Hiding Secret Audio Message Inside Smaller Size Audio Cover

By

***Assis. Prof. Ziyad Tariq Mustafa Al-Ta'i and Abd Al-Basset Kadhim**

Diyala University\ College of Sciences\ Computer Science Department

Abstract

Information hiding techniques have recently become important in a number of application areas. Information hiding and cryptography are cousins in the secrecy family. Information can be hidden many different ways in audio. But, most of these ways suffer from the suitable choice of the message (and\or) cover size. Nearly, all hiding techniques can not hide audio message in smaller audio cover. Therefore, this paper presents proposed system that capable from solving this problem. The proposed system depends on two main ideas. First idea is wavelet compression for secret audio message. This idea converts the secret audio message to half size at least. The second idea is masking steganography for the compressed secret audio message. This idea has the ability to hide an audio message in similar size audio cover. These two ideas are applied in transmitter side, in order to obtain the audio stegocover that hide message of double size larger than the cover. However, the reverse of these processes are applied in receiver side, in order to decompress\ extract the secret audio message.

الخلاصة

أصبحت تقنيات إخفاء المعلومات في الوقت الحاضر من الأمور المهمة في العديد من التطبيقات. إخفاء المعلومات و التشفير أولاد عم في عائلة الأمنية. المعلومات يمكن إخفائها بطرق مختلفة في الصوت. ولكن معظم هذه الطرق تعاني من الأختيار المناسب لحجم الرسالة و(أو) حجم الغطاء. تقريباً جميع تقنيات الأخفاء لاتستطيع إخفاء رسالة صوتية داخل غطاء صوتي أصغر. تقدم هذه المقالة نظام مقترح يستطيع حل هذه المشكلة. يعتمد النظام المقترح على فكرتين أساسيتين. الفكرة الأولى هي تقليص (ويف لت) للرسالة الصوتية السرية، وهذه الفكرة تحوّل الرسالة الصوتية الى نصف حجمها على الأقل. الفكرة الثانية هي (ستيكانوكرافي)

التقنيع للرسالة الصوتية المقلصة، وهذه الفكرة لها القابلية لأخفاء الرسالة الصوتية داخل غطاء صوتي بنفس الحجم. تُطبق هاتين الفكرتين في جهة الأرسال من أجل الحصول على (ستيكوكوفر) صوتي يخفي رسالة أكبر حجماً منه مرتين. على كل حال المعكوس لهذه العمليتين يُطبق في جهة الأستقبال من أجل فتح تشفير وأستخلاص الرسالة الصوتية السرية.

Keywords: Information Hiding, Steganography, Wavelet Transform, Masking, Coding (compression).

1. Introduction

This introduction describes the most important topics that are considered as background for this paper.

Audio Coding

In this paper, the terms coding and compression will be used interchangeably. Peter [1] reported that we have seen rapid progress in bit rate compression techniques for speech and audio signals. Linear Prediction, Subband coding, Transform coding, as well as various forms of Vector Quantization and Entropy coding techniques have been used to design efficient coding algorithms. Speech and audio coding are similar in that, quality is based on the properties of human auditory perception. On the other hand, speech can be coded very efficiently because a speech production model is available, whereas nothing similar exists for audio signals. In this paper wavelet transform coding is used, in order to compress audio message size to half size.

Wavelet Transform

For no stationary signals (i.e., signals whose frequency characteristics are time varying like (music, speech, images, etc.)) the Fourier basis is ill-suited because of the poor localization time. The classical solution to this problem is to use the Short-Time (or windowed) Fourier Transform (STFT). However, the STFT has several problems; the most sever being the fixed time-frequency resolution of the basis functions. Wavelet techniques give new class of signal dependent bases that have desired time-frequency resolution properties. Wavelets are well-suited for approximating data with sharp discontinuities [2]. The special significance that the wavelets transform comes from its ability to divide the time-bandwidth product differently at various frequencies or times. The continuous wavelet transform is given by [5]:

