

SURVEY ON RECENT DIGITAL IMAGE STEGANOGRAPHY TECHNIQUES

^{1,2}SUHAD SHAKIR JABER, ¹HILAL ADNAN FADHIL, ¹ZAHEREEL I. ABDUL KHALIB,
¹RASIM AZEEZ KADHIM

¹School of Computer and Communication Engineering, Universiti Malaysia Perlis (UniMAP),
Pauh Putra, Arau, Perlis 02600, Malaysia

²Ministry of Sciences and Technology, Baghdad, Iraq.

E-mail: ¹shaker_suhad@yahoo.com

ABSTRACT

Steganography is the science of hiding information that involves all the techniques used to exchange the secret message with low distortion of the cover medium. Many different cover medium formats such as (image, audio, video, and text) can be used to hide the secret message. Image files are mostly used because of their frequency on the Internet. Some of the old techniques used schemes to hide information are: invisible ink, null ciphers, micro-codes, and pink-pricks. Modern Steganography has gained a lot of attention for the last two decades because of the rapid growth of communication technologies such as Internet and the need of a secure channel to transmit the important information. In this paper, a concept of image steganography is explained and a review of recent image steganography methods and its applications is presented.

Keywords: *Discrete cosine transform, Ddiscrete wavelet transform, Least significant bit, Vector quantization, Steganography.*

1. INTRODUCTION

Information hiding system generally involves both watermarking and steganography. A main goal of watermarking system is the high level of robustness, in order to make the removing of watermark impossible by a third party without degrading the data object's quality. Steganography, on the other hand, is used to pursue high capacity and security [1]. As shown in Fig.(1), three aspects represent a triangle in information hiding systems that are used to rate the system performance. Besides, capacity is the amount of secret data that can be embedded into the cover medium. It is preferable to be as high as possible. Security or imperceptibility is the ability of the hiding information system to make the cover medium indistinguishable from the stego object by the human and the computer statistical methods. Robustness is the amount of modification the stego medium can withstand before an adversary can destroy hidden information. Furthermore, Steganography is the art of science that deals with sending secret information by embedding into the cover object invisibly. Here, the authorized party is only aware of the existence of the hidden message. An ideal steganographic technique conceals a large amount of information ensuring that the modified

object is not visually or audibly distinguishable from the original object [2].

The steganography components are:

- a- The cover medium: is any medium used for transporting the secret message.
- b- The secret message: is the important data that want to be transported securely.
- c- The stego key: is a sequence of data generated by the sender during the embedding process and should be used by the receiver to recover the embedded message. A person having the stego key can only access the secret message.
- d- The stego-medium: is the cover medium after the secret message is embedded.

Steganography is a good candidate in secure communications in cases where the use of cryptography is not allowed or raises suspicion. One of these applications is military and intelligence agents where a high priority is saving an agent's life. Even if the cryptography is used, the detection of secret messages by cryptanalysis techniques may rapidly lead to an attack on the agent. In addition, steganography can be used in the protection of data alteration, in companies for the safe circulation of secret data, and in accessing the control systems for digital content distribution and so on [3].

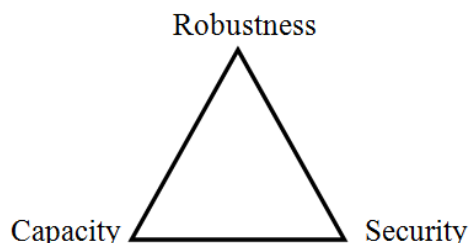


Figure 1: The Competing Factors In Hiding Information System.

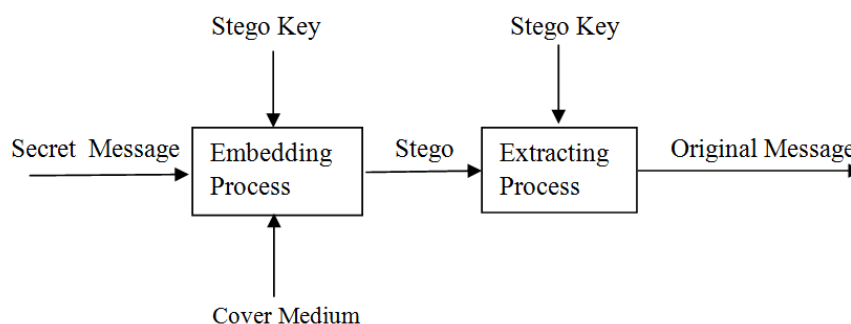


Figure 2: General Hiding Information System.

Steganography uses many types of digital medium as a cover for hiding the secret data such as: Text, Image, Audio, Video, and protocol. The digital files have redundant bits that can be altered without the alteration being detected easily. The image file is the most appropriate carrier type because it has more redundant bits that can be replaced with secret information bits with minimum suspicion. Also, images are frequently used through Internet in websites or an e-mail attached. Image steganography get more attentions for the last two decades. Therefore, this review focuses on the image steganography approach.

The recent image steganography techniques can be classified into [4, 5]:

- 1- Spatial (image) domain.
- 2- Compressed domain based on vector quantization (VQ).
- 3- Transform domain.
- 4- Spread spectrum.
- 5- Statistical technique
- 6- Distortion technique

The rest of this paper is organized as follows: section II explains the concept of spatial domain techniques and the related works. The compressed domain based on vector quantization is presented in section III. Section IV discusses the transform

domain techniques. The concepts of spread spectrum, statistical technique, and distortion technique are introduced in sections V, VI, and VII respectively. Finally, the conclusions are drawn in section VIII.

2. SPATIAL DOMAIN

Spatial domain steganographic techniques, also known as substitution techniques, include a class of relatively simple techniques attempting to create a covert channel in portions of the cover image where changes are likely to be imperceptible visually to an outsider. One method is to hide information in the least significant bit (LSB) of cover image data. LSB based steganography is one of the simplest techniques used to embed a secret data in the LSBs of image pixel values with a very low distortion. There are several variations, but the basic technique involves replacing the pixels LSB of the 8-bit blocks of the cover image with the secret message [6]. The embedding operation of LSB steganography is described by the following equation:

$$y_i = 2 \left\lfloor \frac{x_i}{2} \right\rfloor + m_i \quad (1)$$

where m_i , x_i and y_i are the i -th message bit, and the i -th chose pixel value before and after embedding, respectively.

