 16-bits of *SM* are embedded in the cover detail coefficients. The above *SD* is used to specify the depth of the first insertion position. The purpose of using this parameter is to increase robustness level against additive noise.
5. In reconstruction step, the output stego frame is constructed by performing invers of 2-levels Haar Int2Int LWT after converting all sub-bands to decimal. The above procedure is repeated for all frames. Finally, merging the stego frames together to obtain the output stego-signal.

Illustration Example:

Suppose we have the frame of the cover signal is [12000, 14600, 15000, 18000], the frame of the message is [200 250], *SD* = 1, rotating key = 7, and P_K = [7, 5, 6, 1, 4, 12, 9, 2]. Subsequently, the hiding process corresponding to the proposed technique as follows:

1. Apply 2-levels of Haar Int2Int LWT analysis to obtain: first detail sub-band = [2600, 3000], Second detail sub-band = [3300], Smooth sub-band = [14900].
2. Smooth coefficient = $(14900)_{10} = (0011101000110100)_2$, copy the first 12-bits (101000110100) to the shift register, then achieve left rotation by 7 bits to obtain

(101001010001), now according to the P_K, we choose the 8-bits of the S_K (11010100).

3. Message frame = $[200, 250]_{10} = [11001000, 11111010]_2$, S_K = 11010100, perform XOR between each binary sample of the message frame and S_K to get [00011100, 00101110], finally SM = [0001110000101110].
4. The embedding process is shown in the Table 1.
5. The inverse of 2-levels Haar LWT is applied onto the sub-bans after embedding process to obtain the following output stego signal: [11925, 14577, 15029, and 18069].

For the above example the stego signal closes to the cover signal and the stego signal has excellent imperceptibility (the signal to noise ratio = 48.8695 dB).

2.2 Message Recovery Algorithm

The steps of the recovery algorithm inmost are similar to the hiding steps as shown in Figure 2. In the first step, an input stego audio signal is framed to 4 samples/frame, then; each frame is decomposed by 2-levels of Haar LWT. Subsequently, from the first detail sub-bands coefficients extract the first 12 bits of the *MS* vector (6-bits are extracted from each coefficient), and from the second detail sub-band coefficient extract the reminder 4-bits of *SM* vector. After that, the 16-bits of *SM* vector are separated into two 8-bits samples. After that, exactly same the S_K procedure, which is used in the hiding algorithm, is used to generate S_K, then, perform XOR between each extracted encrypted message sample and S_K, and then, convert the results to the binary to get the secret message frame samples. Repeat the above procedure for all the frames and finally, all the extracted message frames are merging in a single vector to convert to the required message type.

Illustration Example:

In order to demonstrate the steps of the recovery algorithm, here the steps of data retrieval for the same example that is used in the hiding. The received frame in the receiver side is [11925, 14577, 15029, and 18069] is processed as follows:

1. Firstly, this frame is analyzed by using 2-levels of integer Haar LWT to obtain first detail sub-bans = [2652, 3040], second detail sub-band = [3298] and smooth sub-band = [14900].
2. Next, the hidden *SM* binary vector is extracted from the detail sub-bands as shown in the Table 2.

