Hiding Coded Speech in Color Images using Zaslavsky-Map and AMBTC Technique

Suhad Fakhri Hussein

1Ministry of Education, Al-resaffa 1 education, Baghdad, Iraq. Email: Suhad7242@gmail.com

1. Introduction

Image-based data hiding involves embedding secret data inside the image without affecting image quality (cover medium) [1]. The Secret digital data can be text, image, video, or audio, while the cover medium that hold secret data is multimedia file [2]. Different methods for hiding information depend on the characteristics of embedding secret data, and some depend on substitution methods [3]. In contrast, others depend on modifying the cover medium's least significant bit (LSB) and may use a mapping function for embedding and transformation methods [4]. Some methods are based on modification or reordering of the coefficients of the transform domain, such as DWT, DCT, or DFT, using some set of rules. Other miscellaneous methods depend upon techniques like matrix decomposition, fractals, coding, and quantization [5]. In various information technology and communications (ITC), all information hiding methods are used, some of them being used as original or fake in communications like intelligence applications, authentication records, and military. Besides ownership defining and demonstrating for commercial utilizing the watermarking algorithms, the consumer traces user IDs embedding the product in such applications before selling them [6]. The AMBTC method [7] used two levels of thresholding a method of representation image block called bitmap. The method is compatible for real-time application because of its
complexity is low. The quantization level modification and bitmap representation hide secret information about the proposed method with DWT.

2. Related works
There are some works related to proposed method are illustrates as follow:
• In [8] used a pseudo-random generator using Zaslavsky chaotic map is used for image encryption that follows the permutation process and substitution process.
• In [9], used Genetic Algorithm and Particle Swarm Optimization for selecting the best place in coefficients for modification in hiding process. The best placed that caused minimum effect on the cover medium and the partition blocks is classified as smooth or complex texture.
• In [10] proposed a method of encryption based on DLCT "discrete linear chirp transform" that is used to build a method of cipher algorithm.
• In [11] proposed a method of hiding process using AMBTC technique with lossless coding. The bitmap of partition is merge into a decimal number that recognize of embedding process. the Huffman code is used to avoided the redundancy of each block.
• In [12] presents a method of data hiding using image partitioning into specific block sizes. A Hessenberg transform is applied on each block one value is modified to hold the secret data.

3. Proposed Method
Hiding secret speech in image is proposed in this work for transmitting via a communication channel by compressed the secret into bits and the cover image is codded using AMBTC technique. The main diagram of proposal shown in Figure (1). The proposed method consists of hiding secret speech hiding, the embedding method Quantization level Modification. Speech hiding may be used as a password in the network.

3.1 Speech Hiding
The cover image decomposed into its three-color bands red, green, and blue then a two-dimensional DWT applied on each color band to split high values approximation coefficients from the low values detail coefficients. The AMBTC embedding method, shown in Figure (2), is used. The resulting image contains the speech that has been hidden inside the image.
3.2 Split Image to RGB
The split image into its color band is a preprocessing of cover medium in proposed method as explain in Figure (3). Each of them considered as gray image that will used in transformation process.

![Split color image into three color bands (RGB).](image)

3.3 Discrete Wavelet Transforms
Two-Dimensional DWT is used to transform each color image to decompose for regions as shown in Figure (4). All regions are used in embedding process except the first one that have an important detail of image.
The LL coefficients of each RGB color band in DWT contain image approximation coefficients to be concatenated together without embedding any information inside. This action will keep the quality of the image after reconstruction. The LL coefficients are reshaped in a one-dimension vector. Figure (5) shows the concatenation of LL coefficients of DWT.

The HL, LH, and HH coefficients of all RGB colors are concatenated in matrix size (3×3), with the secret information embedded inside, as shown in Figure (6).
3.4 Blocks partition
The matrix resulting from the previous step will be used for hiding after partitioning in a predefined size. The cover medium size is (512×512), while the size of coefficients after applying DWT is (256×256) (LL, HL, LH, and HH). All coefficients except (LL) are collected in a single matrix, as explained before. The matrix size is (768×768) and is partitioned into blocks, each of which is (4×4). Therefore, the total number of blocks is (36864), as shown in Figure (7).

![Fig. 7- Block partition of the matrix (for hiding).](image)

3.5 Block coding
The block coding process represented by m×m size has been used for hiding secret process. The coding process is represented by finding the mean value of block value which used as threshold. All values greater than threshold are converting to a value (a) which is mean of all values lower than threshold, while all values greater than threshold is replaced with the second mean (b) the bitmap (B) is indicate the zeros and ones of a and b respectively, finally the trio (a, b, and B) replace each block values. The decoding of trio (a, b, B) is replaced the value of a and b with the values of ones and zeros in bitmap B as explain in figure 8.