$$F(a, b) = \int f(t)\, \Psi((t-b)/a)\, dt \quad \ldots\ldots(1)$$

in this equation, $\Psi(t)$ is the mother wavelet, b represents a time shift, and a is a scaling factor used with (t) time. The mother wavelet is translated or shifted in time producing the wavelets. The wavelet transform is founded on basis functions formed by dilation (spreading a function over a larger domain) and translation of the prototype function $\Psi(t)$. This prototype function is similar to the function STFT, except that the basis functions are high-frequency, short-time pulses, as well as low-frequency long-time pulses, whose contraction in one domain is accompanied by an expansion in the other, with a contrast Root-Mean-Square (RMS) bandwidth to center frequency; that is, it is logarithmic. In contrast, the STFT RMS bandwidth is constant on a linear scale.

The wavelet transform of analog signal f localizes the signal in a time window:

$$[\, b + at - a\Delta\Psi \,,\, b + at + a\Delta\Psi \,]$$

The center of this window is at b+at with the width 2 $a\Delta\Psi$, the wavelet is:

$$\Psi ab(t) = (1/|a|0.5)\, \Psi((t-b)/a) \quad \ldots\ldots(2)$$

The continuous inverse wavelet transform is given by [2]:

$$f(t) = \iint F(a,b)\, \Psi((t-b)/a)\, dbda \quad c \quad \ldots\ldots(3)$$

There are several wavelet forms such as: Haar wavelets, Sinc wavelets, Spline and Battle-Lemrie' wavelets, and others [3].

Audio Steganography

Steganography is the art of hiding information in ways that prevent the detection of hidden messages [4]. Different methods for audio data steganography are proposed such as: Low Bit encoding, Phase coding, Spread Spectrum, and Echo hide [5]. In this paper different method for audio steganography is used by applying masking effect, in order to obtain steganographic method that provides cover size equal to message size.

Masking is very familiar to us in our daily lives, for instance, it accounts for why we can not hear someone whisper when someone else is shouting [6]. In fact, Peter [1] reported that this example is called simultaneous masking. Simultaneous masking is a frequency domain phenomenon where a low-level signal (the maskee) can made inaudible (masked) by simultaneously occurring stronger signal (the masker), if masker and maskee are close enough to each other in frequency. A masking threshold can be measured; low level signals below this threshold will not be audible. Take the example of the masking threshold for the sound pressure level (SPL) =60dB narrow band masker in figure (1): around 1 KHz the four maskees will be masked as long as their individual sound pressure levels are below the masking threshold.