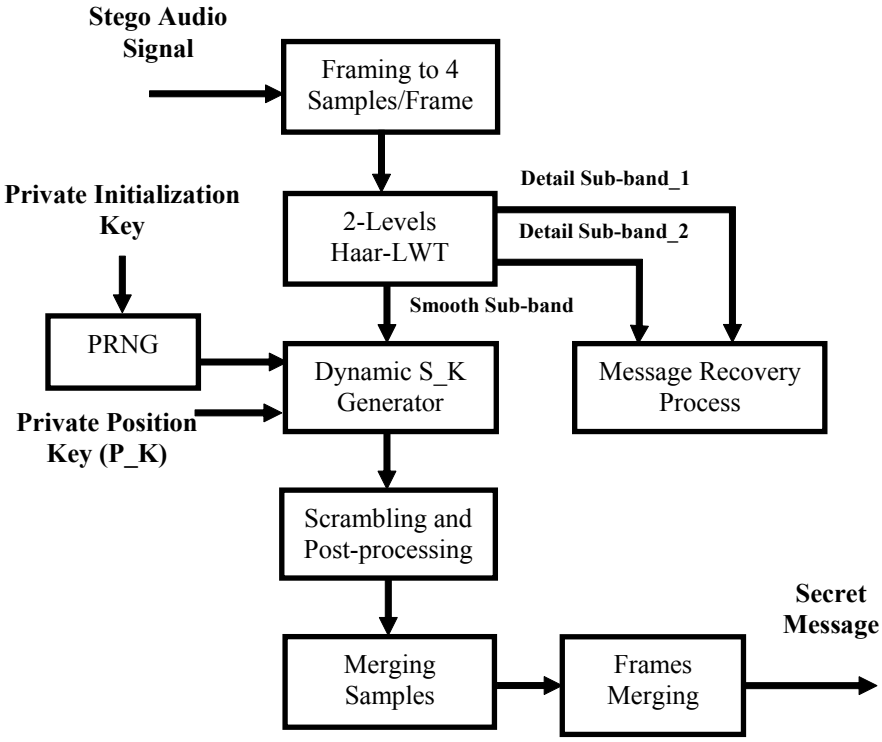


Figure 2. The general block diagram of the Hidden Message Recovery Algorithm.

Table 1. Embedding process steps for the example of the hiding algorithm

The first Coefficient from the first detail sub-band $(2600)_{10}$	0	0	0	0	1	0	1	0	0	0	1	0	1	0	0	0	
First 6-bits From *SM*										1	0	1	1	1	0		
The first Coefficient from the first detail sub-band after embedding $(2652)_{10}$	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0	*SD* =1
The second Coefficient from the first detail sub-band $(3000)_{10}$	0	0	0	0	1	0	1	1	1	0	1	1	1	0	0	0	
Second 6-bits From *SM*										1	1	0	0	0	0		
The second Coefficient from the first detail sub-band after embedding $(3040)_{10}$	0	0	0	0	1	0	1	1	1	1	1	0	0	0	0	0	*SD* =1
The second sub-band coefficient $(3300)_{10}$	0	0	0	0	1	1	0	0	1	1	1	0	0	1	0	0	
Reminder 4-bits From *SM*												0	0	0	1		
The second sub-band coefficient after embedding $(3298)_{10}$	0	0	0	0	1	1	0	0	1	1	1	0	0	0	1	0	*SD* =1

Table 2. The Retrieval of the hidden data steps from the LWT sub-bands

The first Coefficient from the first detail sub-band $(2652)_{10}$	0	0	0	0	1	0	1	0	0	1	0	1	1	1	0	0	SD =1
Extracted first 6-bits of the SM binary vector										1	0	1	1	1	0		
The second Coefficient from the first detail sub-band $(3040)_{10}$	0	0	0	0	1	0	1	1	1	1	1	0	0	0	0	0	SD =1
Extracted second 6-bits of the SM binary vector										1	1	0	0	0	0		
The second sub-band coefficient $(3298)_{10}$	0	0	0	0	1	1	0	0	1	1	1	0	0	0	1	0	SD =1
Extracted final 4-bits of the encrypted *SM*												0	0	0	1		
Final extracted *SM*	0	0	0	1	1	1	0	0	0	0	1	0	1	1	1	0	
Separate extracted *SM* vector into two 8-bits samples						[00011100, 00101110]											

3. Subsequently, the S_K is generated from the smooth coefficient $(14900)_{10}$ = $(0011101000110100)_2$ by the same procedure in the hiding example to obtain S_K = (11010100).
4. Next step is to descramble *SM* by performing XOR between each sample of *SM* and the generated S_K.
5. Finally, the descrambling results are [11001000, 11111010], these results are converted to decimal to obtain the required messages [200, 250].

The above scheme provides full recovery for the retrieval messages without any distortion, also the stego audio data can be saved or transmitted over channels by the same resolution of the cover audio with lossless data retrieval.

