

A Novel Information Security System based on Steganography and Compression Techniques for Higher Education Institutions

Hosnia M. M. Ahmed
Department of Computer Science Mansoura
University, Mansoura, Egypt

Ahmed A. A. Kamel
Department of Computer Science Mansoura
University, Mansoura, Egypt

ABSTRACT

With the advancement of technology and the methods of information storage and exchange, the security of information and data became a vital issue for many institutions, especially higher education institutions. This paper introduces a novel and efficient information security system based on steganography and compression techniques to address the issues related to the security and efficient transmission of secret data in higher education institutions. In this system, the Lempel Ziv Welch (LZW) compression technique is used to compress the data to be hidden, to reduce their size and thereby reduce the percentage of distortion that may occur to the cover image. In addition, a new data hiding algorithm has been developed depending on the second bit, where data are hidden based on LSB without making any changes to the values of the second bit. If any changes are made, these will be only within LSB, which reduces the distortion in the cover image and increases the similarity between the cover image before and after hiding. In the retrieval stage, hidden data are retrieved from the second bit only. The proposed system performance was evaluated in terms of MSE (mean square of error), PSNR (peak signal-to-noise ratio), PD (pixel distortion) and BER (bit error rate). The experimental results confirm the efficiency of the proposed system that achieves good results with higher PSNR and lower MSE.

General Terms

Information Security

Keywords

Steganography, Data Compression, LZW Algorithm

1. INTRODUCTION

Data and information constitute a wealth for any institution, especially higher education ones, to conduct several academic, research and administrative activities, which need to be transmitted securely through insecure transmission methods [1]. With the rapid advancement of information and internet technology, this technology has become an important tool for transferring and interchanging all kinds of multimedia information such as text, audio, video and images [2]. Moreover, the exchange of communication and information has become much easier and quicker. However, the issues connected with the security and confidentiality of information in an open network environment have become a major cause of concern today [3]. The question here is how to transfer secret or private data in a secure way, especially in higher education institutions. Steganography is one of the most important techniques used to secure information [4]. Steganography is the art and science of hiding secret data within seemingly innocent carriers, which can be employed to achieve covert communications [5, 6]. A steganographic

system hides secret data into cover media (like text, image, audio, and video) and generates a stego-media [2], which is then sent to the receiver without anyone else knowing that it contains the hidden data [7]. The receiver retrieves the hidden data by applying the de-steganography process. A stego-key is used for the embedding or encoding process to prevent decoding or extraction of data embedded in cover media [3]. This stego-key is the same as the one used by the sender. Thus, security is achieved by hiding the existence of the message [7]. The main objectives of steganography are imperceptibility, robustness, capacity of the hidden message, and resistance to tampering [8, 9]. Imperceptibility refers to hiding data in such a way that the cover media can be discovered only by the receiver [10]. Robustness is the level of difficulty required to destroy embedded information without destroying the cover media and extracting the required information [11]. Capacity means the amount of information or data that can be embedded without affecting the imperceptibility of cover media. Therefore, the embedding capacity remains a factor that proves the importance of steganography [10, 11]. Although these three objectives are much interrelated, they should meet within steganography without disturbing each other [10]. Several steganographic techniques have been used to ensure the security of data. These techniques can be classified into two domains: spatial domain techniques and transform domain (or frequency domain) techniques [3]. In transform domain technique, the cover media is transformed into coefficients with the use of the well-known transform techniques (the integer wavelet transform and the integer discrete cosine transform). After that, the message is embedded in the media [7, 12]. In spatial domain techniques, the message bits are encoded directly, which results in few changes in the intensities of the sample. This, in turn, results almost in no perceivable changes in the cover media [13]. The Least Significant Bit (LSB) method is the most widely used method for spatial-domain-based steganography [4, 14]. This method is the simplest and best-known way for embedding secret data [15]. In this method, the least significant bits of some or all of the bytes inside a cover media are replaced with bits of the secret message [16]. This paper presents another important technique, which is data compression. The purpose of data compression is to reduce the quantity of data or the redundancy of the image and to store or transmit data in an efficient form [12, 17]. Compression techniques can be applied to digital content to ensure communication efficiency and save network bandwidth conveniently [18-21]. Data are compressed in two schemes: the lossy and lossless schemes. Although both schemes save storage space, they differ in the procedures each applies [22, 23]. In lossless compression, original data can be rebuilt without any loss, while in lossy compression a part of the data is lost in the quantization process [17, 23]. In the proposed