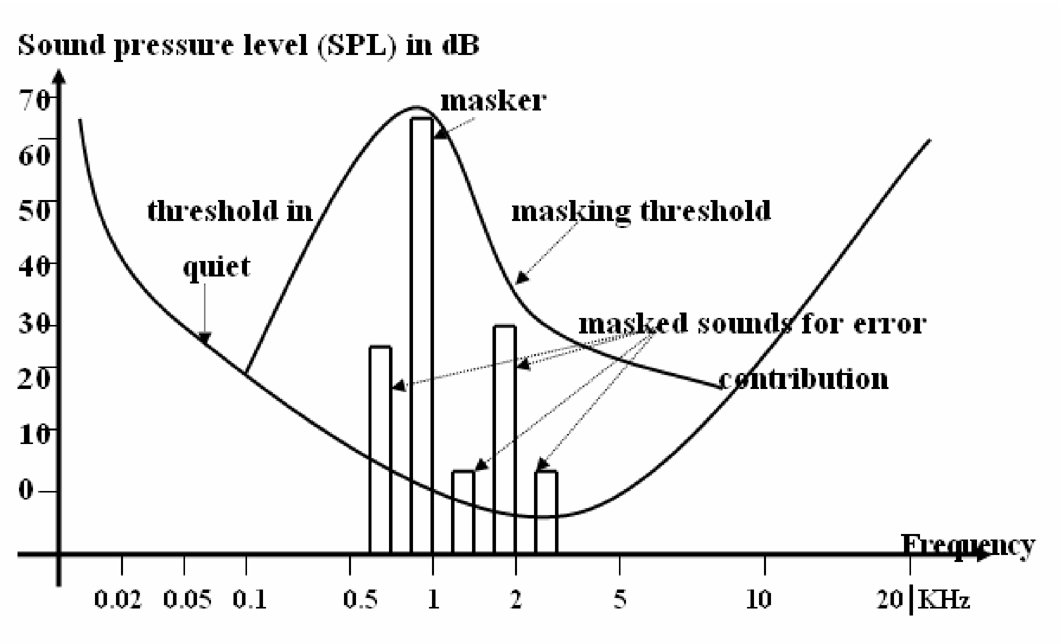
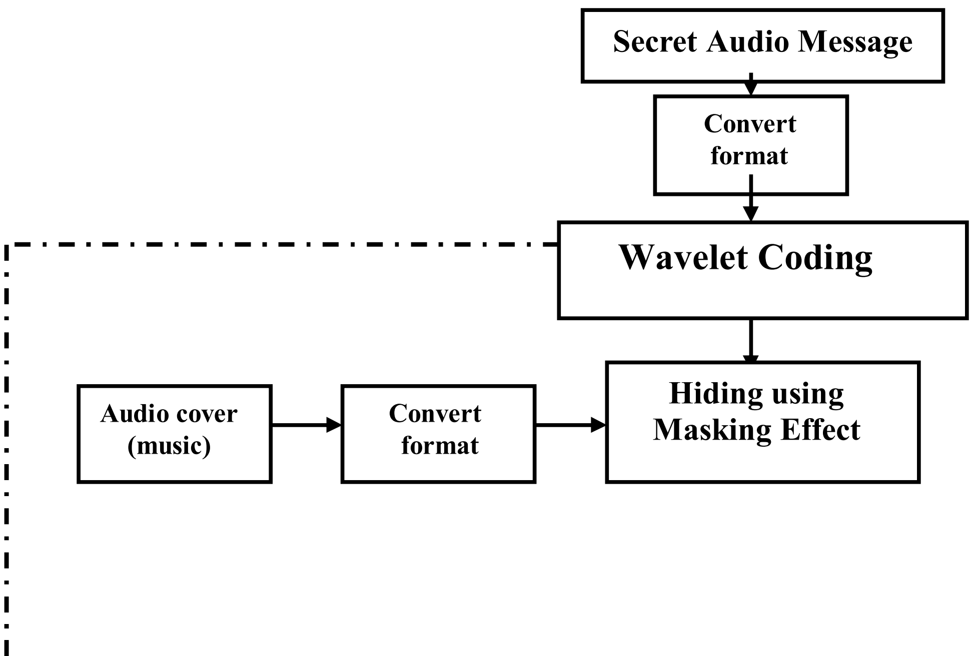


Figure (1) Threshold in quite and masking threshold acoustical events in the closed curve area will not be audible

2. The Proposed System

The details of proposed system can be shown in figure (2). The details of wavelet coding and decoding are shown in figures (3 and 4).



