



# A Novel Joint Data Hiding and Compression Scheme Based on SMVQ and Image In-Painting

**K. Ravi Babu<sup>1st</sup>, G.Keerthi<sup>2</sup>, P.Preethi<sup>3</sup>,  
Y.Sri Harsha<sup>4</sup>, P.Yagnesh<sup>5</sup>, I. Sai Priya<sup>6</sup>**

<sup>1</sup>Assistant Professor ,Dept:ECE ,PBR VITS, Kavali, Andhra Pradesh – 524201

<sup>2,3,4,5,6</sup> UG students Dept. of ECE,PBR VITS, Kavali, Andhra Pradesh -524201 India

*Abstract* : In order to guarantee communication efficiency and save network bandwidth, compression techniques can be implemented on digital content to reduce redundancy, and the quality of the decompressed versions should also be preserved. The two functions of data hiding and image compression can be integrated into one single module, which can avoid the risk of the attack from interceptors and increase the implementation efficiency. On the sender side, the blocks in the leftmost and topmost of the image are compressed by main code book, each of the other residual blocks in raster-scanning order can be embedded with secret data and compressed simultaneously by SMVQ according to the current embedding bit. SMVQ is developing to reduce the block artifact of the decompressed image and increase compression ratio, because the correlation of the neighboring block is consider and the indices of the sub code books are stored. After segmenting the image compressed codes into a series of sections by the indicator bits, the receiver can achieve the extraction of secret bits and image decompression successfully according to the index values in then segmented sections. On the receiver side image edge based harmonic in painting is used for reconstructing lost or deteriorated parts of images. The proposed scheme shows the performances for compression ratio and Better Peak Signal to noise Ratio (PSNR).

*Index Terms* – VQ, SMVQ, IMAGE IN-PAINTING, DATA HIDING, IMAGE COMPRESSION

## I. INTRODUCTION

The rapid development of Internet technology that in development, people can transmit the data and share digital (images, video) content with each other conveniently and it is rapidly used. In order to guarantee communication (that is internet) efficiency and save the network bandwidth, efficiency, compression techniques can be implemented in digital components to reduce redundancy, noise and the quality of the decompressed should also be preserved. Most digital content, digital images and videos, are converted into the compressed form of transmission. Another important issue in a network environment is how to transmit the secret or private data securely through the internet. In the traditional cryptographic methods, the encryption process is used to convert the plaintext into cipher text using the encryption algorithm; the meaningless random data of the cipher text may also arouse the suspicion from the attacker. On the other side decryption process are used to convert the Cipher text into plain text. Cipher text implies meaningless random data. Even though cryptographic methods are providing better security, there may be a chance of finding a plain text by the attacker. To solve this problem in steganography technique is developed in both academia, industry and more. The goal of cryptography is to make text/information unreadable by a third party or attacker, whereas the goal of steganography is to hide the data from a third party or attacker. Due to the rapid use of digital images on the Internet, how to compress the images and hide the secret data into the compressed form of images efficiently deserves in depth study. There many data hiding schemes for compressed codes that it is applied to hide the data , i. e stenography etc. which apply to various different compression techniques of images, that may be JPEG200, JPEG, vector quantization (VQ). But digital images are most popular because of their usage on the internet. Different application uses different Steganography techniques on their requirements Lossy data compression techniques that create smaller image by discarding excess image pixel from the original image. VQ is used due to its simplicity and cost effectiveness for digital image compression. The Euclidean distance is taken to evaluate the similarity between the code words in the codebook and image block for the VQ compression process. The block is represented, that is recorded which is having the index of the code word with smallest distance. The index values containing

in the table for all blocks are generated as code of VQ compression, . And only index values are stored instead of pixel values. And through lookup table for each received index, that is a VQ decompression process. The side match vector quantization (SMVQ) is an improved version of VQ. Both the sub code books and codebook are used to generate index value. In Lin et al.'s [1] method, a key stream (KS) is generated for a original image H according to its content and its watermark W. Fig. 2 shows the flow chart of their key stream generating procedure. In Chang et al.'s reversible embedding method [2], they used the concept of SMVQ to create three state codebooks  $G_0$ ,  $G_1$ , and  $G_2$  for each processed block X. [3] shows the relationship of three  $4 \times 4$  blocks, where block U is the upper adjacent block of X and block L is the left adjacent block of X. Both U and L have been encoded. Then, the corresponding position of  $G_1$  is set to null. After  $G_0$  and  $G_1$  are constructed, they can be used to perform the embedding process. If X is equal to the  $i$ th code word of  $G_0$  and the  $i$ th code word of  $G_1$  is available, one secret bit is embedded in X. If the secret bit is 0, X is unchanged; otherwise, X is replaced with the  $i$ th code word of  $G_1$ . If the  $i$ th code word of  $G_1$  is null, X cannot embed any secret bit.