![Fig. 8 - Absolute moment block truncation coding.](image)

3.6 Conversion Process
The secret speech is compressed using a method containing DCT transform with quantization, and the results are coded in binary form. The embedding process will take the non-uniform length of the binary sequence and embed it in the covered band depending on some threshold explained in the embedding stage.

3.7. Zaslavsky map
The bits represented the secret speech are encrypted by applying exclusive-or operation with a bit generated with Zaslavsky map that is a two-dimensional chaotic map that applied equation (1 and 2). The method generates pseudo real random numbers.

\[ x_{n+1} = x_n + v(1 + \mu y_n) + \varepsilon v u \cos(2\pi x_n) \mod 1 \quad (1) \]

\[ y_{n+1} = e^{-f}[y_n + \varepsilon \cos(2\pi x_n)] \quad (2) \]
here $\mu = \frac{1-e^{-t}}{T}$, $x_n, y_n$ are variables of the chaotic map. The initials parameters $x_0$ and $y_0$. The constants $\epsilon, \tau, \text{ and } \nu$ are controlling parameters.

### 3.8 DCT Compression

The input speech is a spoken word, its size should be reduced by a compression technique that is utilized by DCT for transformation the data vector, then finds the best threshold that balance between speech quality and compression ratio. The absolute coefficients of DCT are sorted in descending order then find the Norm value of these coefficients values in increment order until reaching the specific ratio of the selected Norm with respect to total Norm. The speech signal before and after compression is shown in Figure (9).

![Speech signal before and after compression.](image)

### 3.9 Embedded method

The method for hiding secret bits is represented by replacing the bitmap B with this secret in selected block which it used for extraction in destination. After embedding all secret bits, a decoding is applied for all blocks and reconstruction the stego-image.

### 4. Experimental Results

The proposed method requirements are hardware and software; the hardware requirement consists of multiple computers; the Software requirements are MATLAB programs.

#### 4.1. dataset

A standard image dataset, as shown in Figure (11), is used for analysis the efficiency of proposed method which are car, Lenna, Woman, paper, and House.

![Image dataset.](image)

#### 4.2. Stego-Image Quality Test

The proposed method is applied in three ways, and the quality of stego-image in all are tested by finding the distortion in images, this test is illustrated as follows:

##### a. Speech Hiding

Speech hiding (word) in the cover image is the proposal that hides compressed speech using DCT transform in the coded image that hides 16 bits (equal to bitmap) in each block to be extracted in the receiver side with decompression to reconstruct the original speech. The quality of the Stego-image is listed in Table (1) for hiding the speech of five spoken words.
Table 1- Image quality measurement of different spoken word

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Woman</th>
<th>Car</th>
<th>Lena</th>
<th>Peppers</th>
<th>House</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1</strong></td>
<td>MSE</td>
<td>8.1254</td>
<td>37.2672</td>
<td>18.9575</td>
<td>21.6698</td>
<td>30.9512</td>
</tr>
<tr>
<td></td>
<td>PSNR</td>
<td>35.6254</td>
<td>31.6619</td>
<td>37.1297</td>
<td>35.9896</td>
<td>32.6685</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.9125</td>
<td>0.9521</td>
<td>0.9485</td>
<td>0.9785</td>
<td>0.9816</td>
</tr>
<tr>
<td><strong>2</strong></td>
<td>MSE</td>
<td>11.5252</td>
<td>29.2570</td>
<td>21.0959</td>
<td>22.1908</td>
<td>31.2264</td>
</tr>
<tr>
<td></td>
<td>PSNR</td>
<td>38.9612</td>
<td>31.1825</td>
<td>34.8845</td>
<td>34.6691</td>
<td>33.1856</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.9880</td>
<td>0.9850</td>
<td>0.8918</td>
<td>0.9914</td>
<td>0.9789</td>
</tr>
<tr>
<td><strong>3</strong></td>
<td>MSE</td>
<td>10.1205</td>
<td>32.9677</td>
<td>18.9452</td>
<td>20.0782</td>
<td>25.9884</td>
</tr>
<tr>
<td></td>
<td>PSNR</td>
<td>32.6615</td>
<td>33.8256</td>
<td>38.1478</td>
<td>35.1036</td>
<td>33.7619</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.9825</td>
<td>0.9568</td>
<td>0.9875</td>
<td>0.9922</td>
<td>0.9877</td>
</tr>
<tr>
<td><strong>4</strong></td>
<td>MSE</td>
<td>9.8677</td>
<td>36.1298</td>
<td>25.6789</td>
<td>19.1387</td>
<td>25.4587</td>
</tr>
<tr>
<td></td>
<td>PSNR</td>
<td>39.6252</td>
<td>33.7859</td>
<td>35.9421</td>
<td>35.3117</td>
<td>33.7555</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.9409</td>
<td>0.8974</td>
<td>0.9915</td>
<td>0.9925</td>
<td>0.9819</td>
</tr>
<tr>
<td><strong>5</strong></td>
<td>MSE</td>
<td>11.0712</td>
<td>37.8957</td>
<td>28.6230</td>
<td>21.4431</td>
<td>38.3294</td>
</tr>
<tr>
<td></td>
<td>PSNR</td>
<td>39.1254</td>
<td>33.7599</td>
<td>38.9885</td>
<td>34.8179</td>
<td>33.3122</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.7458</td>
<td>0.8526</td>
<td>0.9587</td>
<td>0.9917</td>
<td>0.9869</td>
</tr>
</tbody>
</table>

