A Highly Secure Skin Tone Based Optimal Parity Assignment Steganographic Scheme Using Double Density Discrete Wavelet Transform

Aruna Mittal

Department of CSE, Disha Institute of Management and Technology, Chhattisgarh Swami Vivekanand Technical University, Bhilai, India

ABSTRACT: Steganography is the art of hiding the existence of data in another transmission medium i.e. image, audio, video files to achieve secret communication. It does not replace cryptography but rather boosts the security using its obscurity features. The proposed method uses both Cryptography and Steganography to enhance the security of the message. The secret message is first encrypted using RSA algorithm and then randomized using OAEP[1]. This encoded message is then embedded in the bitmap cover image using frequency domain approach. For embedding the encrypted message, initially skin tone regions of the cover image are detected using HSV (Hue, Saturation, Value) model. Thereafter, a region from skin detected area is selected, which is known as the cropped region. In this cropped region secret message is embedded using DD-DWT (Double Density Discrete Wavelet Transform). DD-DWT overcomes the intertwined shortcomings of DWT (like poor directional selectivity, Shift invariance, oscillations and aliasing)[2]. Hence the image obtained after embedding secret message (i.e. Stego Image) is far more secure and has an acceptable range of PSNR. The proposed method is much better than the previous works both in terms of PSNR and robustness against various noises (like Poisson, Gaussian, salt and pepper, rotation, translation etc.)

Keywords: Cropping, DD DWT, DWT, Gaussian, HSV, OAEP, Poisson, PSNR, RSA, Skin tone detection, Stego Image.

I. INTRODUCTION

In this highly digitalized world, the Internet serves as an important role for data transmission and sharing. However, since it is a worldwide and publicized medium, some confidential data might be stolen, copied, modified, or destroyed by an unintended observer. Therefore, security problems become an essential issue. Encryption is a well known procedure for secured data transmission [3]. Frequently used encryption methods include RSA, DES (Data encryption standard). Although encryption achieves certain security effects, they make the secret messages unreadable and unnatural or meaningless. These unnatural messages usually attract some unintended observers’ attention. This is the reason a new security approach called “steganography” arises. As an example, the cover text [4]: “I’m feeling really stuffy. Emily’s medicine wasn’t strong enough without another febrifuge.” Hides the sentence “Meet me at nine”, if the reader retains the second letter of each word in sequence. In steganography secret message is the data that the sender wishes to remain confidential and can be text, images, audio, video, or any other data that can be represented by a stream of bits. The cover or host is the medium in which the message is embedded and serves to hide the presence of the message. The message embedding technique is strongly dependent on the structure of the cover, and in this paper covers and secret messages are restricted to being digital images. The cover-image with the secret data embedded is called the “Stego-Image”. The Stego-Image should resemble the cover image under casual inspection and analysis. In addition, for higher security requirements, we can encrypt the message data before embedding them in the cover-image to provide further protection [5]. For this the encoder usually employs a stego-key which ensures that only recipients who know the corresponding decoding key will be able to extract the message from a stego-image. For proposed method cover image is cropped interactively and that cropped region works as a key at decoding side yielding improved security.

There are two things that need to be considered while designing the Steganographic system. (a) Invisibility: Human eyes cannot distinguish the difference between original and stego image. (b) Capacity: The more data an image can carry the better it is. However large embedded data may degrade image quality significantly [6].

Rest of the paper is organized as follows. Section II presents literature review and theoretical background. In section III proposed method is described in detail with skin tone detection, DWT, embedding and extraction procedure step by step. Section IV demonstrates the experimental results. Finally conclusions are provided in section V.