For example, consider an 8-bit gray scale image where each pixel is stored as a byte representing a gray scale value. Suppose the first eight pixels of the original image have the following gray scale values:

Pixels:

```
00100110 10100010 11001111 00000000
01010010 00011100 11100111 01011101
```

To hide the letter 'A' whose binary representation is 01000001, we would replace the LSBs of these pixels to have the following new gray scale values:

Pixels after embedding process:-

```
00100111 10100010 11001110 00000000
01010010 00011100 11100111 01011100
```

Note that, on average, less than half of the LSBs need to change. The difference between the cover (i.e. original) image and the stego image will be hardly noticeable to the human eye. Another example of LSB technique is: Consider a 24-bit colour image and the letter A is to be embedded using LSB technique. The resulting grid is as follows:

```
Pixels: (01000100 00010110 11111000) (10101010
01111100 01110111) (11001110 10110110
10100101)
```

Letter A: 01000001

Pixels after embedding process:-

```
(01000101 00010110 11111000)
(10101010 01111100 01110110) (11001111
10110110 10100101)
```

Generally speaking, the LSB technique has some advantages. There is a little chance for degradation of the original image. Besides, more secret data can be stored in an image. On the other hand, the disadvantages are represented in the less robust, the hidden data can be lost with the image manipulation and easily destroyed by simple attacks. Many techniques of spatial domain steganography are proposed, all of these modify some bits of the pixel values in the cover image. Spatial domain techniques can be classified into [7]:

1. Least significant bit (LSB)
2. Pixel value differencing (PVD)
3. Edges based data embedding method (EBE)
4. Random pixel embedding method (RPE)
5. Mapping pixel to hidden data method
6. Labelling or connectivity method
7. Pixel intensity based method

8. Texture based method

9. Histogram shifting methods

In [8], Chao Hsu et al. presented a data-hiding technique namely the bipolar multiple number base to provide capabilities of authentication, integration, and confidentiality for an electronic patient record (EPR) transmitted among hospitals through the Internet. This technique can be used for hiding those EPR related data such as diagnostic reports, electrocardiogram, and digital signatures from doctors or a hospital into a mark image representing the mark of a hospital which is used to identify the source of an EPR. The bipolar multiple bases are a number system use multiple positive and negative base numbers to represent the numbers. The proposed technique uses the tolerant error (TER) between the mark image and its JPEG decompressed image to obtain the base numbers.

In [9], C. C. Thein et al. al. proposed a simple and fast method for high-hiding capacity based on the modulus operation. A good image vision quality can be achieved by using this method without the need for post-processing. Although, this method is almost as simple as the LSB method in both embedding and extracting, it has a high-hiding capacity in which it can hide a 256×256 or 256×512 image in a 512×512 host image.

In [6], Chan and Cheng developed a data hiding scheme based on an optimal pixel adjustment process to the stego-image obtained by the simple LSB substitution method. The optimal pixel adjustment process depends on checking the embedding error between the original cover-image and the stego-image obtained by the simple LSB substitution method to form the final stego-image. The authors approved that the quality of the stego-image can be improved with low computational complexity.

A novel steganographic method based on least-significant-bit (LSB) replacement and pixel-value differencing (PVD) method is developed by Wu et al. [10]. The PVD method depends on the difference between two consecutive pixels. A smooth area in an image gives a small difference value but an edged area gives a large difference value. In the smooth areas, LSB method is used to hide the secret data into the cover image while using the PVD method in the edged areas. The proposed method can hide much larger information and maintains a good visual quality of stego-image as compared with the PVD method being used alone.

In [11], Yu, Chang et al. presented a prediction-based image-hiding scheme that embeds secret data into compression codes during image compression. This scheme employs a two-stage structure: a prediction stage and an entropy coding stage. The secret data is embedded into the difference values of a given image after the prediction stage is performed. According to the experimental results, the hiding capacity is high and image quality is better than Jpeg-Jsteg.

In[12], Li, Leung et al. introduced a new image-hiding scheme by exploring the block similarity between the cover-image and the secret-image. Both of the cover image and the secret image are 8-bit gray scale images. Based on the block difference, the best match cover image block of the secret image is selected. Then, the error-matrix, the normalized-error-matrix, the difference-degree and the quantized-error-matrix between the cover-image block and secret-image block are computed. After that, the normalized-error-matrix and quantized-error-matrix is used to modify the cover-image block. This scheme provides a high quality and the secret-image is completely extracted.

In[13], EL-Emam introduced a new algorithm based on hiding large data into colour bitmap(BMP) image by using adaptive image segmentation and adaptive image filtering of the cover-image. All pixels in cover-image are segmented into random numbers of uniform or non-uniform segments,(non-uniform which is more secure than uniform segments because it used to carry the input data) after that applying a compression scheme on the secret data file to increase the amount of hiding secret data and perform encryption on secret data. In addition, the pixels in the cover image are selected randomly rather than sequentially by using a new concept called 'main cases and sub cases' for each byte in one pixel. The algorithm which is described by pseudo-code is presented and it is possible to implement a steganography algorithm to hide a large amount of data into a carrier bitmap image.

In[14], Yu, Chang et al. proposed a steganographic method for hiding a colour or a gray scale secret image in a true colour host image. There are three image-hiding types in the scheme: hiding a colour secret image in a true colour image, hiding a palette-based 256-color secret image in a true colour image, and hiding a gray scale image in a true colour image, which depends on the secret image that is to be hidden in a true colour host image. In all three types of hiding, secret data are encrypted by data encryption standard (DES) method before they are embedded into the host

image. The hiding capacity and good image quality are the results in that proposed method .

In[15], Lin and Hsueh introduced scheme to embed a secret message into eight bits gray scale image with size (512*512), which is divided into non-overlapping three-neighbouring-pixel blocks. The absolute difference between pixels one and two, and the difference between pixels two and three are calculated. Thus three pixels are used to embed two secret bits by modifying only the central pixel.

In[16], Chang, Lin et al presented a novel reversible hiding method for colour images that are compressed by block truncation coding (BTC), which can reconstruct an original image effectively after extracting the embedded secret data. The colour image with size (N*N) is divided into non-overlapping blocks before getting compressed. Each block usually contains three bitmaps and three pairs of quantization levels for R, G, and B, all the blocks are compressed by BTC. Moreover, this method has applied genetic algorithm (GA) to find an approximate optimal common bitmap to replace the original three. The results of this method increased the embedding capacity and decreased the stego image size where the average usually embeds more than three secret bits in each BTC.

In[17], Yang introduced a new LSB-based approach, called the inverted pattern (IP) LSB substitution approach. In this scheme, the cover image is the gray scale image (8-bit). A pre-processing is done on the secret messages before embedding by determining if each section of it is capable of being inverted or not (some secret messages are transformed by inverting operation and some secret messages are not). Moreover, the secret message is transformed into another message to make it more suitable for embedding. A post-processing on the stego-image is done by treating the bits which are used to record the transformation as keys and then by applying to the stego-image the optimal pixel adjustment process. The idea is both simple and efficient and the result is of high capacity and high quality.

In[18], Yang, Weng et al. proposed a new adaptive LSB steganographic method using PVD which is based on the concept that edge areas can tolerate a larger number of changes than smooth areas by using pixel-value differencing to distinguish between edge areas and smooth areas . The cover images used in this method are gray scale images. The range [0, 255] is divided into different levels for different values: lower level, middle level, and higher level respectively. All the pixels are embedded by the -bit LSB substitution, but the value is decided by the level which their difference

value belongs to. A higher level will use a larger value of K . At the end, the experimental results showed that this approach obtained both a larger capacity and a higher image quality when compared to the past study in [9].