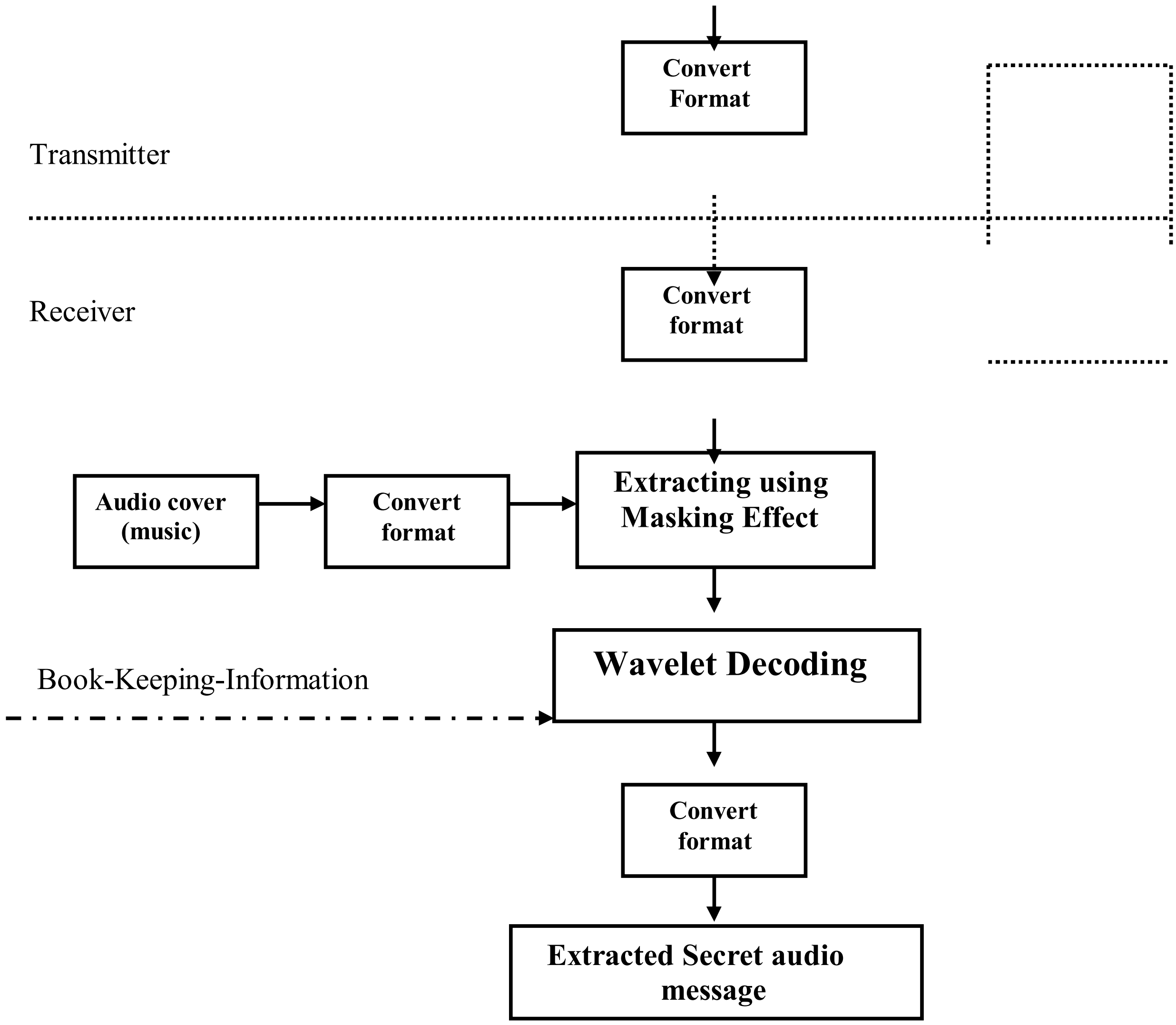


Figure (2) : Block Diagram of The Proposed System

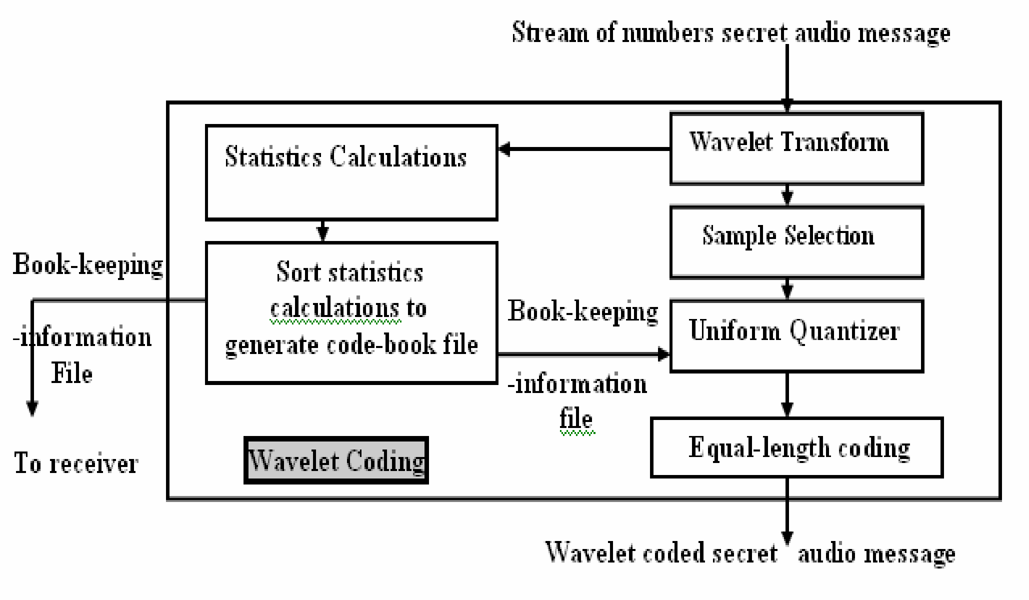


Fig.(3) Wavelet Coding

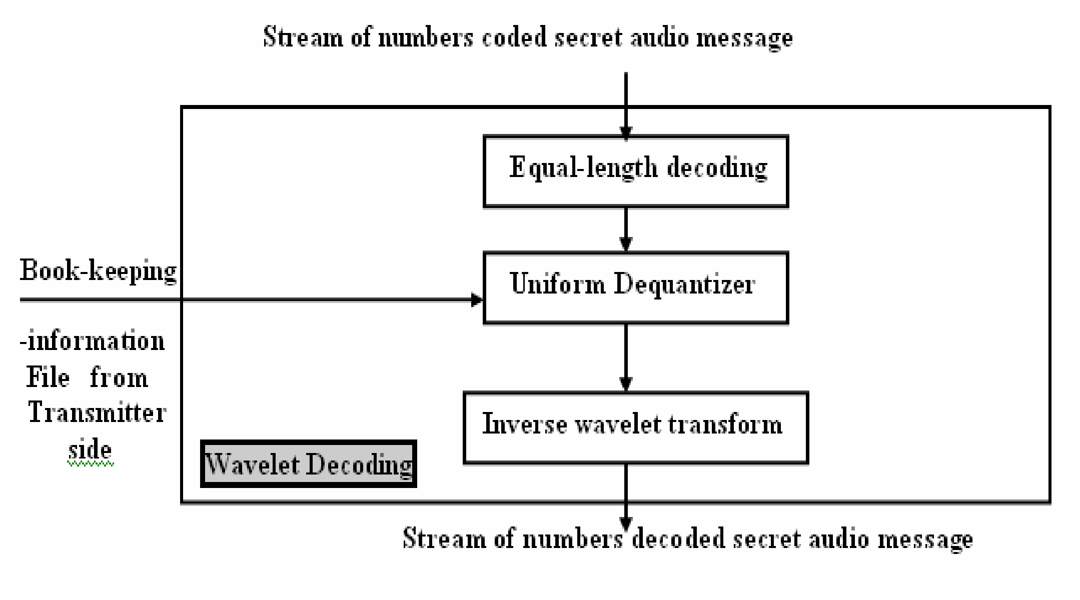


Fig.(4) Wavelet Decoding

First step in transmitter side of proposed system as shown in figure (2) is converting the format of the secret audio message and the format of the audio cover into stream of numbers files. Second step is wavelet coding of the secret message as following:

Wavelet Coding:

First step in wavelet coding as shown in figure (3) is wavelet transformed the stream of numbers secret audio message depending on equation (1). Second step is generating book-keeping-information file. This can be done using some statistical calculations which are collected about each coefficient in the transformed block (one-dimensional window). These calculations are: x-position, mean, variance, required quantizer-bins, and required number of bits. X-position refers to the one dimensional location of the coefficient in the block. Required number of bits and quantizer-bins for each position in the block are generated by the formulas:

$$\text{Bit-number}(i)=\log2\ [J\times\log10\ (\text{variance}(i))\ ] \qquad \ldots\ldots(4)$$

$$\text{Quantizer-bin}(i)=2\text{bit-number}(i) \qquad \ldots\ldots(5)$$

where: J is constant, and (i) is specific position of the block.