II. LITERATURE REVIEW

A. Steganography in Spatial Domain

This is the simplest Steganographic technique that embeds the bits of secret message directly into the least significant bit (LSB) plane of the cover image. In a gray level image, every pixel consists of 8 bits. The basic concept of LSB substitution is to embed the confidential data at the rightmost bits, (bits with the smallest weighting) so that the embedding procedure does not affect the original pixel value greatly [7]. The mathematical representation for LSB is:

\[ x'_i = x_i - \left( x_i \mod 2k \right) + m_i \quad \ldots \ldots (1) \]

In equation (1), \( x'_i \) represents the \( i^{th} \) pixel value of the stego-image and \( x_i \) represents that of the original cover image. \( m_i \) represents the decimal value of the \( i^{th} \) block in the confidential data. The number of LSBs to be substituted is \( k \). The extraction process is to copy the \( k \) rightmost bits directly. Mathematically the extracted message is represented as:

\[ m_i = x_i \mod 2k \quad \ldots \ldots (2) \]
Hence, a simple permutation of the extracted m, gives the original confidential data [8]. This method is easy and straightforward but this has low ability to bear some signal processing or noises. And secret data can be easily stolen by extracting the whole LSB plane.

B. Steganography in Frequency Domain

Robustness of steganography can be improved if properties of the cover image could be exploited. For example it is generally preferable to hide the message in noisy regions rather than smoother regions as degradation in smoother regions is more noticeable to HVS (Human Visual System). Taking these aspects into consideration working in frequency domain becomes more attractive. Here, sender transforms the cover image into frequency domain coefficients before embedding secret messages in it [9]. Different sub-bands of frequency domain coefficients give significant information about where vital and non- vital pixels of image resides. These methods are more complex and slower than spatial domain methods; however they are more secure and tolerant to noises. Frequency domain transformation can be applied either using DCT or DWT.

C. Adaptive Steganography

Adaptive steganography is special case of two former methods. It is also known as “Statistics aware embedding” [10] and “Masking” [11]. This method takes statistical global features of the image before attempting to embed secret data in DCT or DWT coefficients. The statistics will dictate where to make changes.

III. PROPOSED METHOD

Proposed method introduces a new method of embedding secret data within skin and as well as in the edge area, as it is not that much sensitive to HVS (Human Visual System). This method takes advantage of Biometrics features such as skin tone edge detection, instead of embedding data anywhere in Image, data will be embedded in selected regions like skin region. Overview of method is briefly introduced as follows. At first skin tone detection is performed on input image using HSV (Hue, Saturation, Value) color model. Secondly cover image is transformed in Frequency domain. This is performed by applying DD-DWT. Then payload (number of bits in which we can hide data) is calculated. Finally secret data embedding is performed in one of the high frequency sub-band by tracing skin pixels in that band. Before performing all steps cropping on input image is performed and then in only cropped region embedding is done, not in whole image. Cropping results into enhanced security, since cropped region works as a key at the decoding side. Here embedding process affects only certain Regions of Interest (ROI) rather than the entire image. So utilizing objects within images can be more advantageous. This is also called as Object Oriented steganography. Then a stego DD-DWT image is produced, so the IDD-DWT is performed on that. Thereafter IDD-DWT image is merged with original image, and we get the final stego image.

A. Skin Color Tone Detection

A skin detector typically transforms a given pixel into an appropriate color space and then uses a skin classifier to label the pixel whether it is a skin or a non-skin pixel. The Skin classifier used for the proposed method is based on the following values of RGB [12]:

\[
\begin{align*}
R & > 95 \text{ and } G > 40 \text{ and } B > 20 \\
\max(R, G, B) - \min(R, G, B) & > 15 \\
|R| - |G| & > 15 \text{ and } R > G \text{ and } R > B
\end{align*}
\]

These RGB values can be converted into HSV by using eq\(^n\) (3)

\[
H = \begin{cases} 
\text{if } h, B \leq G \\
\text{if } h, B > G
\end{cases}
\]

\[
\cos^{-1} \left( \frac{1}{\sqrt{(R - G)^2 + (R - B)^2}} \right) + (R - B)
\]

\[
S = \max(R, G, B) - \min(R, G, B)
\]

\[
V = \max(R, G, B)
\]

B. Discrete Wavelet Transform (DWT)

This is another frequency domain in which steganography can be implemented. DCT is calculated on blocks of independent pixels, a coding error causes discontinuity between blocks resulting in annoying blocking artifacts. This drawback of DCT is eliminated using DWT. DWT applies on entire image. DWT offers better energy 40 compaction than DCT without any blocking artifact. DWT splits component into numerous frequency bands called sub bands known as

- LL – Horizontally and vertically low pass
- LH – Horizontally low pass and vertically high pass
- HL – Horizontally high pass and vertically low pass
- HH – Horizontally and vertically high pass

Since Human eyes are much more sensitive to the low frequency part (LL subband) we can hide secret message in other three parts without making any alteration in LL subband [12]. As other three sub-bands are high frequency sub-band they contain insignificant data. Hiding secret data in these sub-bands don’t degrade image quality that much. DWT used in this work is DD-DWT.