In [19], Lee et, Kim et al. proposed a new data hiding scheme for binary image. This scheme requires a small image distortion of the cover image and employs a technique based on Hamming codes (H-EMBED) to embed a large amount of data by flipping only a small number of pixels into the cover image. A new algorithm which is based on Edge Line Segment Similarity Measure (ELSS) is used in this method to select flappable pixels to reduce the visual distortion of the cover image.

In [20], Tsai et al. proposed a reversible data hiding scheme based on prediction and histogram-shifting techniques. For a histogram-shifting based reversible data hiding technique, the payload is determined by the peak height of a histogram. Generally speaking, the higher the peak height of a histogram, the more the payload is.

In [21], Wien Hong et al. introduced a reversible data hiding scheme for high quality images based upon Modification of Prediction Errors (MPE). MPE technique employs the idea of histogram-shifting and prepares vacant enough space to embed secret bits. Firstly, the pixel values are predicted then error values of pixels are obtained. Secondly, by modifying the values of prediction errors, the secret message bits are embedded. This method has the capability to control the PSNR and in addition to embedding fewer messages, it keeps the distortion low.

In [22], Hong and Chen produced a method to extend Tsai et al.'s [18]. This is done by providing a better prediction algorithm based reversible data hiding the technique and the local variance-controlled mechanism, so that a better performance can be achieved. The proposed scheme explores the neighbouring similarity of pixels in the medical image to improve the histogram-based reversible data hiding. The goal of the proposed scheme is to provide a higher hiding capacity while keeping the good quality of the stego- image.

In [23], Chang presented a simple reversible data hiding scheme that uses a complementary embedding strategy. The secret data have been encrypted by using advanced encryption standard (AES) before embedding. The method embeds one secret bit at a time in a raster scan order into one gray scale cover image size ($M \times N$) by increasing or decreasing the cover pixels values by one. Specifically, the horizontal embedding process

decreases the cover pixels having odd values by one and the vertical embedding process increases the cover pixels having even values by one. Such a scheme obtained both a good image quality and a high hiding capacity.

In [24], Yang, Weng et al. explored a new view in spatial domain by using varied combined method of PVD and LSB replacement method to make the programme of embedding data secret more stronger. This method achieves a larger capacity and promotes the peak signal to a noisy ratio (PSNR) value.

In [25], Chen, Chang et al. applied a novel method for steganography by using a hybrid edge detector scheme which is based on LSB. This approach depends on the idea of combining a Fuzzy edge detector and Canny edge detector for gray scale image, to embed secret data in this area. A higher embedding payload and higher quality stego images for the human visual system are a result of this novelty.

In [26], Weiqi Luo et al. proposed an edge adaptive image steganography based on LSB Matching Revisited (LSBMR). The embedding regions can be selected according to a threshold calculated by the secret message size and the difference between two consecutive pixels in the cover image. Therefore, adaptively sharper edge regions are used firstly to embed a small size message, while other regions can be used with the increase of message size.

In [27], Liao, Wen et al., presented steganographic method for digital images with a four-pixel difference and a modified LSB substitution. At the beginning, the pixels located in the cover image are determined whether from edge areas or smooth areas. They are then partitioned into two different levels: high level for edge areas and low level for smooth areas because the first level can embed the amount of a secret data more than the second level without making distortion. The secret data are embedded by using the modified LSB substitution method to increase the embedding capacity and to minimize the distortion.

In [28], Sun, Weng et al. proposed two novel anti-forensic steganography methods: the highlight of exploiting modification direction (HoEMD) and adaptive EMD (ADEMD), to embed a secret message into gray scale images. In HoEMD approach, all directions of pixel are exploited. The pixel which has more changes gives more directions and a high embedding capacity. The adaptive EMD (AdEMD) method is based on the human visual system (HVS) where the edge areas have a higher tolerance of a pixel change than the smooth areas. Therefore, a quantization table of two

difference levels is used to classify the pixel difference under the smooth or edge areas: the higher level for edge areas and the lower level for smooth areas. An important condition for totally extracting the embedded messages is the pixel difference after shifting the data embedded into another level. The proposed scheme provides a high image quality and a large capacity as compared with the PVD and LSB.

In [29], Yang, Weng et al. introduced a steganographic scheme for embedding a secret data imperceptibly into a host gray image using the varieties of PVD method based on the technique of human eye. The host image is divided into four blocks, and each block is processed by being divided into two groups based on PVD. In order to hide more numbers of secret data in each group correctly, which depends on the image area (edge or smooth), this proposition exploits the edge area because it is less sensitive to the change from the smooth area. Besides, it exploits the aspect of increasing of the chance for embedding the data based on a pixel-value shifting scheme. The large embedding capacity and the image with more quality constitute the final result of this method.

In [30], Chen presented a module-based LSB substitution method with lossless secret data compression method. The repetition property of the secret data is used to reduce the replacement of a fewer pixels which are concealed in a smooth area in the stego-image. If the secret data will be hidden, it has heavy repetition, then a fewer pixels will replace. When the secret data have a slight repetition, more replacement is obtained. This method is very good in terms of hiding heavy repetition data. The limitation of this method is the use of only four bits to refer to repetition length. Accordingly, if the repetition length of the data is more than 19, the data needs additional three digit vectors to be recorded in it.

A novel reversible data hiding using block median preservation for gray level images is proposed by Hao Luo et al. [31]. This method employs the median pixel of each block as the reference one and it is based on a multi-level histogram shifting mechanism. The image blocks are divided into four categories and the secret message embedded is based upon a multi histogram-shifting method with reference to the integer median of each block. This technique leads to a perfect result of both capacity and marked image quality.

In [32], Hong and Chen introduced a novel method for data-embedding using Adaptive Pixel Pair Matching (APPM). The values of two pixels are paired and used as a reference coordinator and as an

embedding unit, by replacing it with a scanned coordinator to hide the digit. Diamond encoding (DE) and Exploiting modification direction (EMD) are exploited in this proposed method to extend the payload and to achieve a high capacity for the stego-image.

In [33], Lou and Hu proposed a Reversible Histogram Transformation Function (RHTF) based on LSB steganographic technique to resist statistical steganalysis. Despite the zero points, double-frequencies, and non-accurate detection, which constitute the main problems in RHTF method, the authors present the pixel grouping, cover selection, and dynamic secret key adjustment schemes to solve these problems. The security of this method is higher than the other LSB steganography methods with a high embedding capacity.

In [34], Hong introduced two types for embedding techniques, firstly, Patched reference table (PRT) and secondly PRT-PVD. A special patch is used with type PRT, which selects a pixel pair as an embedding unit, to hide the secret digit based on a reference table. This method achieves a high embedding capacity and the conversion between the two bases is significantly reduced. The second proposed PRT-PVD chose more bits to conceal in the edges area as well as used a special strategy of embedding a sequence to provide a high stego-image quality and more robustness against steganalysis techniques than that of other traditional PVD.