The hit map [4] is created by training each block X as follows: If X is equal to the  $i$ th code word of its sorted codebook; the  $i$ th bit of the hit map is set to 1. Otherwise, the  $i$ th bit of the hit map remains as 0. The sorted codebook consists of the code words sorted by Eq. (2) in the super codebook. [5] The Techniques have been evolved into two categories. Those applicable to original or uncompressed images fall in the first category (Category I), and those applicable to compressed images are considered to be the second one (Category II). In Category I, outstanding reversible watermarking algorithms have been advanced including integer transform, lossless compression, difference expansion histogram modification, prediction expansion, and accurate sorting and prediction methods. [6] For color images with moderately complex scenes, all DCT based modes of operation typically produce the following levels of picture quality for the indicated ranges of compression. These levels are only a guideline quality and compression can vary significantly according to source image characteristics and scene content. [7] Also, the space for storing seven codebooks is  $7 \times 256 \times 16 \times 8 = 229376$  bits, where a codebook contains 256 code words and a code word contains  $4 \times 4$  pixels of 8-bit gray values. For group length  $t = 2$  and a watermark value  $W_k$ , if each secret  $S_{i,j}$  ( $j=0, 1, \dots, 6$ ) created from block  $H_i$  is not equal to  $W_k$ , block  $H_i$  will not connect to any watermark bit. [8] Fig. 13 shows an example to describe the embedding process, where 32 state codebooks are used. Indexes located in the first row or the first column of the index table remain unchanged and are used as seek blocks. [9] Blocks located in the left-up part of the cover image use their down neighboring blocks and their right neighboring blocks to create their own state codebooks  $S_0, S_1, \dots, S_{m-1}$  [10] Since side match vector quantization (SMVQ) provides better image quality of reconstructed image and compression bit rate than vector quantization (VQ) does, it becomes another choice to compress the transmitting images when the bandwidth is limited.

## II. EXISTING METHOD

### *Vector Quantization:*

Vector quantization (VQ) is a classical quantization technique from signal processing which allows the modeling of probability density functions by the distribution of prototype vectors. It was originally used for data compression. It works by dividing a large set of points (vectors) into having approximately the same number of points closest to them. Each group is represented by its centroid point, as in k-means and some other clustering algorithms. The density matching property of vector quantization is powerful, especially for identifying the density of large and high-dimensional data. Since data points are represented by the index of their closest centroid, commonly occurring data have low error, and rare data high error. This is why VQ is suitable for lossy data compression. It can also be used for lossy data correction and estimation. Vector quantization is based on the competitive learning paradigm, so it is closely related to the self-organizing map mode.

VQ [6, 7] is a well-known lossy compression method especially designed for digital images due to its simple yet efficient encoding and decoding procedures. Fig. 1 shows the VQ encoding and decoding processes.

Before encoding a grayscale image, the image is first partitioned into nonoverlapping blocks of  $r \times l$  pixels, so each block can be represented by an  $r \times l$ -dimensional vector. The basic function of VQ is to map each block using a mapping function Q from  $r \times l$ -dimensional Euclidean space  $R^{r \times l}$  to a finite subset  $\Psi$  of  $R^{r \times l}$ ; that is,

$Q: R^{r \times l} \rightarrow \Psi$ , where  $\Psi = \{Y_1, Y_2, \dots, Y_n\}$  is called the codebook and  $Y_i$  is the  $i$ -th codeword in  $\Psi$ .