The above information are sorted according to the variance of each position in the block. Hence, the book-keeping-information is ready. The book-keeping-information file is sent to the quantizer process and to the receiver side. Third process is sample selection, i.e., how to discard the low energy elements? Because of wavelet transform clusters the high energy elements in first positions of each block, therefore the last 8 elements of each block (block=16 elements) are discarded. The number of discarded coefficients affects directly the compression and mean square error factors. Fourth process is uniform quantizer which simply means using same quantizer with equal spacing between the bins for all coefficients. This type of quantizer is implemented, in this paper using the book-keeping-information, as follows:

Upper-limit=mean+3×(variance)0.5 ……(6)

Lower-limit=mean-3×(variance)0.5 ……(7)

Range=Upper-limit – Lower-limit …… (8)

Bin-location=((selected value-lower-limit)/range)×quantizer-Bin ……(9)

Fifth (last) process is equal length coding which means converting quantized coefficients into bits form. An equal-length coding is a set of code words each of which has the same number of bits. The implementation of an equal-length coding in this work is done by giving each quantized value (not more than 16 because the quantizer-Bin is 16) only 4-bits. Therefore each byte may contain 2 quantized values.

Third step in transmitter side of proposed system is hiding the coded audio message inside the audio cover using masking technique as following.

Audio Hiding using Masking technique:

The masking principle is applied by quieting the secret message in comparison with the audio cover. This can be done using equation (10).

$$\text{Stegocover} = k \times m + (1-k) \times c \quad \text{....(10)}$$

where:

k=masking factor (0.01→1).

m=wavelet coded secret audio message.

c= audio cover.

By applying equation (10) to every byte of the coded message and the cover, with appropriate choice for (k) we can get audio stegocover. The format of the stegocover is reconverted in order to listen to it.

The receiver side must receive the book-keeping-information file and the original audio cover. First step at the receiver side of the proposed system as shown in figure (2) is converting the format of the audio stegocover into stream of numbers file.

Second step is extracting the coded secret audio message using masking technique as the following:

Audio Extracting using Masking technique:

The extracting of the coded secret message is done using Equation (11).

$$m = (stegocover - (1-k) \times c) / k \quad \ldots\ldots (11)$$

By applying equation (11) to every byte of the stegocover and the original cover with same appropriate choice of (k), we can extract the coded secret audio message.

Third step in receiver side is wavelet decoded of the converted stegocover as the following.

Wavelet Decoding:

First step in wavelet decoding as shown in figure (4) is equal-length decoding. The reverse of equal-length coding process is followed exactly in decoding coded data. Second step is uniform dequantized decoded data. The reverse of the uniform quantization process is followed exactly in dequantizing the decoded data. The quantizer/dequantizer process introduces an error that can not be recovered, which affects the quality of the extracted secret audio message. Third (last) step is inverse wavelet transformed the dequantized data depending on equation (3). This makes the (stream of numbers decoded secret audio message) ready.

Last process in receiver side is reconverting the format of the stream of numbers extracted audio message in order to listen to the secret message.

3. Results

The results are intended by applying the proposed algorithm with Windows98 Audio-Video (WAV) format. The Secret message and the cover are recorded using simple (available) microphone of (Pentium III) PC, with the following attributes: 8-bit, sampling rate (11.025 KHz), stereo. The transmission environment is considered as a digital end to end. The results are calculated with (sample) speech message of length (7 sec) and (sample) music cover of length (76 sec). The required time for hiding process (transmitter side) is (2.6 min). The required time for extracting process (receiver side) is (2 min). These results are shown below.

