Evaluation of Wavelet Transform Audio Hiding

By
Dr. Thaher Abd Al-Hadi , Dr. Ziyad Tariq Mustafa Al-Ta'i
Diyala University / College of Science / Computer Science Department

Abstract

Audio hiding is a method for embedding information into an audio signal. It seeks to do so in a robust fashion, while not perceivably degrading the host signal (cover audio). Hiding data in audio signals presents a variety of challenges; due in part to the wider dynamic and differential range of the Human Auditory System (HAS) as compared to other senses. Transforms are usually used for robust audio hiding (audio watermarking). But, the audio hiding process is affected by the type of transform used. Therefore, this paper presents an evaluation of wavelet transform hiding in comparison with selected types of transforms: (walsh transform and cosine transform) hiding.

In order to generate the audio stegocover, this paper concludes (wavelet, walsh, or cosine) transform of the audio cover, replacing some transformed cover coefficients with secret audio message coefficients, and inverse (wavelet, walsh, or cosine) transform for audio cover with replaced coefficients. While, the extracting method concludes (wavelet, walsh, or cosine) transform of the stegocover and extracting the secret audio message. The generated stegocovers and the extracted audio messages are used to create the proposed evaluation.
Keywords: Information hiding, Audio hiding, Wavelet transform, Walsh transform, Cosine transform.

1. Introduction

Tuomas [1] called information hiding as Invisible Communication. Laurunce et al. [2] revealed that audio data hiding techniques are divided into two categories: Audio steganography:

Audio steganography refers to the techniques that utilize the existence of redundant information in a communication process. Digital sound naturally contains such redundancies in the form of a noise component. The most important requirement in audio steganography is that the presence of the hidden message be undetectable. For audio data the stego data must look like a "typical audio" [3]. The general methods of audio steganography are: (low bit encoding, phase encoding, echo hiding, and spread spectrum) [4].

Audio Copyright Protection:

Todd et al. [5] reported that audio copyright protection may be either content-based, or it can be accomplished through watermarking. Watermarking is an application which embeds the least amount of data, but requires the greatest robustness because the watermark is required for copyright protection. Watermark robustness is enabled using: (i) redundant spread-spectrum (ii) psycho-acoustic frequency masking and (iii) transform techniques. Digital watermarking is a technology which potentially can be used to enforce the copyrights and integrity of digital multimedia data [6]. However, audio watermarking is useful as a general audio steganography tool.
Selected Transform Techniques
Wavelet Transform

The wavelet transform converts a data input sequence of a given length to a sequence of real numbers of equal length in which the vertical size of the wavelets changes at each of the set horizontal positions (and scales) so that the additions of all the wavelets reproduces the original [7]. 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 [8]:

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

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 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 $2a\Delta\psi$, 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 [5]:

$$f(t) = \int\int F(a,b) \Psi((t-b)/a) \, db \, da \quad \ldots\ldots(3)$$

There are several wavelet forms such as: Haar wavelets, Sinc wavelets, Spline and Battle-Lemrie' wavelets, and others [8].

Walsh Transform

Beauchamp [9] reported that Walsh functions form an ordered set of rectangular waveforms taking only two amplitude values, +1 and -1, defined over a limited time interval.

The one-dimensional discrete Walsh transform of a function $f(x)$, denoted by $F(u)$, is given by:

$$F(u) = (1/N) \times \sum_{x=0}^{N-1} f(x) \times \prod_{i=0}^{n-1} (-1)^{b_i(x) \times \text{bn}-1-i(u)} \quad \ldots\ldots(4)$$

The inverse Walsh transform is given by:

$$f(x) = \sum_{u=0}^{N-1} F(u) \times \prod_{i=0}^{n-1} (-1)^{b_i(x) \times \text{bn}-1-i(u)} \quad \ldots\ldots(5)$$