C. Implementation of DWT in 1D

In separable DWT the analysis filter bank decomposes the input signal x(n) into two sub band signals, c(n) and d(n). The signal c(n) represents the low frequency part of x(n), while the signal d(n) represents the high frequency part of x(n). We denote the low pass filter by af1 (analysis filter 1) and the high pass filter by af2 (synthesis filter 2). As depicted in figure(1), the output of each filter is then down sampled by 2 to obtain the two sub band signals c(n) & d(n)[13].
The Synthesis filter bank combines the two sub band signals c(n) & d(n) to obtain a single signal y(n). The synthesis filter bank up-samples each of the two sub band signals. The signals are then filtered using a low pass and high pass filter. We denote the low pass filter by sf1(synthesis filter 1) and the high pass filter by sf2(synthesis filter 2). The signals are then added together to obtain the signal y(n). If the four filters are designed so as to guarantee that the output signal y(n) equals the input signal x(n), then the filters are said to satisfy the perfect reconstruction condition.

D. 2-D Discrete Wavelet Transform

Image-processing applications require two-dimensional implementation of wavelet transform. Implementation of 2D DWT [14],[15],[16] is also referred to as multidimensional wavelet transform in literature. In the 2D case, the 1D analysis filter bank is first applied to the columns of the image and then applied to the rows. If the image has N1 rows and N2 columns, then after applying the 1D analysis filter bank to each column we have two sub band images, each having N1/2 rows and N2 columns; after applying the 1D analysis filter bank to each row of both of the two sub band images, four sub band images are obtained, each having N1/2 rows & N2/2 columns. This is depicted in figure (2) given below. The 2D synthesis filter bank combines the four sub band images to obtain the original image of size N1 by N2 [15][16].

The analysis filter bank consists of three analysis filters h0(n), h1(n) and h2(n) such that y(n)=x(n).

G. Thresholding Techniques

Generally two types of thresholding techniques are there in spread spectrum denoising [17]:

1) Hard Thresholding: Hard Thresholding is a straightforward technique for implementing wavelet denoising (i.e., ‘keep’ or ‘kill’ strategy). If T is the threshold, then hard thresholding operation on the wavelet coefficient w(n) is given by

\[ \delta^1_W(W_i) = \begin{cases} W_i & \text{if} |W_i| > T \\ 0 & \text{otherwise} \end{cases} \tag{4} \]

This operation is not a continuous mapping and only affects the input coefficients that are less than or equal to the threshold. Proposed method uses hard Thresholding.

2) Soft Thresholding: The other standard technique for denoising is soft Thresholding [18] of the wavelet coefficient w(n) via,

\[ \delta^2_W(W_i) = \text{sign}(W_i)(|W_i| - T) \]

where, \( \text{sign}(W_i) = +1 \text{ if } W_i > 0, \)
\(-1 \text{ if } W_i < 0, \)
\(0 \text{ if } W_i = 0, \)
and \( x_+ = x \text{ if } x \geq 0, \)
\(0 \text{ if } x < 0 \)

Where, sign(w(n)) is the signum function. Instead of forcing w_i to zero or leaving it untouched, soft thresholding pushes all the coefficients towards zero. Hence the smoothing effect is better in soft thresholding than the hard thresholding. So, soft thresholding is preferred in this work.

H. RSA and OAEP Encryption
1) **RSA Encryption:** The RSA cryptosystem, named after its inventors R. Rivest, A. Shamir, and L. Adleman, is the most widely used public key Cryptosystem. It may be used to provide both secrecy and digital signatures and its security is based on the intractability of the integer factorization.