work, to increase capacity, LZW data compression technique is used for compressing the secret data. The proposed work presents a novel technique that integrates steganography with data compression into a single module seamlessly to solve the problems related to the security and efficient transmission of secret data. This will increase the efficiency of transmission. The remaining part of this paper is organized as follows: section 2 discusses related work in this research area, section 3 describes the proposed system, section 4 analyses the performance of the proposed system, while section 5 presents the conclusion.

2. RELATED WORK

To protect medical data, **R. Karakis and et al** [8] transformed them into a single file format with the use of steganographic methods, choosing an electroencephalogram (EEG) to be the hidden data. In addition, image headers contained the doctor's comments and patient information. Magnetic resonance (MR) images were used to hide data. Two new image steganography methods were proposed using fuzzy-logic and similarity. These were used to choose image pixels non-sequential LSB. To embed the message, they used similarity values of the gray levels in the pixels. Then, the message was protected from attacks using lossless compression and symmetric encryption algorithms. They used MSE, PSNR, and structural similarity measure (SSIM), universal quality index (UQI) and correlation coefficient (R) to calculate the quality of the stego image. Their findings show that the suggested method ensured the secrecy of patient information, increasing data volume and transmission ability of both MR images and EEG signals.

M. Nazrul Islam, M. Faysal Islam and K. Shahrabi [24] proposed an innovated and efficient system for information security to protect biometric signatures from unauthorized access. They used individual orthogonal codes to encode several biometric signatures. After that, they multiplexed these codes together and embedded them within a cover image. They used a new steganography technique for this. First, a colored cover image was decomposed into its red, green and blue colors. Each of these components was used to hide one group of biometric signatures. They used another secret key to choose a bit from the three LSBs in the cover image. After that, the selected bit was replaced by the information bit and constituted the stego image by combining stego images together. They used the MRJTC technique to encrypt the color stego image again, which greatly increased security against unauthorized access. Any steganalysis method can easily fail to recover any information. This shows that the steganography technique is very efficient in making the information completely hidden. In addition, the orthogonal encoding method increased information security by making accessing the biometric data without authorization nearly impossible. Although their findings were limited to the four biometric signatures, the orthogonal coding scheme can be theoretically extended to any amount of information. Simulation experiments prove that the proposed method can produce an effective system to protect any kind of information including biometric signatures, personal identification information, as well as confidential documents. Real-time identity verification is also one of the fields where the information security system can be used.

A. Subash and N. George [21] presented an improved integrated scheme to hide and compress digital images. Their scheme uses vector quantization (VQ) and side match. The data could be hidden and images could be compressed in the

same module. Secret data could be embedded in the blocks and compressed at the same time using VQ or side match, except for non-residual blocks (i. e., the blocks located in the leftmost and topmost of the image). They calculated the similarity between neighboring blocks using side match. They encoded residual blocks using the redundancy between the image blocks and applied the Huffman encoding method to the resulting compressed code stream to improve compression more. They divided the compressed codes of the image resulting from Huffman decoding into sections by indicator bits. Thus, the receiver could extract secret bits and compress the image successfully. As the proposed technique uses the redundancy of blocks, it could embed two secret data bits with each side match encoded blocks. In addition, the proposed system applies a procedure to hide and compress data without using a threshold. According to experimental results, the performance of this scheme is better in hiding capacity, compression ratio and decompression quality.