Cosine Transform

Kramer and Mathews [10] mentioned that the Fourier series representation of a continuous real symmetric function contains only real coefficients corresponding to the cosine terms of the series. This condition can be realized with an image or audio field and a compact discrete cosine transform is obtained. The one-dimensional discrete cosine transform pair is given by the expressions:

$$F(0) = (1/N) \times \sum_{x=0}^{N-1} f(x) \quad \ldots\ldots(6)$$
\[ F(u) = \left(\frac{1}{2N^3}\right) \times \sum_{X=0}^{u} f(x) \times \cos(2x + 1)u \] 

\[ f(x) = \left(\frac{1}{N}\right) F(0) + \left(\frac{1}{2N^3}\right) \sum_{u=1}^{N-1} F(u) \times \cos(2x + 1)u \] 

2. The Proposed Algorithm

The proposed algorithm can be shown in figure (1).
Figure (1) Block Diagram of Proposed Algorithm

The main steps of the proposed algorithm in transmitter side are:

Forward Transform:

This technique is based on forward (normalized wavelet, walsh, or cosine) transform, which is applied to the stream of numbers audio cover. The theoretical background of the selected types of transforms is described in the introduction. the implementation algorithms of these transforms are depend on equations: (1, 4, 6, and 7).

Replacing some Transformed Cover Coefficients with Secret message Coefficients:

After employing forward (wavelet, walsh, or cosine) transform on cover file, the high energy elements are clustered at certain positions with each transformed window (block). Therefore, some coefficients with low energy can be discarded from each block without distorting the reconstructed stegocover. The principle idea of this technique is done by discarding low energy coefficients using Zonal Sampling method [11], which depends on discarding the elements that have small variances and keeping the elements that have large variances. The number of discarded coefficients depends on the type of transform used and the discarding method. The number of discarded coefficients affects directly the mean square error factor for the reconstructed stegocover. The discarded coefficients in each transformed block are replaced with scaled secret message coefficients sequentially. The discarding and replacement processes are continued until secret message file is finished. Otherwise, the transformed blocks are kept as they are.
One good feature of wavelet transform, it clusters high energy coefficients in first positions of the transformed block and the low energy coefficients at the end positions. This feature makes this technique easier to implement.

Inverse Transform:

This technique is based on inverse (normalized wavelet, walsh, and cosine) transform, which is applied to the (wavelet, walsh, or cosine) transformed audio cover with scaled message coefficients. The implementation algorithms are depend on equations: (3, 5, or 8).

After these three techniques, the audio stegocover is prepared after converting the format.

At receiver side, the main processes are (wavelet, walsh, or cosine) transform of the audio stegocover and extracting process. Wavelet, walsh, or cosine transform is done exactly as described before. The extracting process is done according to specific positions in transform domain. Now the secret audio message is extracted, and can be listen to after converting the format.

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 II) 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 evaluation of wavelet transform in comparison with walsh and cosine transforms is shown in table (1). The length and size of
secret speech message are (7 sec, 141 KB). The length and size of audio cover are (71 sec, 1.58 MB).

<table>
<thead>
<tr>
<th>Transfo-rm Type</th>
<th>Cryptographic Covering Stegocover</th>
<th>Time</th>
<th>Cryptographic Extracting Extracted Message</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelet</td>
<td>e_{rms} % 45.1633 SNR_{ms} (dB)</td>
<td>2.5</td>
<td>e_{rms} % 1.1902 SNR_{ms} (dB)</td>
<td>1.2</td>
</tr>
<tr>
<td>Walsh</td>
<td>0.0075 40.7761</td>
<td>2.8</td>
<td>0.6842 19.5736</td>
<td>1.4</td>
</tr>
<tr>
<td>Cosine</td>
<td>0.0118 36.794</td>
<td>2.8</td>
<td>2.168 10.161</td>
<td>1.4</td>
</tr>
</tbody>
</table>

The results of table (1) are calculated with different scaling factors: (S=(50→100) for wavelet transform, and S=(100→200) for walsh or cosine transform).