The scheme makes use of an expression with exponentials. Plaintext is encrypted in blocks, with each block having a binary value less than some number n. That is, the block size must be less than or equal to \( \log_2(n) \); in practice, the block size is 1 bit, where \( 2^k < n < 2^{k+1} \). Encryption and decryption are of the following form, for some plaintext block \( M \) and cipher block \( C \):

\[
C = M^e \mod n
\]

\[
M = C^d \mod n = (M^e)^d \mod n = M^{ed} \mod n
\]

Both sender and receiver must know the value of n. The sender knows the value of e, and only the receiver knows the value of d. Thus, this is a public-key encryption algorithm with a public key of \( PU = \{ e, n \} \) and a private key of \( PU = \{ d, n \} \). For this algorithm to be satisfactory for public-key encryption, the following requirements must be met:

i) It is possible to find values of e, d, n such that \( M^{ed} \mod n = M \) for all \( M < n \).

ii) It is relatively easy to calculate \( M^{ed} \mod n \) and \( C^d \) for all values of \( M < n \).

iii) It is infeasible to determine \( d \) given \( e \) and \( n \).

The algorithm is described as

**Key Generation:**

i) Select two random numbers \( p \) and \( q \) such that both are prime and \( p \neq q \).

ii) Calculate \( n = pq \).

iii) Calculate \( \varphi(n) = (p-1)(q-1) \).

iv) Select public key e such that \( \gcd(\varphi(n),e) = 1 \) and \( 1 < e < \varphi(n) \).

v) Calculate private key d such that \( de - 1 \equiv 1 \mod(\varphi(n)) \).

vi) Public key is given by \( PU = \{ e,n \} \).

vii) Private key is given by \( PR = \{ d,n \} \).

**Encryption:**

i) Plaintext \( M \) should be such that \( M < n \).

ii) Cipher text \( C = M^e \mod n \).

**Decryption:**

i) \( M = C^d \mod n \).

2) **Security of RSA:** Four possible approaches to attacking the RSA algorithm are as follows:

i) Brute force: This involves trying all possible private keys.

ii) Mathematical attacks: There are several approaches, all equivalent in effort to factoring the product of two primes.

iii) Timing attacks: These depend on the running time of the decryption algorithm.

iv) Chosen cipher text attacks: This type of attack exploits properties of the RSA algorithm.

The defence against the brute-force approach is the same for RSA as for other cryptosystems, namely, use a large key space. Thus, the larger the number of bits in \( d \), the better. However, because the calculations involved, both in key generation and in encryption/decryption, are complex, the larger the size of the key, the slower the system will run. We can identify three approaches to attacking RSA mathematically:

i) **Factor n into its two prime factors.** This enables calculation of \( f(n) = (p - 1) \times (q - 1) \), which, in turn, enables determination of \( d \).

ii) Determine \( f(n) \) directly, without first determining \( p \) and \( q \). Again, this enables determination of \( d \).

iii) Determine \( d \) directly, without first determining \( f(n) \).

3) **OAEP:** To overcome the drawbacks of RSA, a randomization approach is combined to it namely OAEP. Optimal Asymmetric Encryption Padding (OAEP) is a padding scheme in the form of a Feistel network which uses a pair of random oracles G and H to process the plaintext prior to asymmetric encryption. When combined with any secure trapdoor one-way permutation f, this processing is proved in the random oracle model to result in a combined scheme which is semantically secure under chosen plaintext attack (IND-CPA). When implemented with certain trapdoor permutations (e.g., RSA), OAEP is also proved secure against chosen ciphertext attack (IND-CPA). OAEP can be used to build an all-or-nothing transform. OAEP satisfies the following two goals:

i) Add an element of randomness which can be used to convert a deterministic encryption scheme (e.g., traditional RSA) into a probabilistic scheme.

ii) Prevent partial decryption of cipher texts (or other information leakage) by ensuring that an adversary cannot recover any portion of the plaintext without being able to invert the trapdoor one-way permutation f.