S. Bader and Noufal P. [25] used a new method that hides and compresses data for more efficient communication. Data hiding and compression were performed in a single unit, thereby increasing the efficiency of communication. At first, they segmented the image into non-overlapping blocks. Then, they processed the image. They used VQ to compress the topmost and leftmost blocks. They embedded the remaining blocks with hidden data and compressed them at the same time using SMVQ or image inpainting in an adaptive way, using the secret bit to be hidden as a basis for this. They used VQ with some blocks as a threshold, to control the distortion. The received linked codes were divided into indices and secret bits according to the indicator bits. The receiver used FoE-based inpainting to recreate the lost parts of the image. According to experimental results, the performance of the scheme was satisfactory as regards compression ratio and image quality.

Rupali Bhardwaj and Vaishali Sharma. [3] Proposed a technique to achieve three levels of security. The secret message was first complemented and then hidden in the cover image pixels that were selected randomly using pseudo-random number generator. Finally, as a steganographic technique, they used the inverted bit LSB method, rather than the simple LSB. This limited the detectability of the hidden message. They used MSE and PSNR to evaluate the difference between the cover-image and the stego-image. According to the findings of their study, the proposed technique outperformed the simple LSB and inverted LSB methods. PSNR was higher and MSE was lower. The proposed technique is better than the basic LSB method as regards higher visual quality shown by high PSNR values of hiding secret message bits in the image. This reduces the detectability of the secret message and makes secret communication possible.

3. PROPOSED SYSTEM

In this paper, a novel information security system is presented using the seamless integration of steganography and data compression into a single module that is utilized to achieve data compression and secret data embedding at one and the same time. The proposed system can be mainly divided into two phases. The first phase includes data compression and secret data embedding, while the second phase includes data decompression and secret data extraction. Fig. (1) Illustrates the framework of the proposed system.

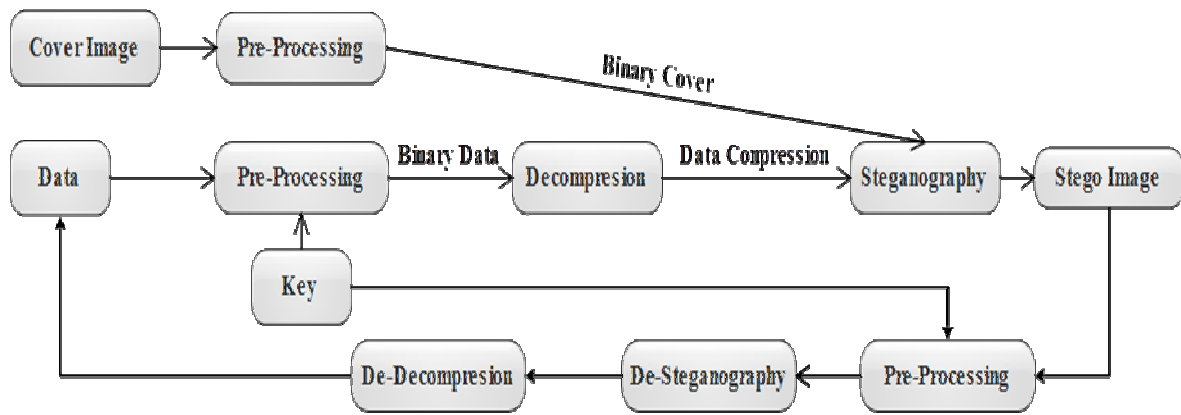


Fig.1. Proposed System Framework

In the proposed system, the issues related to capacity and security has been addressed to achieve efficient and fast transmission. The objective of compression is to reduce the amount of data to be hidden, while the objective of steganography is to improve data security. In the following, the operation of these two phases during the embedding and extraction procedures is detailed.

First phase: Data Compression and Secret Data Embedding

The embedding in the first phase relies on using three steps to embed secret bits in the image simultaneously to produce stego-image.