The results of proposed algorithm using wavelet transform are shown in figure (1) and figure (2). The results of proposed algorithm using walsh transform are shown in figure (3) and figure (4). The results of proposed algorithm using cosine transform are shown in figure (5) and figure (6).
Figure (1) Wavelet Transform Proposed Algorithm Comparison among Left-Right Waveforms of (a) Original Music Cover (b) Stegocover
Figure (2) Wavelet Transform Proposed Algorithm Comparison among Left-Right waveforms of
(a) Original Secret Speech message (b) Hided Message in stegocover (c) Extracted Secret Speech message

Figure (3) Walsh Transform Proposed Algorithm Comparison among Left-Right Waveforms of (a) Original Music Cover (b) Stegocover
Figure (4) Walsh Transform Proposed Algorithm Comparison among Left-Right waveforms of
(a) Original Secret Speech message (b) Hided Message in stegocover (c) Extracted Secret Speech message
Figure (5) Cosine Transform Proposed Algorithm Comparison among Left-Right Waveforms of (a) Original Music Cover (b) Stegocover
Figure (6) Cosine Transform Proposed Algorithm Comparison among Left-Right waveforms of
(a) Original Secret Speech message (b) Hided Message in stegocover (c) Extracted Secret Speech message
Table (2) shows the effect of an important scaling factor on wavelet transform audio hiding.

<table>
<thead>
<tr>
<th>Scaling factor (S)</th>
<th>Cryptographic Covering</th>
<th>Cryptographic Extracting</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Stegocover1</td>
<td>Stegocover2</td>
</tr>
<tr>
<td></td>
<td>( e_{\text{rms}} )</td>
<td>SNR(_{\text{rms}} )</td>
</tr>
<tr>
<td>S=10</td>
<td>0.0234</td>
<td>30.846</td>
</tr>
<tr>
<td>S=60</td>
<td>0.0045</td>
<td>45.208</td>
</tr>
<tr>
<td>S=100</td>
<td>0.0033</td>
<td>47.836</td>
</tr>
<tr>
<td>S=1000</td>
<td>0.0026</td>
<td>49.936</td>
</tr>
</tbody>
</table>

The results of table (2) can be shown in figures (7 and 8).
Figure (7) Wavelet Transform Audio Hiding Results showing the Effect of Scaling Factor (S) between Left-Right Waveforms of (a) Original Music Cover and Stego Cover when (b) S=10
Figure (8) Wavelet Transform Audio Hiding Results showing the Effect of Scaling Factor (S) between Left-Right Waveforms of
(a) Original Secret Speech message and Extracted message when
(b) $S=10$ (c) $S=60$ (d) $S=100$ (e) $S=1000$.

4. Evaluation and Conclusions

Calculated results showed that the proposed algorithm is a good method for hiding audio in audio. The evaluation of the proposed algorithm using (wavelet, walsh, and cosine) transform can be deduced from table (1):

1. Wavelet transform is better than the other transforms for the proposed algorithm if the comparison factor is the time required.
2. Wavelet transform is better than the other transforms for the proposed algorithm if the comparison factor is the quality of the stegocover.
3. Walsh transform is better than the other transforms for the proposed algorithm if the comparison factor is the quality of the extracted message.

Some conclusions can be inferred from the proposed algorithm:

a- From the above evaluation we can deduce that wavelet transform is a suitable transform type for hiding audio in audio in the proposed algorithm.

b- The quality of the stegocover is very good (which is the goal). However, the quality of the extracted message is not good enough but it is understandable.

c- The proposed algorithm is wideband method. I.e., the length of secret message can be equal to (0.25) the length of cover. This ratio is depended on number of the replaced coefficients.
d- The secret message may be speech, music, or any recorded audio.

e- Scaling factor is an important factor in hiding and extracting processes. This factor is considered as a control factor; therefore it must be selected with a suitable choice. This factor is affected by the type of transform used, therefore it is (50→100) for wavelet transform and (100→200) for walsh and cosine transforms.

f- The number of replaced coefficients affects directly the quality of the stegocover and the extracted message. Therefore, this number must be selected with suitable choice (the balance between the bandwidth and quality is required). This number is affected directly by the type of transform used.
References