3. Results and Discussion

In this section, we present the experimental results of the proposed scheme with discussion; also we compare our results with the results of the methods[10–12]. We have named the methods that we selected for compression based on their title as follows: High Capacity and Inaudible Audio Steganography Scheme (HCIASS)[10], Adaptive Digital Audio Steganography based on Integer Wavelet Transform (ADAS-IWT)[11], and LSB-based Audio Steganography Method by using Lifting Wavelet Transform (LSB-ASM-LWT)[12]. We have achieved the tests for numerous audio cover signals with different sampling frequencies; also we have used images and audio signals as secret messages.

3.1 Tests of Perceptual Quality

The perceptual quality or imperceptibility in audio steganography means inability of human hearing to recognize the difference between stego and cover signals. It can be measured either by hearing to both the cover and stego signals (by several people whose have excellent hearing), or by using mathematical measurements. A good mathematical way to measure a signal quality is by calculating the signal-to-noise power ratio (SNR), by regarding the difference between the cover and stego-signal as a noise, as shown in Eq. (1) [18].

$$SNR = 10\log_{10}\frac{\sum_{i=1}^{N} C^2(i)}{\sum_{ij=1}^{N}[C(i)-C'(i)]^2} \quad (1)$$

where, C and C' are the cover and stego signals, respectively; N represents the number of samples in every one of them.

The proposed scheme has excellent perceptual quality with fixed high embedding capacity that equal to equal to 25% from the size of the audio cover signal. Table 3 shows several tests of perceptual quality in terms of SNR for the proposed approach and the others related selected methods. In these tests, three samples music that exist within the music samples of windows 7 (Kalimba, maid with the flaxen hair, and sleep away) and four recorded speech in different sampling frequencies are used as cover signals. Audio converter has been used to reproduce the music signals in a mono channel and different sampling frequencies for each one as shown in the table. Also, two types of messages have been used: image with size 256*256, and recorded audio with sampling frequency 16 kHz and resolution 8 bits/sample. The duration of cover signals varied depends on the number of message samples, where each 8 bits sample of the message requires two 16 bits samples of the cover. The results in the table show that the perceptual quality is excellent (above 45 dB) for various cover audio signals and for the both message types. In addition, the perceptual quality of the proposed method is very little dependent the chosen cover audio and its sampling frequency. The results in Table 3 show the superior of the proposed method over the other selected methods.

Table 3. Perceptual Transparency Tests for the proposed approach and some related approaches with embedding capacity 25% from the audio cover size

Cover Audio Signal	Sampling Frequency/Hz	Perceptual quality in term of SNR/dB							
		The Proposed Approach		ADAS-IWT		HCIASS		LSB-ASM-LWT	
		In case of audio message	In case of Image message	In case of audio message	In case of Image message	In case of audio message	In case of Image message	In case of audio message	In case of Image message
M-F	22050	47.4701	49.0583	40.9542	41.1816	48.335	47.6539	44.7930	43.4368
M-F	32000	47.7760	49.1456	39.8766	42.7822	49.4069	49.1878	45.2560	43.7385
M-F	44100	49.4207	50.3155	40.4581	40.5317	50.2387	51.8601	45.7654	44.0932
S-a	22050	48.7851	49.1624	41.0172	41.4472	49.8530	49.441	43.4741	41.7741
S-a	32000	48.5576	49.6350	42.8365	43.9347	50.4381	49.6590	44.6548	43.1115
S-a	44100	48.6634	46.8901	42.9593	41.8427	51.5736	50.8487	45.1832	44.9450
Kal	22050	52.0271	51.7362	42.8218	44.3422	49.3933	48.8862	44.8465	43.2854
Kal	32000	51.8778	51.2645	42.768	44.1180	51.5071	50.6296	45.4682	44.1792
Kal	44100	51.8529	50.8292	42.3851	43.6901	53.3633	53.1817	46.6663	44.5815
SP1	44100	53.1346	53.5875	39.8731	40.7293	52.2729	51.9515	41.3419	40.0781
SP2	22050	51.1204	50.8731	38.8810	38.9217	49.7743	49.1089	40.2762	38.9221
SP3	16000	48.1702	48.0634	38.1127	38.4928	48.4888	47.7657	39.8861	38.2349
SP4	11025	47.96712	48.04720	38.4321	38.5582	47.9692	47.0455	39.4541	37.3326
SP5	8000	46.7147	46.7147	37.6164	38.3127	46.7631	45.2694	38.3634	37.09651