Step1: Data Pre-Processing

In order to generate secret data for embedding, the key was first converted into ASCII code. Then, the data that the user wants to embed were converted into the same into ASCII code and then into binary pattern. Thus, the final secret bit sequence $S = \{s_1, s_2, \dots, s_n\}$ is obtained.

Step2: Data Compression

The LZW algorithm is directly applied to the final secret bit sequence and the obtained bit stream is hidden into the cover image. Thus, the compressed final secret bit $Z = \{z_1, z_2, \dots, z_n\}$ is obtained as shown in Fig (2). IF $Z \leq 255$, it will be represented in 8 bits. Otherwise, $(Z \bmod 256)$ will be represented in 8 bits and the value of the MSB is 1.

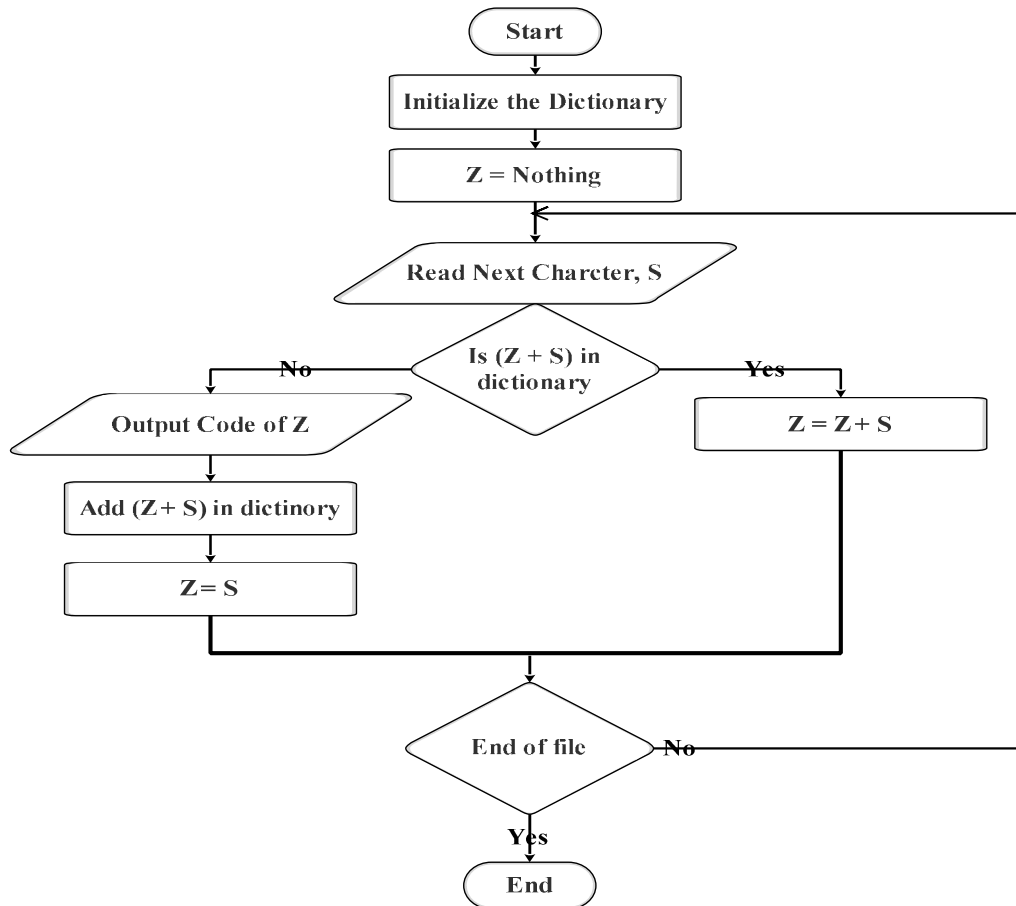


Fig.2. LZW algorithm

Step3: Steganography

The embedding step starts by preprocessing the cover image to generate a new cover-image. The cover image pixels are then converted into a binary pattern. The image is now used as a cover to embed secret information. This process is done by the LSB encoder, which replaces the least significant bit of pixel values with secret information bits. The modified image is now termed as stego-image. The LSB embedding scheme is used to hide the final compressed secret data in image files. Hiding data in any bit of the cover has the same efficiency of hiding them in LSB. The steps for embedding process are described in the following algorithm:

<u>Steganography Algorithm</u>
<p><i>If</i> $Bit_z = Bit_2$ <i>Then</i> $Bit_1 = 0$</p> <p><i>Else If</i> $Bit_z = 0$ <i>and</i> $Bit_2 = 1$ <i>Then</i> $Bit_2 = Bit_z$ $Modify\ Byte_i = Original\ Byte_i + (2)_2$ $Bit_1 = 1$</p> <p><i>Else</i> $Bit_z = Bit_2$ $Modify\ Byte_i = Original\ Byte_i - (2)_2$ $Bit_1 = 1$</p> <p><i>End If</i></p>
<p><i>Where:</i></p> <p>Bit_z: <i>The message to be hidden</i></p> <p>$Original\ Byte_i$: <i>The cover before hiding</i></p> <p>Bit_2: <i>The bit where data will be hidden</i></p> <p>$Modify\ Byte_i$: <i>The cover after hiding</i></p>

For example, if the compressed final secret datum to be hidden is (01101100) and the cover datum is (01010110), then, the hiding process will result in (01010111). It is noted that the value of the second bit didn't change, whereas the LSB changed. After finishing the hiding process, the cover is obtained. Its bits are divided into groups of 8 bits, converted into an ASCII code and saved as an image.

Phase two: data decompression and secret data extraction

The second phase of the proposed system is the reverse process of embedding secret data. After receiving the stego-image from the sender, the receiver can extract secret digits and restore the cover image. The steps for extraction phase are described in the following sections.

Step1: Data Pre-Processing

Data-preprocessing is the first step in the extraction phase. The stego-image and the key are read and converted into ASCII and then into the binary pattern.

Step2: De-Steganography

The de-steganography procedures in the proposed system depend on applying the following algorithm to the binary pattern to extract the data hidden in the stego image.

<u>De-Steganography Algorithm</u>
If $Bit_1 = 0$ Then
$Bit_z = Bit_2$
Else
$Bit_z = 1 - Bit_2$
End If

Step3: Data De-Compression

The LZW algorithm is applied to the extracted data using the same algorithm depicted in Fig.2, to decompress the bit stream to get the original secret data. Finally, the original secret data is converted into ASCII code then into their corresponding characters. Thus, the original secret data are obtained correctly and the cover image can be obtained successfully.

4. EXPERIMENTAL RESULTS

This section discusses the experimental results of the proposed system. The system is completely implemented using the C#.NET2017 programming environment. Experiments were conducted on a group of test images, to verify the efficiency of the proposed system. These test images are Lena, Baboon, Pappers and F16. In this study, the performance of the proposed system was measured between the cover and stego-image using four different comparison methods. These include: MSE, PSNR, BER and PD, which are given as following in equations (1) - (4), respectively [14, 26].

$$MSE = \frac{1}{n} \sum_{i=1}^n (p_i - p'_i)^2, \tag{1}$$

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right), \tag{2}$$

$$PD = \frac{\text{total number of bits changes}}{\text{total number of bits embedded}} \tag{3}$$

$$BER = \frac{\text{total number of bit changes}}{\text{total number of bits}} \tag{4}$$

Where n is the total number of pixels, p_i is the original pixel value in the cover-image and p'_i the modified pixel value in the stego-image. Table 1 shows the results of applying the proposed algorithm to the cover and stego- images.