4) **Implementation of OAEP:**

![Fig 4. OAEP Diagram](image-url)
To decode,

i) recover the random string as \( r = Y \oplus H(X) \)

ii) recover the message as \( m_{00...0} = X \oplus G(r) \)

The "all-or-nothing" security is from the fact that to recover \( m \), you must recover the entire \( X \) and the entire \( Y \); \( X \) is required to recover \( r \) from \( Y \), and \( r \) is required to recover \( m \) from \( X \). Since any changed bit of a cryptographic hash completely changes the result, the entire \( X \), and the entire \( Y \) must both be completely recovered.

### I. Encoding and Data Hiding Process

Suppose \( C \) is original 24-bit color cover image of \( PXQ \) Size.

\[ C = \{x_{ij}, y_{ij}, z_{ij} | 1 \leq i \leq P, 1 \leq j \leq Q, x_{ij}, y_{ij}, z_{ij} \in \{0, 1,..., 255\} \} \]

Let size of cropped image is \( Pc \times Qc \) where \( Pc \leq P \) and \( Qc \leq Q \) and \( Pc = Qc \). i.e. Cropped region must be exact square as we have to apply DWT later on this region. Let \( S \) is secret data. Here secret data considered is binary image of size \( a \times b \).

Figure 5 represents flowchart of embedding process. Different steps of flowchart are given in detail below.

**Fig 5. Flowchart of Encoding Process**

Steps:

1. Initially load the cover object in which we will hide the secret message (text).

2. After loading the cover object, skin tone detection is performed. This enables us to know where and how much data can be hidden.

3. Cropping: From the detected skin portion, cropping is performed. This is done so that within skin pixels data is hidden at only limited pixel positions. This feature of cropping enhances security, as any eavesdropper cannot detect secret message just by detecting the skin pixels.

4. Histogram Modification: This is performed to adjust the contrast of the colors.

5. Key Generation: This is the step where the secret message to be selected and is encrypted using RSA and OAEP.

6. Double Density DWT: Double Density Discrete Wavelet Transform is applied to the cropped skin portion.

7. Secret encrypted message is now merged into the transformed skin pixels.

8. Optimal Parity Assignment is used to assign secret code values to limited areas of cropped skin portion, so as to have least effect over the HVS (human visual system).

9. Inverse DD-DWT: Now the transformed image has secret code as well, so it is ready to be merged with the original cover object. The first step to merge this transformed secret message embedded image, with cover object is to inverse transform it.

10. After applying inverse DD-DWT, we get the original cropped image along with secret code. This image is now called stego image. This stego image is now merged with original cover image to get the final reconstructed cover image along with secret data embedded in it. This Stego image is now sent to the receiver by some transmission medium.

**Fig 6. Flowchart of Decoding Process**

At the Decoding End following steps are performed:

1. From the Stego Image skin pixels are detected and cropping of Stego image is performed.

2. Now the DD-DWT is performed to get the transformed cropped image.

3. Secret encrypted message is extracted from the transformed cropped stego image. This encrypted message is decrypted (using RSA+ OAEP decryption) to get the secret message.

4. Results of Extraction process are measured in terms of PSNR and MSE. This are discussed below in detail.
IV. RESULTS

In this section we demonstrate simulation results for the proposed scheme. These have been implemented using MATLAB 7.6.0. A 24 bit color image is employed as cover-image of size 256×256, shown in Fig. 7. Fig.8 shows sample secret message image to hide inside cover image.

![Fig 7: Cover image](image)

![Fig 8: Secret message image](image)

Performance measurement for image distortion is well known as peak signal to noise ratio (PSNR) which is classified under the difference distortion metrics and can be applied on stego images. PSNR is used to evaluate quality of stego image after embedding the secret message. Secret message can be any word. The performance in terms of capacity and PSNR (in dB) is demonstrated for the method in the following subsections. PSNR is defined as per Eq.9

$$PSNR = 10 \log_{10} \left( \frac{255^2}{MSE} \right) \quad \text{(9)}$$

$$MSE = \left( \frac{1}{M \times N} \right) \sum_{i=1}^{M} \sum_{j=1}^{N} (a_{ij} - b_{ij})^2 \quad \text{(10)}$$

$a_{ij}$ and $b_{ij}$ represents pixel values of original cover image and stego image respectively as in Eq.10 The calculated PSNR as in Eq.9 usually adopts dB value for quality judgement, the larger PSNR is higher the image quality (which means there is a little difference between cover image and stego image).On the contrary smaller dB value means there is a more distortion. PSNR values falling below 30dB indicate fairly a low quality.