In [35], Hong, Chen et al. proposed a new data embedding method based on PVD and Diamond encoding (DE) with the consideration of HVS. Based on the modified DE method, multiple bases can be used to embed the digit according to the local complexity of the cover image. In addition, this method overcomes the noise problem of DE by finding better pixel pairs to replace the underflow or overflow pixel pairs. This method has improved the image quality and the robust against histogram analysis and the detection methods such as RS scheme.

In [36], Amirtharajan and Balaguru Rayappan proposed a new adaptive hiding method based on four different random walks (Z scan SFC, Hilbert SFC, Zigzag SFC and Moore SFC). At first, the cover image is divided into non-overlapping small blocks with the same size. Then, in each block, the four aforementioned random walks are tested for n-bit embedding. The best random walk which provides a minimum mean square error (MSE) is adopted to embed the confidential data in that block. The information of the chosen walk for each

block is registered as a secret key that is used in the extracting process.

In [37], Bedi, Bansal et al. proposed an efficient image embedding technique in the spatial domain using Particle Swarm Optimization (PSO) for digital image. The best pixel locations in a gray scale cover image are determined by using PSO, where the secret image pixel data, by optimizing an objective function, can be embedded based on the human visual perception and also the tolerance to certain distortions. The results of stego image were good in quality as well as ability to be tolerant to distortions while transmitting over a network.

Two reversible embedding methods based on block median preservation and modification on prediction errors are proposed by H.Y. Leung et al. [38]. The first scheme called Accurate Gradient Selection Predictor (AGSP) which is employed if the secret message bits are smaller than the capacity limit of a single layer hiding. While the second method Median Edge Detection (MED) used to conceal the remaining data. The capacity of the embedding secret data are increased compare with Hong et al.'s (2009) and Luo et al.'s (2011).

In [39], Yuan proposed two simple techniques for secret sharing (SS) with multi-cover adaptive steganography, which share the secret information into one or two LSB planes of cover images. These methods apply some of simple Boolean operations (XOR and AND) to reconstruct the original secret image exactly. The length of secret message that can be embedded by using these methods is greater than the LSB method matched with approximately four times with the same security level. Also, the secret message can be extracted by the receiver without the need of a stego key. Moreover, these techniques have the ability to resist the steganalysis.

3. COMPRESSED DOMAIN BASED ON VQ

Vector Quantization is one of the most popular techniques used for image compression. This technique is proposed by Linde et al. for compressing the gray scale images in 1980. The VQ method starts with the construction of a codebook (Cb) from a set of training images by using Linde-Buzo-Gray (LBG) algorithm [40]. The

codebook involves a set of code words that $Cb = \{C_i | i = 1, 2, \dots, N\}$, where N is the codebook size, and C_i is the i th codeword. Each image to be quantized is partitioned into non-overlapping blocks of size $n \times n$. Each block X represents a vector of h dimensions, where $h = n \times n$, and $X = (x_1, x_2, \dots, x_h)$. Each block in the image is independently encoded with an index of the nearest codeword in the codebook by making a comparison between the current image block and all the codewords of the codebook to find the nearest index.

The similarity between block X and a codeword C_i is often measured by the squared Euclidean distance as follows:

$$D(X, C_i) = \sum_{j=1}^h (x_j - c_{ij})^2 \quad (2)$$

where x_j and c_{ij} denotes the j -th component of the block X and the codeword C_i , respectively.

The block diagram of image compression using VQ is shown in Figure 3.

VQ provides a low bit rate and a simple decoding structure. But, it has a complex codebook design and coding procedures. Data hiding methods for VQ images can be classified into two groups according to the reversibility:-

3.1 Reversible Data Hiding

In this scheme, a secret message can be extracted, and the original cover images also can be restored. The advantage of this scheme is to provide a stego image with a low distortion to avoid the detection of the hidden message and the ability to restore the original cover image. It may have a low hiding capacity to embed secret message.

In [5], Chang, Chen et al. provided a steganographic technique based on SOC compression method of VQ indices. At first, the index table of the VQ is encoded with SOC or original index values (OIV). Then, the bits of the secret message are embedded according to the type of the index by assigning one type for bit "0" and the other for bit "1".

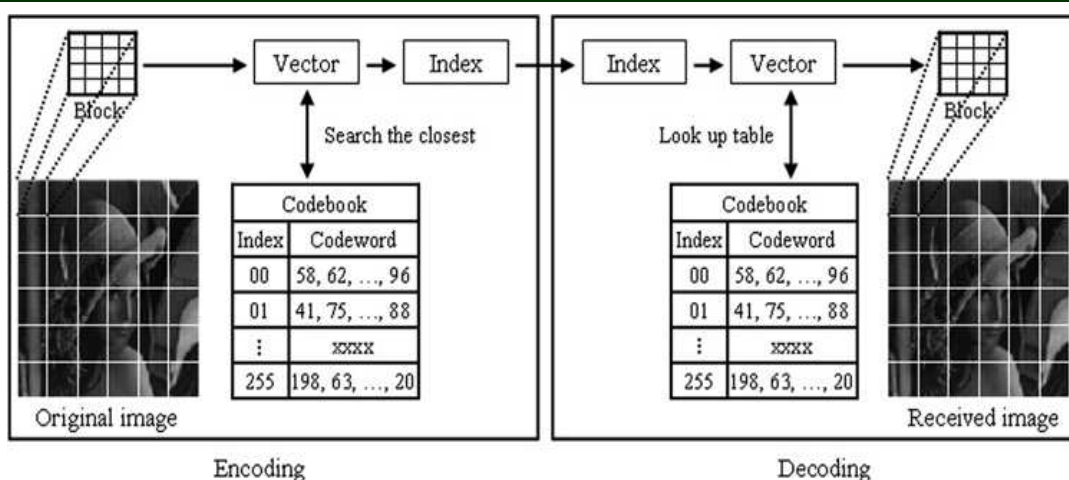


Figure 3: VQ for compressing image [41].

The coding distortion induced is very low and the compression ratio of the image is good.

In [42], Hu proposed a scheme of gray scale image for hiding multiple secret images into the same size cover image for higher hiding capacity and security. This scheme consists of three steps. Firstly, VQ with additional index compression process is used to compress the secret images. Secondly, the DES is used to encrypt the compressed secret images. Finally, the encrypted images are embedded into the host image by using the modulus LSB technique. A secure and high capacity can be obtained by this scheme.

In [43], Chang et al. provided a VQ-based embedding method that can lossless recover the VQ index table for extending the embedding capacity. For one bit embedding algorithm, the codebook is partitioned into three clusters, while for two bits embedding algorithm, the codebook is partitioned into six clusters to increase the hiding capacity. Some indexes of the codebook are reserved as indicators. The secret data are embedded into the VQ index table by transferring index values from one cluster into another. Sometimes, the index values are led by indicators. Only the front one-third of the codebook is used to embed secret data. The distortion of the stego-image is not considered in their method.