Table.1: MSE, PSNR, PD AND BER of cover- image and stego-image









Cover Image	Stego Image	MSE	PSNR	PD	BER
		0.03832	51.296	0.05	0.00
Lena					
		0.0406	53.045	0.1	0.00
Baboon					
		0.0601	51.877	0.075	0.00
Pappers					
		0.0305	53.034	0.125	0.00
F16					

Table1 shows that the quality of the stego- image is good. This is due to the high embedding efficiency. In addition, high SPNR values were obtained after hiding secret data in the test images using the proposed system. The higher PSNR values of the proposed system reflect the good quality of the resulting image. Thus, the proposed system is proved to provide enhanced security. From this comparison, we can conclude that the proposed system can generate a resulting image that is closest to the original cover-image, with a low level of distortion. The experimental results confirm the efficiency of the proposed system that achieves good results with higher PSNR and lower MSE.

5. CONCLUSION

Information security is an important research field in data communication, especially in higher education institutions. In this paper, an integrated steganography and compression techniques were implemented in a single module to increase data security and communication efficiency, while they are transmitted over networks. In this system, the Lempel Ziv Welch (LZW) compression technique is used to compress the data to be hidden, to reduce their size and thereby reduce the percentage of distortion that may occur to the cover image. In addition, a new data hiding algorithm has been developed depending on the second bit, where data are hidden based on the LSB. Data are hidden without making any changes to the values of the second bit. If any changes are made, then, these will be only within LSB, which reduces the distortion in the cover image and increases the similarity between the cover image before and after hiding. In the retrieval stage, hidden data are retrieved from the second bit only. The proposed system performance was measured in terms of MSE, PSNR, PD and BER. As shown in the experimental results, the results demonstrate the efficiency of the proposed system. The proposed system can be applied to all kinds of multimedia information like audio, video.