(M-f: Maid with the Flaxen Hair (Music), S-a: Sleep Away (Music), Kal: Kalimba (Music), SP: Speaker (Speech), SNR: cover to noise ratio)

3.2 Tests of Robustness against Additive Noise

Robustness means an ability to retrieve a hidden message from a stego signal without or with an acceptable distortion after a stego signal has been affected by factors such as channel additive noise. Similarity between the retrieved and original hidden messages is the usual method being used to measure robustness. The most popular method is Normalized Correlation (*NC)* presently being adopted to measure the similarities. The *NC* formula for one dimensional message signal such as the audio is shown in Eq. (2) [19].

$$NC(M,M') = \frac{\sum_{i=1}^{L} M_i M_i'}{\sqrt{\sum_{i=1}^{L} M_i^2}\sqrt{\sum_{i=1}^{L} M_i'^2}} \quad (2)$$

where, M and M', are the original and recovered secret message, respectively; L indicates number of samples in each one of them. For the two dimensional embedded message such as an image, the *NC* formula is shown in Eq. (3) [15].

$$NC(M,M') = \frac{\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} M_{i,j} M'_{i,j}}{\sqrt{\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} M_{i,j}^2}\sqrt{\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} M'^2}} \quad (3)$$

where, M and M', are the original and recovered secret message, respectively; L_1 is the number of rows and L_2 is the number of pixels in each row for an image matrix.

In the proposed scheme, there are two factors increase the robustness: firstly, using integer to integer LWT transformation and secondly, the embedding insertion depth that is controlled by the parameter of the Starting Depth (*SD*). Table 4 shows the robustness tests against Additive Wight Gaussian Noise (AWGN) in terms of NC between the original and retrieved message. In these tests the cameraman image with size 256*256 pixels has been hided in a recorded speech cover signal with sampling frequency 44100 Hz. The channel noise in terms of Signal to Noise Ratio (SNR) in the table indicates the values of the AWGN noise, small values of SNR means bad communication channel (that has high noise). The table shows the effective of *SD* parameter on the message immunity against AWGN, where with increasing *SD* value the message immunity is increased although there is decreasing in the stego audio quality in terms of SNR, however it still good or acceptable.

Figure 3 shows several robustness tests against AWGN for audio massage that hidden in speech signal by using the proposed scheme with different *SD* values and the selected related methods. The figure shows the superior of the proposed over the other methods.

3.3 Statistical Steganalysis Tests

The proposed scheme is a type of blind steganography that a receiver does not require the original cover signal to detect the hidden message. So, for the steganalyzers also should assume they do not have data base for cover signals and they depend only on the statistical analysis of the signal variations to classify the signal as a cover or stego in case of passive steganalysis or to detect the secret messages in case of active steganalysis. Typically, steganalyzers use intelligent classifiers for the above purposes; accordingly, the core of steganalysis processes is the feature extraction vector of the signal that is usually computed based on statistical analysis. In several researches of audio steganalysis[20,21], the authors employ the variation in the signal histogram that represents the data distribution to calculate the signal features. Some other researchers[22,23] employ the first fourth moment (average (μ), variance (σ2), Skewness (sk), and kurtosis (k)), which yields the nature of function distribution to calculate the signal features.

In this work we have used in our tests both histogram and first fourth moment in time and wavelet domains to measure the robustness of the proposed method against statistical steganalysis. We have used 100 bins in histogram computing for both cover and stego signals and then find the error ratio between them according to Eq. (4).

$$HER = \frac{\sum_{i=1}^{N} (CH_i - SH_i)^2}{\sum_{i=1}^{N} CH_i^2} \quad (4)$$

where, HER is the histogram error ratio, CH and SH is the histogram of the cover and stego signals respectively. For the first fourth moments, we have determined them for cover and stego audio signals according to the equations (5–8) in the time and wavelet domains, then we find the difference ratio percentage according to the Eq. (9):