In [44], Yung Kuei Chiang, and Piyu Tsai proposed a steganography based on VQ compression by using overlapping codebook partitioning method. In this scheme, the codebook is partitioned into sub-clusters with an overlap based on codeword similarity. To overcome the drawback of the related works, the size of the sub-clusters is restricted to the power of two. This method permits any size of

sub-clusters. Thus, the hiding capacity is improved with the same image quality.

In [45], Yang and Lin presented new strategies for improving the reversible data hiding approach based on VQ index table and referred counts. This method uses the front half of codebook to embed secret data for the purpose of increasing the embedding capacity. It can recover the original VQ coding, and also flexibility adjusts the embedding capacity. The strategies using side-match predictions based on (SMVQ) encoding are able to get better compression rates in comparison with Chang et al.'s method which uses the front one third of the codebook to embed secret data.

In [46], Wang, Chang et al. proposed a novel reversible data hiding scheme based on VQ images. The secret data are embedded into the VQ compression code (called VQ index table) by replacing the index values according to index differences. This method uses the principle component analysis (PCA). It can extract the principal component from all code words in a codebook, then; all code words are sorted according to the principal component. This method gives a high hiding capacity and bit rate.

In [47], Chang, Nguyen et al. proposed a new reversible data hiding scheme using a combination of side-match vector quantization (SMVQ) and search order coding (SOC). At first, the index table of the cover image is computed by using the traditional codebook of VQ. After that, this index table is changed into a transformed index table using the SMVQ where the secret data is embedded into. This method achieves a better compression rate bit of the image and a low execution time.

In [48], Lee, Chiou et al. proposed a novel lossless recovery data hiding method by exploiting the

natural feature of the high correlation among neighbouring blocks to hide secret data into VQ compressed image, and also to restore the original cover image after the secret data is extracted. This scheme can be classified into a variety of reversible VQ-based codes with control messages. VQ depends on the concept of comparison of each block in the image with all codewords of the codebook to find the nearest codeword. After that, the encoded word is determined through the nearest codeword index. Normally a VQ compressing have a high correlation and it can achieve a free space to hide secret data.

3.2 Irreversible Data Hiding

In this scheme, a secret message can be extracted, while the original cover images cannot be restored. The main goal of this scheme is to provide the ability of the secret message extracted from the stego-image and to increase the hiding capacity.

In [49], Chang and Wu presented an adaptive data-hiding method based on VQ to mitigate the limited embedding capacity of the fixed VQ-based data hiding method. Firstly, the codebook is partitioned into groups where a codeword of small index is chosen as a seed and all codewords that are close to the seed are merged together in a certain group. These groups have different numbers of codewords. Therefore, the embedding capacity is different and is based on the size of each group. The codeword-order-cycle technique is used to enhance the stego-image quality. The stego-image quality is very well when the embedding capacity is smaller than 16 kb.

4. TRANSFORM DOMAIN TECHNIQUES

The various transform domain techniques which include Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT) and Fast Fourier Transform (FFT) are used to hide information in transform coefficients of the cover images that make much more robust to attacks such as compression, filtering, etc.

In a 2-dimensional DCT phase, the image is partitioned into 8×8 non-overlapping block. Each block is transformed into the DCT coefficients by using the 2-dimensional DCT as given by [1]:

$$F(u,v) = \frac{c(u)c(v)}{4}$$

$$\sum_{t=0}^{7} \sum_{j=0}^{7} \cos\left(\frac{(2t+1)u\pi}{16}\right) \cos\left(\frac{(2j+1)v\pi}{16}\right) f(t,j) \quad (3)$$

$$\text{where } c(m) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } m = 0 \\ 1 & \text{if } m \neq 0 \end{cases}$$

$F(u,v)$ and $f(t,j)$ stand for a DCT coefficient at the (u,v) coordinate and a pixel value at the (t,j) coordinate, respectively. $F(0,0)$ is the DC component, which corresponds to an average intensity value of each block in the spatial domain. $F(u,v)$ is the AC component, in which $u \neq 0$ and $v \neq 0$. For data reduction during the quantization phase, DCT coefficients are quantized by using the standard quantization table.

The second approach is based on DWT. The 1D-DWT analysis separates between the high frequency and the low frequency information at each level of decomposition; therefore, only two sub-signals are produced at each level. Generally, images are represented by two dimensional signals that change spatially in horizontal and vertical directions, thus 2D-DWT analysis must be used for images. 2D-DWT analysis uses the same 'mother wavelets' as 1D-DWT but requires an extra step at each level of decomposition.

The block diagram shown in Fig.(4) illustrates the image decomposition based on 2D-DWT. Firstly, the image with M rows and N columns is filtered horizontally with low pass and high pass filters to obtain two sub-images with size of $M \times N/2$. Then, the outputs are filtered on columns to obtain four sub-images with $M/2 \times N/2$ size [50]. An example of image decomposition is shown in Figure 5.

In [51], Chin-Chen Chang et al. made a new steganography method based upon Joint Photographic Expert-Group (JPEG) and a quantization table modification. After modifying the quantization table process, the secret message is embedded in the middle-frequency area of the cover image. Then a JPEG stego-image is generated. This scheme is better than (Jpeg-Jsteg), which allows hiding a large message capacity.

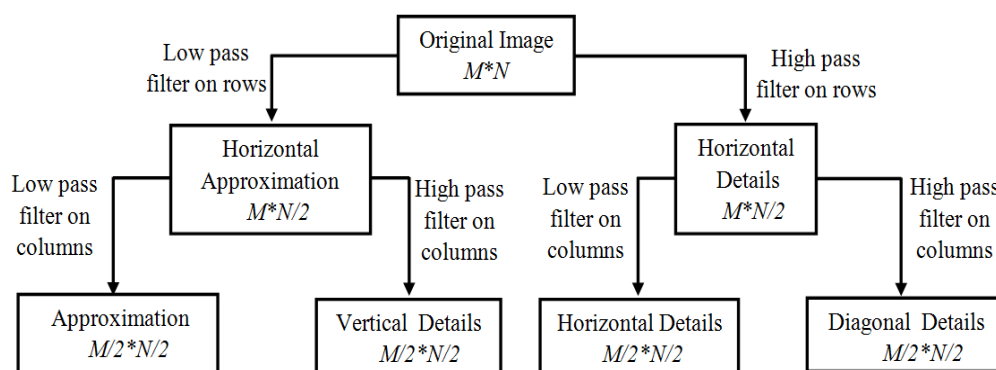


Figure 4: Block diagram of level one 2D-DWT [50].