A-Fidelity of stegocover:

A comparison between original cover and stegocover waveforms can be shown in figure (5). This figure verifies the following calculations:

Root mean square error% =0.0035.

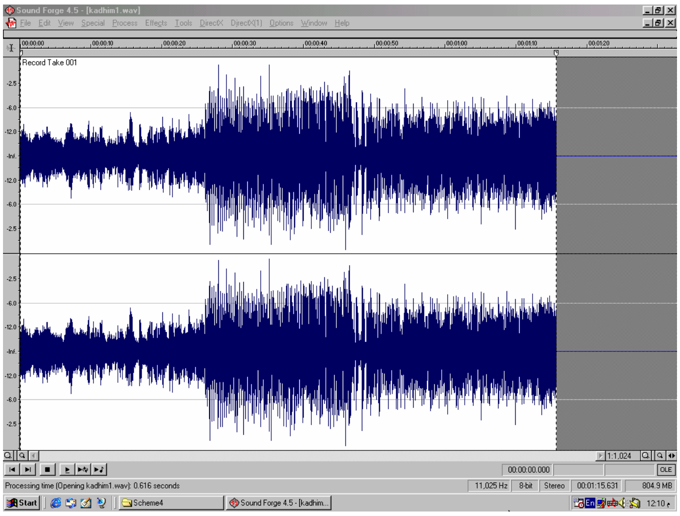
Mean square signal to noise ratio in (dB)=47.103.

B-Fidelity of extracted message:

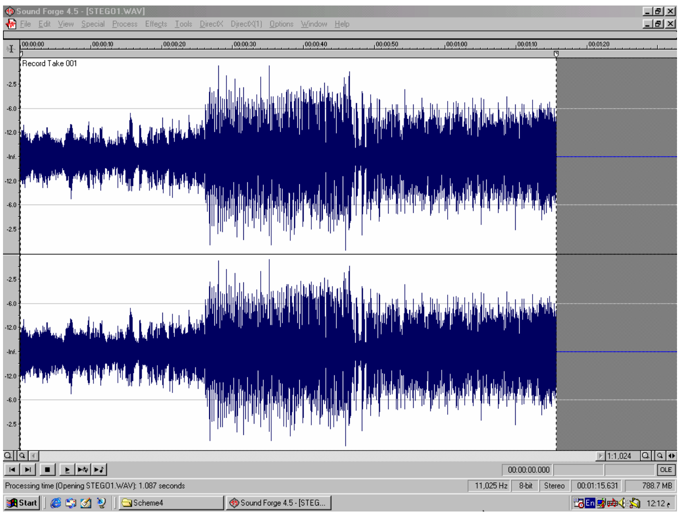
A comparison between original and extracted message waveforms can be shown in figure (6). This figure verifies the following calculations:

Root mean square error %=1.417.

Mean square signal to noise ratio in (dB)=16.1099.



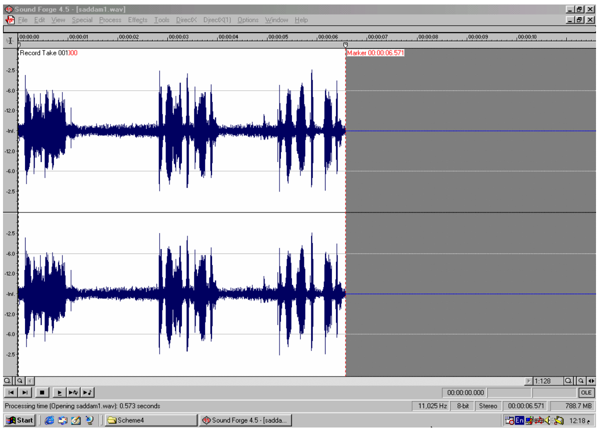
(a)