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (5)$$

$$\sigma^2 = \frac{1}{(N-1)}\sum_{i=1}^{N} (x_i - \mu)^2 \quad (6)$$

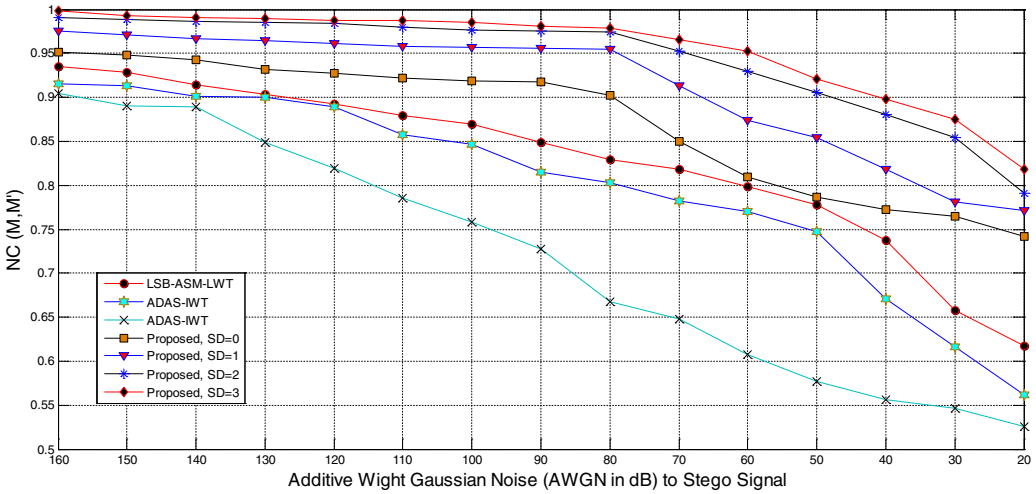
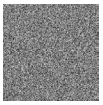


Figure 3. Robustness tests against AWGN for the proposed and selected related methods.

Table 4. Some tests of message immunity against AWGAN for the proposed approach with embedding capacity 25% from the audio cover size

SD / bits	Output quality measured by SNR/dB	Recovered message after adding channel Wight Gaussian Noise /dB					
		Without Noise	CSNR = 100 dB	CSNR = 80 dB	CSNR = 60 dB	CSNR = 40 dB	CSNR = 20 dB
0	53.1634 dB						
		NCC=1.000	NCC=0.9194	NCC= 0.9189	NCC= 0.8316	NCC= 0.7665	NCC= 0.7649
1	47.0448 dB						
		NCC=1.000	NCC= 0.9564	NCC= 0.9559	NCC= 0.8746	NCC= 0.7675	NCC= 0.7669
2	41.1373 dB						
		NCC=1.000	NCC= 0.9748	NCC= 0.9744	NCC= 0.9060	NCC= 0.8414	NCC= 0.7645
3	36.0226						
		NCC=1.000	NCC= 0.9852	NCC= 0.9848	NCC= 0.9431	NCC= 0.8732	NCC= 0.7657
4	29.7964						
		NCC=1.000	NCC= 0.9899	NCC= 0.9897	NCC=0.9618	NCC= 0.9023	NCC=0.8240