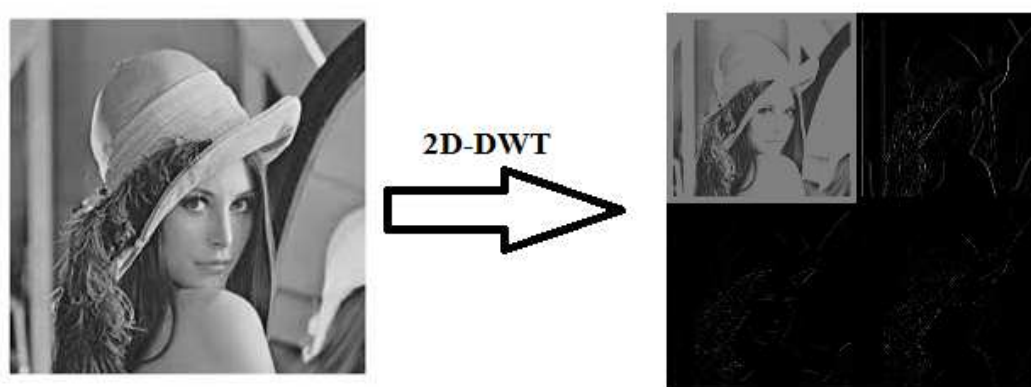


Figure 5: Image decomposition example with 2D-DWT.

In [52], Hideki Noda et al. proposed two JPEG steganography schemes using Quantization Index Modulation (QIM) in DCT domain. Firstly, Histogram Matching (HM-JPEG) method is used and it is based on histogram matching by using two quantizers with a dead zone. Secondly, QIM-JPEG technique is used in a straightforward path with a dead zone. This scheme embeds a secret message during quantization of DCT coefficients while the embedding based on other methods is obtained by modifying DCT coefficients after quantization. The two methods preserve the histogram of quantized DCT coefficients and offer rise in performance with regard to the embedding rate, PSNR of stego image, and particularly histogram preservation.

In [53], KokSheik Wong et al. presented a novel DCT based on a blind Mod4 steganography method for still images. The DCT coefficients are chosen as the valid message carrier from group of (2×2) spatially adjacent quantized. The next step is to apply modulus 4 arithmetic operation on the valid

GQC to embed two bits. To reduce distortion when modification is required for data embedding, the shortest route modification scheme is applied as compared to the ordinary direct modification technique. Mod4 method does not provide a high embedding ability and it achieves a minimum detection ratio against blind steganalyzers.

In [54], Chin-Chen Chang et al. proposed a reversible embedding secret data in each block of DCT coefficients based on compressed image technique JPEG. This method involves three stages. First, the cover image is partitioned into 8×8 pixel blocks. Second, the DCT is applied on each block. Third, the secret data are embedded into the successive zero sequence in the middle frequency components after quantizing the DCT coefficients. This scheme is used to increase the security of the embedding secret data as well as to improve hiding capacity.

In [55], Yung-Kuan Chan et al. presented a simple reversible data hiding technique using Haar Digital

Wavelet Transform (HDWT) method. HDWT is implemented by using two operations (horizontal and vertical) respectively, the first operation is decomposing an image into frequency band high (H) and low (L). The second operation is used to partition the image into LL, LH, HL, and HH of different frequencies respectively. The method of transforming the cover image from a spatial domain into a frequency domain is based on HDWT technique. The coefficients which have a high frequency band will be used to embed a secret data after applying a compression process on it by using the Huffman (or arithmetic) coding method. The receiver can recover the cover image without any distortion, and the stego-image is of high quality.

In [56], Chia-Chen Lin presented a high capacity data hiding scheme which is based on a notation transformation idea for DCT based images. The secret data are embedded in the middle frequency of the DCT coefficients where the coefficients in the middle frequency are divided into six sub-areas. However, this method increases the embedding capacity and stego-image quality when compared to the related works, but the reversibility is not considered as well.

In [57], Yih-Kai Lin proposed a reversible data hiding method which is based on the DCT of the cover image. The cover image is decomposed into different frequency parts, where the secret messages are embedded into the high-frequency parts. The secret data are embedded in the space around zero coefficient after shifting the histogram, in which the positive and negative coefficients are shifted to the right and left respectively. This method has obtained a high improvement when compared to past related work.

In [58], Hsiu-Lien Yeh et al. presented a method to hide a secret image into multi-stage compressed images. The embedding process involves DWT of the cover image, refined bit stream detection, and Set Combination Theory (SCT) for embedding, which is used to determine the relationship between subset distribution and the refined bit stream. An improvement is achieved in stego- image quality with the same hiding capacity of the related works.

In [59], Khamrui and Mandal proposed a Genetic Algorithm using DCT based steganography technique (GASDCT). This method uses four steps to conceal secret information: transformation into frequency domain using DCT, hiding secret data, re-transformation into spatial domain using Inverse Discrete cosine Transformation (IDCT) and finally applying Genetic Algorithm (GA). From the source gray scale image, the sub mask is selected (2*2 nonoverlapping) in row major order and DCT is

applied to it to produce four steps of frequency components. In each lower frequency components two bits of the secret information are embedded into the second and third position. In such a way of embedding procedure, that is no loss during the reversed transformation chosen. After that, IDCT is applied to the resultant image of sub mask to convert it to a spatial domain. GA proposed obtains high image integrity, PSNR and a large capacity of secret image.

In [60], Mohamed Amin et al., proposed a new steganography scheme based on DCT coefficient of host image by using a new quantization technique. The quantization process of DCT coefficients is scaled by applying a predefined mathematical operation. The secret messages are hidden in the 2LSB of the DCT coefficient for all blocks except for the last block which is booked to embed the message size. The proposed method has improved the amount of the secret message that will be imbedded to keep the stego-image of a high quality.

5. SPREAD SPECTRUM TECHNIQUE

Spread Spectrum Image Steganography (SSIS) technique stores a secret message as Gaussian noise in the cover image [61]. The main idea of SSIS is that of spreading the secret message over a large frequency bandwidth under the noise level. SSIS provides the ability to hide and recover, error free, a significant quantity of information bits within digital images, voiding detection by an observer. Furthermore, SSIS is a blind scheme because the original image is not needed to extract the hidden information. The proposed recipient needs only to possess a key in order to reveal the hidden message. The very existence of the hidden information is virtually undetectable.

6. TSTATISTICAL TECHNIQUE

In this method, the information is embedded in the cover image by changing the statistical properties of the cover image and a hypothesis testing is used in the extraction process. The cover image is divided into blocks and the message bits are hidden in each block. The cover image block is modified if the message bit is "1", otherwise the block unchanged [62]. This method is not preferable because it is easily and it can be detected by steganalysis techniques such as the first order statistics or the second order statistics.

7. DISTORTION TECHNIQUE

The secret data is stored by signal distortion. The encoder makes changes to the cover image based on the secret data and the decoder checks for the differences between the original cover image and the stego-image to restore the secret data. Therefore, the decoder needs knowledge about the original cover image during the decoding process [63]. According to the secret message, a sequence of modifications is selected in the encoding process. The message is embedded in random pixels of the cover image. The receiver restores a message bit of "1" when the stego-image is different from the cover image at the given message pixel. Otherwise, the message bit will be "0".