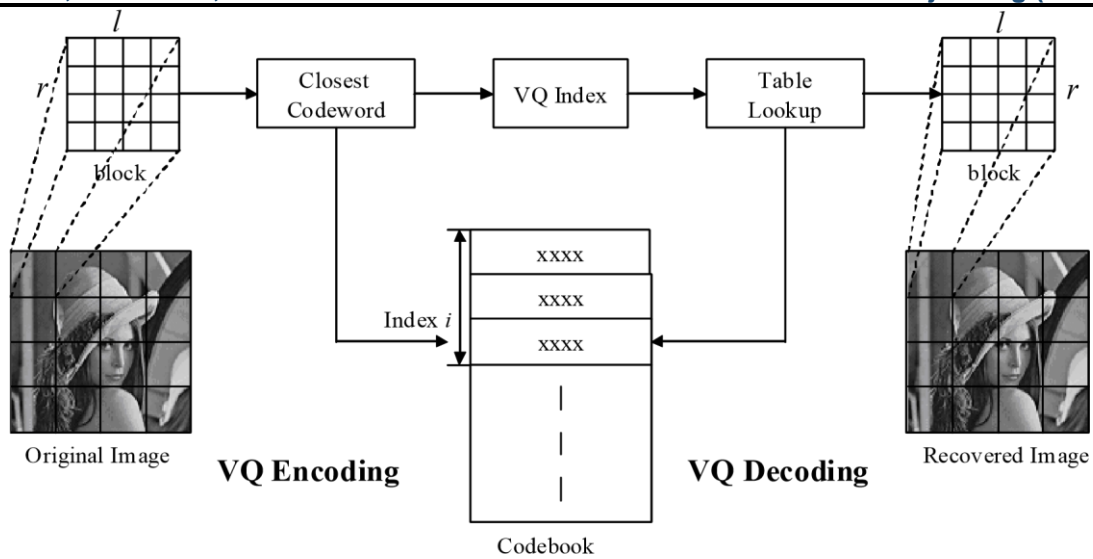


Fig. 1. VQ encoding and decoding processes

During encoding, the closest codeword in the codebook is found for each vector  $X \in R^{r \times l}$  of the original image. The distance between  $X$  and a codeword  $Y_i, i = 1, 2, \dots, n$ , is determined by the Euclidean distance,  $d(X, Y_i)$ :

$$d(X, Y_i) = \|X - Y_i\| = \sum_{j=0}^{r \times l - 1} (x_j - y_{i,j}) \tag{1}$$

where  $x_j$  and  $y_{i,j}$  are the  $j$ -th elements of vectors  $X$  and  $Y_i$ , respectively. When the closest codeword  $Y_i$  of  $X$  is found, index  $i$  is used to encode vector  $X$ , with the original image eventually represented by the indices of these closest codewords.

In the decoding phase, only table lookup operations are required to reconstruct the original image. The VQ decoder requires a codebook that is the same as the VQ encoder. According to the indices compressed by the VQ encoder, the VQ decoder fetches the corresponding codewords to reconstruct the image. Therefore, the VQ-compressed image quality is significantly influenced by the quality of the codebook, which can be well-designed by [3,8].

**III. PROPOSED METHOD**

Since side match vector quantization (SMVQ) provides better image quality of reconstructed image and compression bit rate than vector quantization (VQ) does, it becomes another choice to compress the transmitting images when the bandwidth is limited. To expand the cover media for transmitting confidential information, we propose a novel data hiding scheme which embeds secret data into the SMVQ-compressed image. In terms of the payload capacity, the visual quality, and the compression rate, experimental results confirm that the performances of our scheme are better than that of other information hiding schemes for VQ-based and SMVQ-based compressed images. In addition, the embedded secret data can be extracted from the stego-image without referencing the original cover image.