Figures (12), (13), and (14) display the cumulative effects of mean squared error, peak signal to noise ratio and similarity structure index image of tested images.

![Fig. 12- Mean square error of speech hiding method.](image-url)
Fig. 13- PSNR of the speech hiding method.

Fig. 14- Structural similarity Index Image of proposed method

The histogram of stego images, compared to the original image, is kept in the same shape as possible, as shown in figure (15) for word speech 1 hiding, while for speech 2, 3, 4, and 5 are almost the same.

Image dataset

Before Hiding

After Hiding

Fig. 15- Histogram test for speech hiding (speech 1).
b. Speech Compression

Speech samples always have large-scale volumes of data that represent them with high redundancy. The compression of speech signals is necessary for size reduction of secret data in the cover image. The proposed compression quantized DCT parameters and the remaining parameters will be embedded cover medium (codded image). Table (2) explains the number of samples and a compression ratio of two recorded speeches as a pattern of the person recording with multiple rates of Norm ratio concerning the original Norm: 99%, 98%, 97%, 96%, and 95%, respectively.

Table 2- Compression ratio of speech.

<table>
<thead>
<tr>
<th>Speech samples 24000 samples (3 sec)</th>
<th>Norm Ratio</th>
<th>No of Sample After DCT compression</th>
<th>Compression Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>99%</td>
<td>6754</td>
<td>28.2657</td>
</tr>
<tr>
<td>2</td>
<td>99%</td>
<td>4258</td>
<td>25.5924</td>
</tr>
<tr>
<td>1</td>
<td>98%</td>
<td>5364</td>
<td>15.1536</td>
</tr>
<tr>
<td>2</td>
<td>98%</td>
<td>8925</td>
<td>14.8464</td>
</tr>
<tr>
<td>1</td>
<td>97%</td>
<td>1255</td>
<td>13.2671</td>
</tr>
<tr>
<td>2</td>
<td>97%</td>
<td>3625</td>
<td>8.7124</td>
</tr>
<tr>
<td>1</td>
<td>96%</td>
<td>2989</td>
<td>9.8732</td>
</tr>
<tr>
<td>2</td>
<td>96%</td>
<td>1615</td>
<td>8.1479</td>
</tr>
<tr>
<td>1</td>
<td>95%</td>
<td>1798</td>
<td>7.1548</td>
</tr>
<tr>
<td>2</td>
<td>95%</td>
<td>1218</td>
<td>4.8264</td>
</tr>
</tbody>
</table>

The previous Table shows the average compression ratio when Norm Ratio are: 99% is (28.2657), 98% is (15.1536), 97% is (13.2671), 96% is (9.8732), 95% is (7.1548). The speech signal after compression is kept the same as possible as the original signal to retain speech quality, as shown in Figure (16).

Fig. 16- Speech signal before and after compression for all speech

5. CONCLUSION

Data hiding can be used for various purposes, such as covert communication, digital watermarking, or copyright protection. However, using it responsibly and legally is essential, respecting privacy and ethical considerations. The speech was hidden in the encoded image proposed in this work. The speech is considered a secret that was
compressed using the DCT method, quantized then to a binary form. The Chaotic Map methods Zaslavsky maps are used as the key for embedded location. The system's accuracy reached 98%, and the modification of the quantization level was 16 bits, one in each block (4 × 4). It is noted that speech masking in the image suffered from some limitations because the volume of the speech signal is generally higher than the image. Thus, using DWT will reduce distortion in the image stego by leaving the low subscript unchanged. Speech compression only can embed the cover image.

6. REFERENCES