8. CONCLUSION

According to the cover medium type, steganography can be classified into text, image, video, and audio. The most commonly used is the

image steganography due to the vast use of images through the Internet and the difficulty of figuring out the difference between the original image and the stego-image by human eyes. In this paper, a literature survey on the image steganography techniques is presented. Three approaches are mostly used than the others and they include spatial domain, compressed domain based on VQ, and transform domain. Spatial domain based on LSB has a simple implementation and high capacity with low robustness against signal processing techniques such as filtering, cropping, and modification. Compressed domain based on VQ is more robust than the spatial domain but the hiding capacity is low. Transform domain based on DCT or DWT is the most robust among the others in that it has high resistance to signal processing techniques while the hiding capacity is low. As shown in Table I, there is no ideal technique that can be used but a trade off must be made between the steganography parameters (capacity, robustness, and security) to obtain a suitable result for a certain application.

TABLE I: A comparison of image steganography techniques.

	Spatial domain	Compressed domain based on VQ	Transform domain	Spread spectrum	Statistical techniques	Distortion techniques
Imperceptibility	High*	High	High	High	Medium*	Low
Robustness	Low	Medium	High	Medium	Low	Low
Payload capacity	High	Low	Low	High	Low*	Low

*: Indicates dependency on the used cover image

REFERENCES:

- [1] Provos, N., & Honeyman, P. (2003). Hide and seek: An introduction to steganography. *Security & Privacy, IEEE*, 1(3), 32-44.
- [2] Silman, J. (2001). Steganography and steganalysis: an overview. *SANS Institute*, 3, 61-76.
- [3] Tayana, M. (2012). Image Steganography applications for secure communications. Msc. Thesis.
- [4] Hamid, N., Yahya, A., Badlishah Ahmad, R., Najim, D., & Kanaan, L. (2013). Steganography in image files: A survey. *Australian Journal of Basic & Applied Sciences*, 7(1).
- [5] Chang, C.-C., G.-M. Chen and M.-H. Lin (2004). "Information hiding based on search-order coding for VQ indices." *Pattern Recognition Letters* 25(11): 1253-1261.
- [6] Chan, C.-K. and L.-M. Cheng (2004). "Hiding data in images by simple LSB substitution." *Pattern recognition* 37(3): 469-474.
- [7] Hussain, M., & Hussain, M. (2013). A Survey of Image Steganography Techniques. *International Journal of Advanced Science & Technology*, 54.
- [8] Chao, H.-M., C.-M. Hsu and S.-G. Miaou (2002). "A data-hiding technique with authentication, integration, and confidentiality for electronic patient records." *Information Technology in Biomedicine, IEEE Transactions on* 6(1): 46-53.

- [9] Thien, C.-C. and J.-C. Lin (2003). "A simple and high-hiding capacity method for hiding digit-by-digit data in images based on modulus function." *Pattern Recognition* 36(12): 2875-2881.
- [10] Wu, H.-C., N.-I. Wu, C.-S. Tsai and M.-S. Hwang (2005). "Image steganographic scheme based on pixel-value differencing and LSB replacement methods." *IEE Proceedings-Vision, Image and Signal Processing* 152(5): 611-615.
- [11] Yu, Y.-H., C.-C. Chang and Y.-C. Hu (2005). "Hiding secret data in images via predictive coding." *Pattern Recognition* 38(5): 691-705.
- [12] Li, S.-L., K.-C. Leung, L. Cheng and C.-K. Chan (2006). "A novel image-hiding scheme based on block difference." *Pattern Recognition* 39(6): 1168-1176.
- [13] EL-Emam, N. N. (2007). "Hiding a large amount of data with high security using steganography algorithm." *Journal of Computer Science* 3(4): 223.
- [14] Yu, Y.-H., C.-C. Chang and I.-C. Lin (2007). "A new steganographic method for color and grayscale image hiding." *Computer Vision and Image Understanding* 107(3): 183-194.
- [15] Lin, C.-C. and N.-L. Hsueh (2008). "A lossless data hiding scheme based on three-pixel block differences." *Pattern Recognition* 41(4): 1415-1425.
- [16] Chang, C.-C., C.-Y. Lin and Y.-H. Fan (2008). "Lossless data hiding for color images based on block truncation coding." *Pattern Recognition* 41(7): 2347-2357.
- [17] Yang, C.-H. (2008). "Inverted pattern approach to improve image quality of information hiding by LSB substitution." *Pattern Recognition* 41(8): 2674-2683.
- [18] Yang, C.-H., C.-Y. Weng, S.-J. Wang and H.-M. Sun (2008). "Adaptive data hiding in edge areas of images with spatial LSB domain systems." *Information Forensics and Security, IEEE Transactions on* 3(3): 488-497.
- [19] Lee, Y., H. Kim and Y. Park (2009). "A new data hiding scheme for binary image authentication with small image distortion." *Information Sciences* 179(22): 3866-3884.
- [20] Tsai, P., Y.-C. Hu and H.-L. Yeh (2009). "Reversible image hiding scheme using predictive coding and histogram shifting." *Signal Processing* 89(6): 1129-1143.
- [21] Wien Hong, Tung-Shou Chen, Chih-Wei Shiu. Reversible data hiding for high quality images using modification of prediction errors. *The Journal of Systems and Software* 82 (2009) 1833-1842.
- [22] Hong, W. and T.-S. Chen (2010). "A local variance-controlled reversible data hiding method using prediction and histogram-shifting." *Journal of Systems and Software* 83(12): 2653-2663.
- [23] Chang, C.-C. (2010). "A reversible data hiding scheme using complementary embedding strategy." *Information Sciences* 180(16): 3045-3058.
- [24] Yang, C.-H., C.-Y. Weng, S.-J. Wang and H.-M. Sun (2010). "Varied PVD+ LSB evading detection programs to spatial domain in data embedding systems." *Journal of Systems and Software* 83(10): 1635-1643.
- [25] Chen, W.-J., C.-C. Chang and T. Le (2010). "High payload steganography mechanism using hybrid edge detector." *Expert Systems with Applications* 37(4): 3292-3301.
- [26] Weiqi Luo, Fangjun Huang and Jiwu Huang. Edge Adaptive Image Steganography Based on LSB Matching Revisited. *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 5, NO. 2, JUNE 2010*.
- [27] Liao, X., Q.-y. Wen and J. Zhang (2011). "A steganographic method for digital images with four-pixel differencing and modified LSB substitution." *Journal of Visual Communication and Image Representation* 22(1): 1-8.
- [28] Sun, H.-M., C.-Y. Weng, C.-F. Lee and C.-H. Yang (2011). "Anti-forensics with steganographic data embedding in digital images." *Selected Areas in Communications, IEEE Journal on* 29(7): 1392-1403.
- [29] Yang, C.-H., C.-Y. Weng, H.-K. Tso and S.-J. Wang (2011). "A data hiding scheme using the varieties of pixel-value differencing in multimedia images." *Journal of Systems and Software* 84(4): 669-678.
- [30] Chen, S.-K. (2011). "A module-based LSB substitution method with lossless secret data compression." *Computer Standards & Interfaces* 33(4): 367-371.
- [31] Hao Luo, Fa-Xin Yu, Hua Chen, Zheng-Liang Huang, Hui Li, Ping-Hui Wang. Reversible data hiding based on block median preservation. *Information Sciences* 181 (2011) 308-328.
- [32] Hong, W. and T.-S. Chen (2012). "A novel data embedding method using adaptive pixel