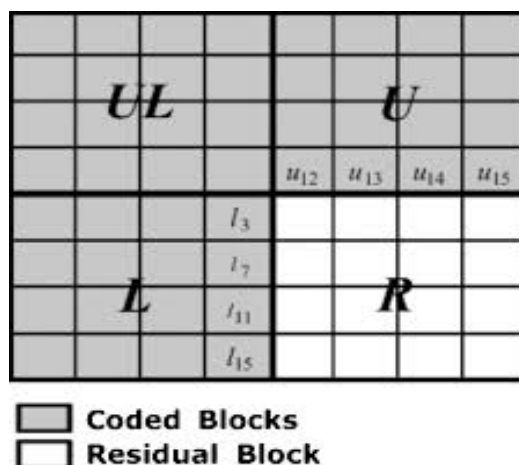


Figure 3.2: Code block and residual block

Side match vector quantization provide

1. Change original seed blocks. More seed blocks or “cross” blocks. Then residual blocks can be recovered by comparing with sides of its upper, right, down, and left blocks.
2. We recover every residual blocks in advance. Compare with corresponding blocks in original image. When its distortion is over predetermined threshold, then encode the block in usual way.
3. The SMVQ algorithm tries to make the gray level transition across the boundaries of the vectors as soon as possible. In our experiments, the improvement using SMVQ to recover the lost blocks is up to 3.618 dB for the image Lena. An interleaved SMVQ (ISMVQ) algorithm is also proposed in this paper. The ISMVQ algorithm combines the VQ algorithm and the SMVQ algorithm. In our experiments, the improvement of ISMVQ over VQ is up to 1.488 dB at the same bit-rate for the image Boat.

3.1 FLOW CHART OF COMPRESSION

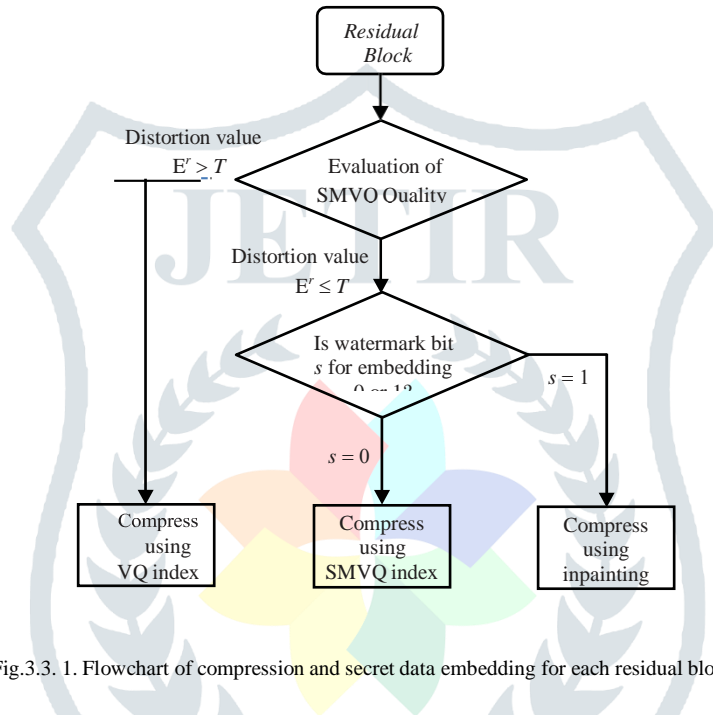


Fig.3.3. 1. Flowchart of compression and secret data embedding for each residual block.

3.2 FLOW CHART OF DE COMPRESSION

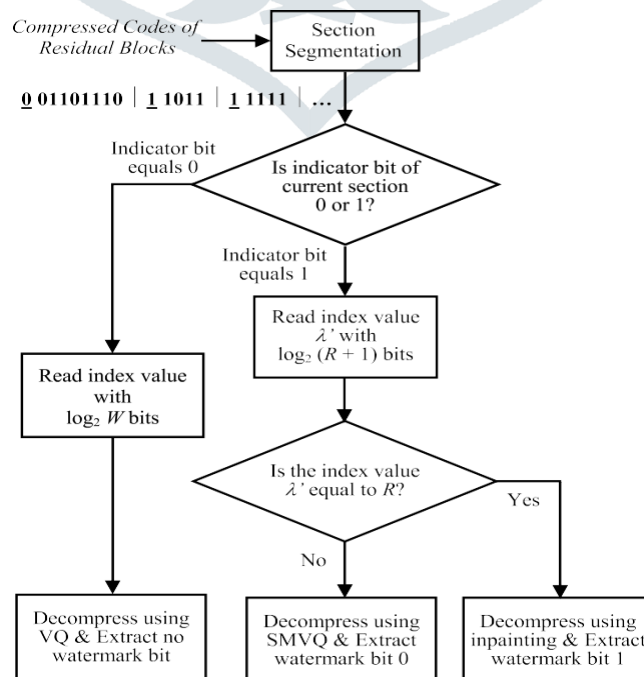


Fig. 3. Flowchart of decompression and secret data extraction for each residual block.