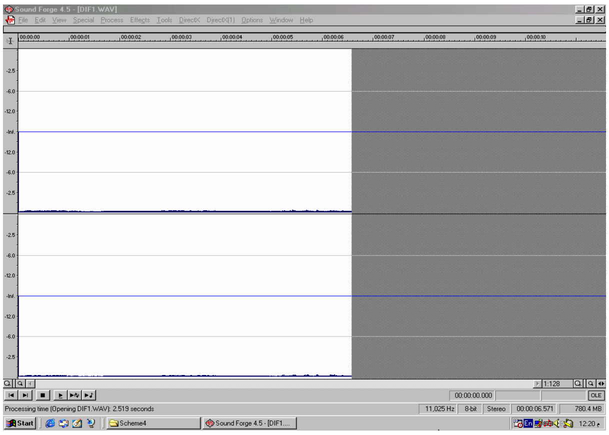
(b)

Fig.(5) Comparison among Left-Right waveforms of

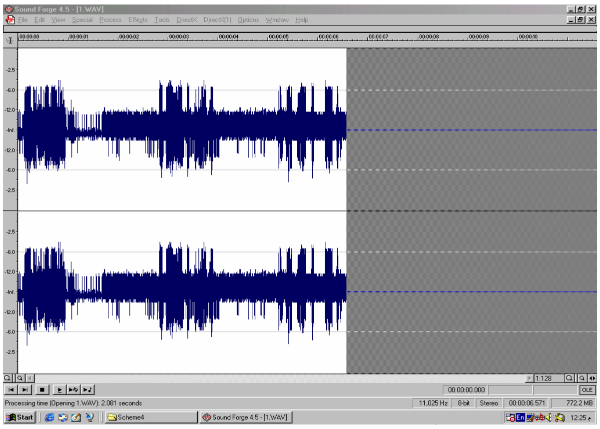
(a) Original Music Cover (b) Stegocover



(a)



(b)



(c)

Figure (6) Comparison among Left-Right waveforms of

(a) Original Secret Speech message (b) Hided coded secret audio message in stegocover

(c) Extracted Secret Speech message

4. Conclusions

Calculated results showed that proposed system is very good hiding method with ability to hide audio message inside half smaller size audio cover, at least. However, some conclusions can be inferred:

a- In spite of quantization error, the proposed coding technique that is included in the proposed system has no noticeable effect on overall performance of the system.

b- One of the points that make the proposed system verifies its aim is wavelet coding, which makes the secret audio message half of its original size. However, wavelet coding is affected by: (transform type, quantizer type, number of discarded coefficients, and coder type).

c- One of the points that make the proposed system verifies its aim is the masking technique, because it is wideband method, i.e., the length of secret audio message can be equal to the length of audio cover.

d- The secret message and the cover may be speech, music, or any recorded voices.

e- Masking factor (k) is an important factor in hiding and extracting processes. (k) is considered as a control factor, therefore it must be selected with a suitable choice.

f- Weak point in proposed system is that the original audio cover must be sent to receiver side. This is because, sending two audio songs to the same side may lead to some suspicious.

References

[1] Peter Noll , " Wideband Audio " , proceedings of the NATO advanced study Institute on signal processing for multimedia, Jim Byrnes Ed., IOS Press, Lucca, Italy, 5-8 July, 1999.

[2] Burrus C. Sidney , Ramesh A. Gopinath , and Haitao Guo, " Introduction to Wavelets and Wavelet Transforms - A primer " , Book published by Prentice-Hall, Inc., 1998.

[3] Don Morgan, " Numerical Methods for DSP Systems in C ", Book from John Wiley and Sons, Inc., 2001.

[4] Neil F. Johnson and Sushil Jajodia, "Exploring Steganography: Seeing the Unseen " , IEEE Computer Society, 0018-9162/98©, pp26-34, Feb., 1998.

[5] Bender W. , Gruhl D. , Morimoto N., and Lu A. ,"Techniques for data hiding " , IBM systems Journal , Vol.35 No.3&4, MIT Media Lab, 1996.

[6] Richard Brice , " Multimedia & Virtual Reality Engineering " , Newnes Press , Linacre House , Jordan Hill, Oxford, 2002.