$$sk = \frac{1}{(N-1)\sigma^3}\sum_{i=1}^{N}(x_i-\mu)^3 \quad (7)$$

$$k = \frac{1}{(N-1)\sigma^4}\sum_{i=1}^{N}(x_i-\mu)^4 \quad (8)$$

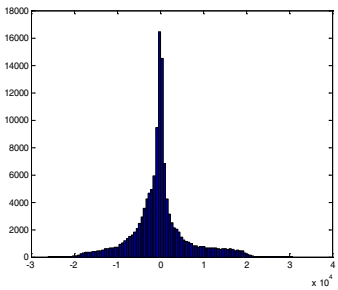
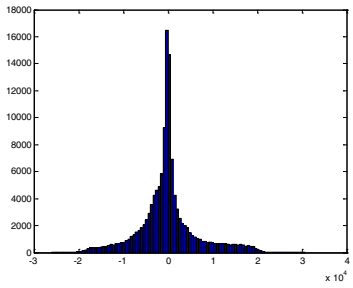
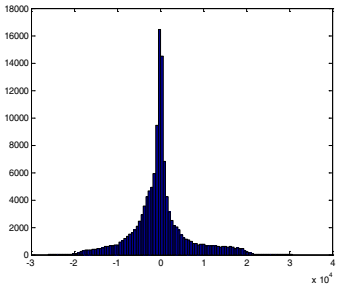
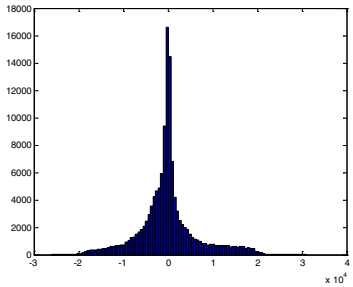
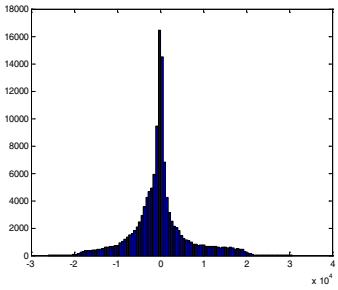
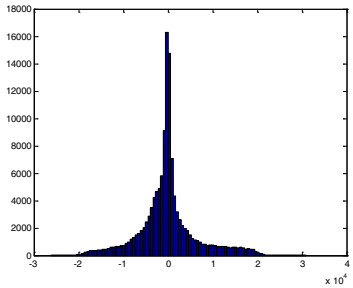
$$DR = 100\times\left|\frac{mfc-mfs}{mfc}\right| \quad (9)$$

where, μ is the average, $\sigma 2$ is the variance, sk is the skewness, and k is the kurtosis, x is the input audio signal to the steganalysis for both tests either in the time or wavelet domain and N is the number of samples in signal x. DR is the ratio of difference for the moment, and mf represent any of the first four statistical moments for either the cover (c) or stego (s) signal.

Table 5 shows the histogram tests for speech signal before and after embedding image message. In this table, the tests have achieved for SD = 0 to 2, in all tests cannot recognize any differences between the two graphs (before and after embedding). Furthermore, the histogram power error ratios are very small because the high correlation between the cover and stego signals.

Table 6 shows the tests of statistical analysis in terms of the difference ratio (DR) of the first fourth moments for the proposed and selected related methods. The tests have been achieved in time and wavelet domains, the small values for DR indicates the output of steganography method (stego signal) has more resistivity against statistical analysis. Table 6 shows the superior of the proposed method over other selected methods for all first fourth moments analysis.

Table 5. Histogram tests for speech signals before and after image message embedding

SD Value	Cover Signal Histogram	Stego Signal Histogram	Histogram Error Ratio *(HER)*
0			1.2274e-04
1			4.0122e-05
2			2.2909e-04

3.4 Computational Complexity Tests

One of the methods that can be used to measure the algorithm computational complexity is the finding of the time that is required for processing. Low processing time required for hiding and recovery denotes one of the substantial features of the proposed approach. The processing time of real time applications such as data embedding through speech conversation between two persons concerns by the processing time that required for each frame (small interval time) and does not concern about the total time required to complete the whole message embedding. In this approach the total time required for hiding or recovery algorithm divides on the number of the cover audio frames to obtain the average approximation processing time for each frame, where each frame has 4 samples. Accordingly, if the computed time less than the frame duration time that depends on the sampling frequency as shown in Table 7, then it can be consider the approach applicable for real time applications. Practically, in real time conversation 8000 Hz sampling rate is the most coding standard used such as CODEC.