However, high quality strives for 40dB or more.

A. Result Discussion of proposed work

After embedding secret data in cropped image, resulted cropped stego image is shown in Fig. 9. Cover image is now merged with cropped embedded Stego image as is shown in Fig.10. For merging, co-ordinates of first and last pixels of cropped image are calculated and then replaced with the one in original cover image. After performing decoding process on stego image, retrieved output text file consisting of the secret message is shown in Fig 11.

![Fig 9: Cropped Stego Image](image)

![Fig 10: Merged Stego Image](image)

![Fig 11: Output Text File (having the secret message)](image)
PSNR for different cases is shown in table 1. Average PSNR of proposed method is calculated based on the obtained PSNR. Average PSNR obtained by the proposed method is much better than the ones proposed by Rekha Nagar and Anjali Shejul (as can be seen in table 2). Table I also includes PSNR of considered image after addition of noises (like Gaussian, Salt and Pepper, Speckle, Poisson, and Image Rotation), which are fairly acceptable (having PSNR greater than 40%). Thus the proposed method is better than previous ones as well as robust against various noises.

**TABLE I**

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Cover Image</th>
<th>With or Without Addition of Noise</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image 1</td>
<td>Without Noise</td>
<td>71.4286</td>
</tr>
<tr>
<td>2</td>
<td>Image 2</td>
<td>Without Noise</td>
<td>51.9769</td>
</tr>
<tr>
<td>3</td>
<td>Image 1</td>
<td>Gaussian Noise</td>
<td>48.6234</td>
</tr>
<tr>
<td>4</td>
<td>Image 1</td>
<td>Salt and Pepper Noise</td>
<td>56.3243</td>
</tr>
<tr>
<td>5</td>
<td>Image 1</td>
<td>Speckle Noise</td>
<td>49.3423</td>
</tr>
<tr>
<td>6</td>
<td>Image 1</td>
<td>Poisson Noise</td>
<td>58.6758</td>
</tr>
<tr>
<td>7</td>
<td>Image 1</td>
<td>Image Rotation</td>
<td>42.6754</td>
</tr>
</tbody>
</table>

**TABLE II**

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Method</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Rekha Nagar</td>
<td>50.5</td>
</tr>
<tr>
<td>2</td>
<td>Anjali Shejul</td>
<td>51</td>
</tr>
<tr>
<td>3</td>
<td>Proposed Method</td>
<td>51.9769</td>
</tr>
</tbody>
</table>

V. CONCLUSION

Digital Steganography is a fascinating scientific area which falls under the umbrella of security systems. Proposed framework is based on steganography that uses Biometric feature i.e. skin tone region. Skin tone detection plays a very important role in Biometrics and can be considered as secure location for data hiding. Secret data embedding is performed in DD-DWT domain than the DWT as DD-DWT outperforms than DWT as well as DCT. Using Biometrics resulting stego image is more tolerant to attacks and more robust than existing methods.

ACKNOWLEDGEMENT

I would like to express my sincere thanks to Reader Mrs. Preeti Tuli for giving me the opportunity to explore this field of Object Oriented Steganography. She has always motivated and supported me at all stages of the project work. Also I would like to thank our Department of Computer Science and Engineering for providing me all the help as and when required.

REFERENCES


BIographies

Aruna Mittal received her degree of B.E. in Computer Science and Engineering from Rajeev Gandhi Technical University, Bhopal, in the year 2007. She is pursuing her M.Tech Degree in Information Security from Chhattisgarh Swami Vivekanand Technical University, Bhilai. She has more than 3 years of industry experience. Her interest areas are data warehousing and information security.