6. REFERENCES

- [1] K.B.Sudeepa .et al. 2016. A New Approach for Video Steganography Based on Randomization and Parallelization , International Conference of Information Security and Privacy (ICISP), Procedia Computer Science , Vol.78, PP.483-490.
- [2] A. Malik, GeetaSikka and H. k.Verma. 2016. A high Capacity Text Steganography Scheme based on LZW Compression and color coding, Engineering Science and Technology an International Journal, Vol.20, issue.1, PP.72-79.
- [3] R.Bhardwaj and V.Sharma. 2016. Image Steganography Based on Complemented Message and Inverted bit LSB Substitution, Procedia Computer Science, 6th International Conference on Advances In Computing & Communications, ICACC 2016, 6-8, Vol.93, PP.832-838.
- [4] M.Ramalingam and NorAshidi M. Isa. 2015. A data – hiding technique using scene-change detection for video steganography, Computers and Electrical Engineering , Vol.000, PP.1-12.
- [5] H.Tian and et al. 2015. Optimal matrix embedding for Voice-over-IP steganography, Signal Processing, Vol.117, PP.33-43.
- [6] R. Moradi Rad, K. Wong and J. Ming. Guo. 2016. Reversible data hiding by adaptive group modification on histogram of prediction errors, Signal Processing, Vol.125, PP.315-328.
- [7] R. Jain and et al. 2012. Efficient data hiding scheme using lossless data compression and image steganography, International Journal of Engineering Science and Technology, Vol.4, No.08, PP.3908-3915.
- [8] R.Karakis and et al. 2015. A novel fuzzy logic-based image steganography method to ensure medical data security, Computers in Biology and Medicine, Vol.67, PP.171-183.
- [9] A. Saleema and T.Amarunnishad. 2016. A new Steganography Algorithm Using Hybrid Fuzzy Neural Networks, Procedia Technology, International Conference of Emerging Trends in Engineering, Science and Technology, Vol.24, PP.1566-1574.
- [10] B.Datta, U.Mukherjee and S. K. Bandyopadhyay. 2016. LSB layer independent Robust Steganography using Binary Addition, Procedia Computer Science, International Conference on Computational Modeling and Security, Vol.85, PP.425-432.
- [11] S.Mungmode, R.R.Sedamkar and N.Kulkarni. 2016. A Modified High Frequency Adaptive Security Approach using Steganography for Region Selection based on Threshold Value,Procedia Computer Science, 7th International Conference on Communication, Computing and Virtualization , Vol.79, PP.912-921.
- [12] K.Ramya .K.Kathiresan and G.T.Kalaiarasi. 2014. A Survey on Data Hiding and Compression Scheme, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.3, Issue .11, PP.13132-13134.
- [13] S.A.El_Rahman. 2016. A comparative analysis of image steganography based on DCT algorithm and steganography tool to hide nuclear reactors confidential information, Computers and Electrical Engineering, Vol.000, PP.1-20.
- [14] T.Tuncer and E.Avci. 2016. A reversible data hiding algorithm based on probabilistic DNA-XOR secret sharing scheme for color images, Displays, Vol.14, PP.1-8.
- [15] W.Chung Kuo and et al. 2016. High capacity data hiding scheme based on multi-bit encoding function, Optik-International Journal for Light and Electron Optics, Vol.17, issue.4, PP.1762-1769.
- [16] S. Gupta, A. Goyal and B. Bhushan. 2012. Information Hiding Least Significant Bit Steganography and Cryptography, I.J.Modern Education and Computer Science, Vol.6, PP.27-34.
- [17] M.M.Shanthi Rani and S.Lakshmanan. 2016. AN INTEGRATED METHOD OF DATA HIDING AND COMPRESSION OF MEDICAL IMAGES, International Journal of Advanced Information Technology, Vol.6, No.1, PP.43-51.
- [18] A. Mohan and M.Nasseena N. 2016. AN EFFICIENT JOINT DATA HIDING AND COMPRESSION TECHNIQUE, International Journal of Engineering Research and General Science, Vol.4, Issue.3, PP.727-733.
- [19] A.S.Shankari and J.Shanthini. 2015. Enhance Novel Joint Data-Hiding and Compression Scheme Based on exemplar approach, International Research Journal of Engineering and Technology, Vol.02, Issue.08, PP.136-139.
- [20] P. .M.Patil and V.R.Udupi. 2015. A Novel Joint Data Hiding and Compression Scheme Based on SVMVQ and Image Inpainting, International Journal of Science Technology & Engineering, Vol.1, Issue. 12, PP. 51-56.
- [21] A. Subash and N.George. 2015. Improved Integrated Data-Hiding and Compression Scheme for Digital Images using Vector Quantization (VQ) and Side Match, International Journal of Innovative Research in

Computer and Communication Engineering,
Vol.3, Issue.11, PP.10589-10597.

- [22] K.Gurusamy and et al. 2015. A LITERATURE SURVEY ON SECURE JOINT DATA HIDING AND COMPRESSION SCHEME TP STORE HIGH CAPACITY DATA IN IMAGE, International Journal of Technical Research and Application, Vol.3, Issue.1, PP.01-04.
- [23] A.Jain and D. Pawar. 2013. Encrypted Reversible Data Hiding on Compressed Image, International Journal of Computer Application, Vol.69, No.25, PP.1-5.
- [24] M. Nazrul Islam, M. Faysal Islam, and K. Shahrabi. 2015. Robust information security system using steganography, orthogonal code and joint transform correlation, Optik-International Journal for Light and Electron Optics, Vol. 126, PP. 4026-4031.
- [25] S.Bader and Noufal.P. 2016. Integrated Data Hiding and Compression Scheme Based on SMVQ and EOF Inpainting, International Conference on Emerging Trends in Engineering, Science and Technology, Procedia Technology ,Vol.24, PP.1008-1015.
- [26] C.Wei Shiu, Y.Chi Chen and W.Hong. 2015. Encrypted image-based reversible data hiding with public key cryptography from difference expansion, Vol.39, PP.226-233.