In the tests of processing time measurements, we have used a Matlab version (2012-a that operates under 64-bits windows-7 operating system on a PC which has the following specification: Intel (R) Core i5, CPU@ 1.8 GHz, and RAM 4 GB. The tests have been achieved demonstrate that the average frame processing time required for hiding or recovery is nearly constant and does not depend on the input message type or cover sampling frequency as shown in Table 8. Also, both hiding and recovery processing time are nearly same because they are symmetry in their processing. The frame average time in all the tests of Table 8 less than 8 μsec, and this duration time is less

Table 6. Statistical analysis tests: the Difference Ratio (*DR*) of the first fourth moments for the proposed and selected related methods

sth-moment, moment name-Domain	Difference Ratio Percentage (*DR*) for the Measured Moment			
	The Proposed Approach	ADAS-IWT	HCIASS	LSB-ASM-LWT
1st-moment, mean-time	0.2304	1.8731	1.4951	0.1928
1st-moment, mean-wavelet	0.7757	7.3362	8.8140	0.7176
2nd-moment, variance-time	0	$2.0863 * 10^{-4}$	$2.531 * 10^{-4}$	$1.2953 * 10^{-6}$
2nd-moment, variance-wavelet	0	$2.0863 * 10^{-4}$	$2.531 * 10^{-4}$	$1.2953 * 10^{-6}$
3rd-moment, Skewness -time	0.0004	0.4705	0.3916	0.0011
3rd-moment, Skewness -wavelet	0.0083	0.5223	0.4828	0.0172
4th-moment, kurtosis -time	0.0005	0.00736	0.0317	0.00161
4th-moment, kurtosis -wavelet	0.0007	0.04157	0.1409	0.00988

Table 7. Cover frame duration time for various sampling frequency, where each frame has 4 samples

Sampling frequency/Hz	96000	48000	44100	32000	22050	16000	8000
Frame duration time/sec	4.1667e-05	8.3333e-05	9.0703e-05	1.2500e-04	1.8141e-04	2.5000e-04	5.0000e-04

Table 8. Average time required for processing each frame in hiding and recovery of the proposed approach

Cover audio	Frame average processing time for data hiding		Frame average processing time for data recovery	
	Audio message	Image message	Image message	Image message
Speech at 44100 Hz	7.5850e-05	7.7054e-05	7.8835e-05	7.7948e-05
Speech at 32000 Hz	7.5905e-05	7.7519e-05	7.8501e-05	7.7622e-05
Speech at 16000 Hz	7.6214e-05	7.5363e-05	7.7972e-05	7.87750e-05
Speech at 8000 Hz	7.6250e-05	7.7230e-05	7.8391e-05	7.7764e-05

than various duration times of the cover frame that are shown in Table 7 except at sampling frequency 96000 Hz. In terms of the ratio between the obtained frame average processing time and the time duration of the cover frame at the standard 8000 Hz, it will be less than 16%. This means, the proposed approach applicable to use in real time applications even under very slower computer from the used one in these tests.

4. Conclusion

We have presented a new audio steganography scheme based on integer LWT. The proposed approach attains adaptive stego key and full recovery for the hidden messages. The embedding technique insert message data in imperceptible positions of the details sub-bands in integer LWT domain, therefore it provides excellent perceptual quality for high embedding capacity (25% from the size of the cover signal). Furthermore, the integers LWT beside the *SD* parameter that have been used to increase the depth of embedding positions uphold the immunity of the secret messages against AWGN. The experimental results have showed not only excellent perceptual quality and immunity against AWGN for the proposed technique, but also it is robust against statistical steganalysis. Another advantages for the proposed scheme, it is simple and can process each frame of only 4 samples individually; therefore it can be used in the real time steganographic systems. Also, it can be easily implemented on hardware chips. In the near future, we intend to implement the proposed approach on reconfigurable hardware devices such as FPGA.