- pair matching." *Information Forensics and Security, IEEE Transactions on* 7(1): 176-184.
- [33] Lou, D.-C. and C.-H. Hu (2012). "LSB steganographic method based on reversible histogram transformation function for resisting statistical steganalysis." *Information Sciences* 188: 346-358.
- [34] Hong, W. (2012). "Adaptive image data hiding in edges using patched reference table and pair-wise embedding technique." *Information Sciences*.
- [35] Hong, W., T.-S. Chen and C.-W. Luo (2012). "Data embedding using pixel value differencing and diamond encoding with multiple-base notational system." *Journal of Systems and Software* 85(5): 1166-1175.
- [36] Amirtharajan, R. and J. B. Balaguru Rayappan (2012). "An intelligent chaotic embedding approach to enhance stego-image quality." *Information Sciences* 193: 115-124.
- [37] Bedi, P., R. Bansal and P. Sehgal (2013). "Using PSO in a spatial domain based image hiding scheme with distortion tolerance." *Computers & Electrical Engineering*.
- [38] H.Y. Leung, L.M. Cheng, F. Liu, Q.K. Fu. Adaptive reversible data hiding based on block median preservation and modification of prediction errors. *The Journal of Systems and Software* 86 (2013) 2204–2219.
- [39] Yuan, H.-D. (2014). "Secret sharing with multi-cover adaptive steganography." *Information Sciences* 254: 197-212.
- [40] Linde, Y., Buzo, A., & Gray, R. M. (1980). An algorithm for vector quantizer design. *Communications, IEEE Transactions on*, 28(1), 84-95.
- [41] Lee, C. C., Ku, W. H., & Huang, S. Y. (2009). A new steganographic scheme based on vector quantization and search-order coding. *Image Processing, IET*, 3(4), 243-248.
- [42] Hu, Y.-C. (2006). "High-capacity image hiding scheme based on vector quantization." *Pattern Recognition* 39(9): 1715-1724.
- [43] Chang, C.-C., W.-C. Wu and Y.-C. Hu (2007). "Lossless recovery of a VQ index table with embedded secret data." *Journal of Visual Communication and Image Representation* 18(3): 207-216.
- [44] Yung Kuei Chiang, Piyu Tsai. Steganography using overlapping codebook partition. *Signal Processing* 88 (2008) 1203–1215.
- [45] Yang, C.-H. and Y.-C. Lin (2009). "Reversible data hiding of a VQ index table based on referred counts." *Journal of Visual Communication and Image Representation* 20(6): 399-407.
- [46] Wang, Z.-H., C.-C. Chang, K.-N. Chen and M.-C. Li (2010). "An encoding method for both image compression and data lossless information hiding." *Journal of Systems and Software* 83(11): 2073-2082.
- [47] Lee, J.-D., Y.-H. Chiou and J.-M. Guo (2012). "Lossless data hiding for VQ indices based on neighboring correlation." *Information Sciences*.
- [48] Chang, C.-C., T. S. Nguyen and C.-C. Lin (2012). "A novel VQ-based reversible data hiding scheme by using hybrid encoding strategies." *Journal of Systems and Software*.
- [49] Chang, C. C., & Wu, W. C. (2006). Hiding secret data adaptively in vector quantization index tables. *IEE Proceedings-Vision, Image and Signal Processing*, 153(5), 589-597.
- [50] Lees, K. (2002). Image compression using Wavelets. Report of MS.?
- [51] Chin-Chen Chang, Tung-Shou Chen, Lou-Zo Chung. "A steganographic method based upon JPEG and quantization table modification". *Information Sciences* 141 (2002) 123–138
- [52] Hideki Noda, Michiharu Niimi, Eiji Kawaguchi. "High-performance JPEG steganography using quantization index modulation in DCT domain". *Pattern Recognition Letters* 27 (2006) 455–461.
- [53] KokSheik Wong, Xiaojun Qi, Kiyoshi Tanaka. "A DCT-based Mod4 steganographic method". *Signal Processing* 87 (2007) 1251–1263.
- [54] Chin-Chen Chang, Chia-Chen Lin, Chun-Sen Tseng, Wei-Liang Tai. "Reversible hiding in DCT-based compressed images". *Information Sciences* 177 (2007) 2768–2786.
- [55] Yung-Kuan Chan, Wen-Tang Chen, Shyr-Shen Yu, Yu-An Ho, Chwei-Shyong Tsai, Yen-Ping Chu. "A HDWT-based reversible data hiding method". *The Journal of Systems and Software* 82 (2009) 411–421.
- [56] Chia-Chen Lin, Pei-Feng Shiu. "High Capacity Data Hiding Scheme for DCT-based Images". *Journal of Information*

- Hiding and Multimedia Signal Processing*,
Volume 1, Number 3, July 2010.
- [57] Yih-Kai Lin. "High capacity reversible data hiding scheme based upon discrete cosine Transformation". *The Journal of Systems and Software* 85 (2012) 2395– 2404.
- [58] Hsiu-Lien Yeh, Shu-Tsai Gue, Piyu Tsai, Wei-Kuan Shih. " Wavelet bit-plane based data hiding for compressed images". *Int. J. Electron. Commun. (AEÜ)* 67 (2013) 808– 815.
- [59] Amrita Khamruia, J K Mandal. "A Genetic Algorithm based Steganography using Discrete Cosine Transformation (GASDCT)". *Procedia Technology* 10 (2013) 105 – 111.
- [60] Mohamed Amin, Hatem M. Abdulkader, Hani M. Ibrahim, and Ahmed S. Sakr. "A Steganographic Method Based on DCT and New Quantization Technique". *International Journal of Network Security*, Vol.16, No.3, PP.214-219, May 2014.
- [61] Marvel, L. M., Boncelet, C. G., Jr., & Retter, C. T. (1999). Spread spectrum image steganography. *Image Processing, IEEE Transactions on*, 8(8), 1075-1083.
- [62] Kruus, P., Scace, C., Heyman, M., & Mundy, M. (2003). A survey of steganographic techniques for image files. *Advanced Security Research Journal*, V(I), 41-52.
- [63] Reddy, H. S. M., & Raja, K. B. (2009). High capacity and security steganography using discrete wavelet transform. *International Journal of Computer Science and Security (IJCSS)*, 3(6), 462-472.