5. References

1. Cheddad A, Condell J, Curran K, Mc Kevitt P. Digital image steganography: Survey and analysis of current methods. Signal Process. 2010; 90(3):727–52.
2. Nissar A, Mir AH. Classification of steganalysis techniques: a study. Digit Signal Process. 2010. 20(6); 1758–70.
3. Bender W, Gruhl D, Morimoto N, Lu A. Techniques for data hiding. IBM Systems Journal. 1996; 35(3.4):313–36.
4. Wang H, Wang S. Cyber warfare: steganography vs. steganalysis. Comm ACM. 2004; 47(10):76–82.
5. Cvejic N, Seppanen T. (2004). Reduced distortion bit-modification for LSB audio steganography. 2004 7th International Conference on Signal Processing Proceedings (ICSP'04); 2004 Aug 31–Sep 4; Beijing, China. p. 4–7.
6. Geiger R, Yokotani Y, Schuller G. Audio data hiding with high data rates based on Intmdct. IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2006; 2006; Toulouse, France. p. 205–08.
7. Cvejic N, Seppanen T. Channel capacity of high bit rate audio data hiding algorithms in diverse transform domains. IEEE Proceeding of International Symposium on Communications and Information Technologies: Smart Info-Media Systems (ISCIT); 2004; Oct 26–29 p. 84–8.
8. Cvejic N, Seppanen T. A wavelet domain LSB insertion algorithm for high capacity audio steganography. Proceedings of 10th IEEE Digital Signal Processing Workshop and 2nd Signal Processing Education Workshop; 2002; Oct 13–16 p. 53–5.
9. Shahadi HI, Jidin R. High capacity and inaudibility audio steganography scheme. 2011 7th International Conference on Information Assurance and Security IAS. Melaka, Malaysia; 2011 Dec 5–8. p. 104–09.
10. Delforouzi A, Pooyan M. Adaptive digital audio steganography based on integer wavelet transform. Circuits, Systems & Signal Processing. 2008; 27(2):247–59.
11. Pooyan M, Delforouzi A. LSB-based audio steganography method based on lifting wavelet transform. Proc. 7th IEEE International Symposium on Signal Processing and Information Technology (ISSPIT'07); 2007 Dec 15–18; Egypt. p. 600–03.
12. Shirali-Shahreza S, Manzuri-Shalmani MT. Adaptive wavelet domain audio steganography with high capacity and low error rate. 2007 IEEE International Conference on Information and Emerging Technolog (ICIET); 2007; Karachi, Pakistan. p. 1–5.
13. Shahreza S, Shalmani M. High capacity error free wavelet domain speech steganography. IEEE International conference on acoustics, speech, and signal processing conference. on acoustics, speech, and signal processing; 2008 Mar 31–Apr 4; Las Vegas,NV, US. p. 1729–32.
14. Dewitte S, Cornelis J. Lossless integer wavelet transform. IEEE Signal Processing Letters. 1997; 4(6):158–60.
15. Tao Z, Zhao H, Wu J, Gu J, Xu Y, Wu D. A lifting wavelet domain audio watermarking algorithm based on the statistical characteristics of sub-band coefficients. Arch Acoust Q. 2010; 35(4), 481–91.
16. Lei B, Soon IY, Zhou F, Li Z, Lei H. (2012). A robust audio watermarking scheme based on lifting wavelet transform and singular value decomposition. Signal Process, 92(9), 1985–2001.
17. Aziz SM, Pham DM. Efficient parallel architecture for multi-level forward discrete wavelet transform processors. Comput Electr Eng. 2012; 38(5):1325–35.
18. Gordy JD, Bruton LT. Performance evaluation of digital audio watermarking algorithms. Proceedings of the 43rd IEEE Midwest Symposium on Circuits and Systems (Cat. No.CH37144), 2000 Aug 8–11; Lansing, US. p. 456–59.

19. Zhang Q, Huang Y, Deng J, Zheng L. A dither modulation with distortion compensation audio watermarking algorithm based on DWT. J Comput Inform Syst. 2011; 7(1):106–13.
20. Fu J, Qi Y, Yuan J. Wavelet domain audio steganalysis based on statistical moments and PCA. IEEE International Conference on Wavelet Analysis and Pattern Recognition, 2007. ICWAPR. 2007 Nov 2–4; Beijing, China. p. 1619– 23.
21. Qi Y, Fu J, Yuan J. Wavelet domain audio steganalysis based on statistical moments of histogram. Journal of System Simulation. 2008; 20(7):1912–14.
22. Qi Y, Ye L, Liu C. Wavelet domain audio steganalysis for multiplicative embedding model. International Conference on Wavelet Analysis and Pattern Recognition ICWAPR2009; 2009 Jul 12–15; Baoding, China. p. 12–5.
23. Li C, Zeng W, Ai H, Hu R. Steganalysis of spread spectrum hiding based on DWT and GMM. 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing (NSWCTC 2009); 2009 Apr 25–26. p. 240–43.