## IMAGE IN PAINTING

In-painting is nothing but reconstruct the lost and remove small regions or small defects .On other important definition is image in-painting is called image interpolation and video or audio interpolation. In basic image in-painting is referring to the sophisticated algorithm. The sophisticated algorithm is used to replace the lost or corrupted part of data is reconstructed.

Applications of image in-painting cinema and photography is used for film reconstruction for reverses distortion e.g. cracks in photographs or dust spots in film and scratched.In our scheme image in painting is ancient technique and it is repair valuable art work an un detectable manner. Digital use in painting repairing damaged part of photography or films and removing chosen areas and wiping of visible watermarks.

The image in-painting is very ancient technique; it is manually repairing valuable artwork in an undetectable manner. Image in painting is used to repairing the damaged photographs and digital images. The image in-painting method is usually depends upon the partial differential equation (PDE) method. Partial differential equation based methods interpolation based methods and patch based method. In this several mathematical physics models that can be used for PDE based method.

In our scheme, a PDE-based image in-painting method using the fluid dynamics model is adopted [34]. Denote  $\mathbf{B}_\chi$  as the region including the current block  $\mathbf{B}_{x,y}$  that needs compression by in-painting and the available neighboring region of  $\mathbf{B}_{x,y}$ . Let  $B_\chi(\zeta, \eta)$  be the gray value of  $\mathbf{B}_\chi$  in the coordinate  $(\zeta, \eta)$ . The Laplacian  $B_\chi(\zeta, \eta)$  is used as a smoothness measure of the region  $\mathbf{B}_\chi$ . By analogizing the in-painting process as the fluid flowing and imitating the practice of a traditional art professional in the manual retouching, details in the unknown region may be created through propagating the available information in the surrounding areas into the unknown region along isopoda directions. The field of isophote is defined as:

$$\nabla^\perp B_x(\xi, \eta) = \left( -\frac{\partial}{\partial \xi} i + \frac{\partial}{\partial \eta} j \right) B_x(\xi, \eta),$$

Having finished the in painting process,  $\nabla^\perp B_x(\xi, \eta)$  should be normal to the gradient of the smoothness  $B_x(\xi, \eta)$ .

## IV. PROPOSED SYSTEM ALGORITHM

The algorithm of data hiding and image compression based on SMVQ include the following steps

Step1: Load input image size MN, M=Width of the image, N=Height of the image.

Step2: Evaluation of SMVQ Quality.

Step3: Compress using SMVQ and Image in painting Based on Index values.

Step4: Data hiding by coding process, a=1 sequential encoding a=2 random encoding

Step5: Data extraction according to the index values with coding process.

Step6: Decompress using SMVQ and In painting

### 4.1 PERFORMANCE EVALUATION

#### 4.1.1 Peak Signal to Noise Ratio

The measure of compression using PSNR is more preferable then the other measures.(Such as SNR measure). The measure makes use of constant value for comparing the noise rather using fluctuating signal as used in SNR. Thus it permits the received PSNR values to be dealt more significantly when compared to the various image-coding algorithms. Thus PSNR is employed through of the image-coding circles in order to compare image-coding algorithm.

$$\text{PSNR} = 10 \log_{10} \left[ \frac{(2^b - 1)^2}{M \times N \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (f(m,n) - g(m,n))^2} \right]$$

Where  $M \times N$  = Size of the image,  $g(m, n)$  = Reconstructed image,  
 $f(m, n)$  = Original image

#### 4.1.2 Compression ratio ( $C_R$ )

The compression ratio is the basic metric used to calculated performance of the compression algorithm. The type of compression metrics used for digital data changes according to the data changes compression used.

The expression for the calculated compression ratio in percentage is

( $C_R$ ): Output file size (bytes)



#### 4.1.3 ADVANTAGES OF PROPOSED METHOD

- Compression ratio.
- Better peak signal to noise ratio
- Reduce redundancy, and
- The quality of the decompressed versions should also be preserved.

#### 4.1.4 APPLICATIONS OF PROPOSED SYSTEM

##### 4.1.4.1 Save network bandwidth

The rapid development of Internet technology, people can transmit and share digital content with each other conveniently. In order to guarantee communication efficiency and save network bandwidth, compression techniques can be implemented on digital content to reduce redundancy, and the quality of the decompressed versions should also be preserved. Nowadays, most digital content, especially digital images and videos are converted into the compressed forms for transmission.

#### RESULTS AND DISCUSSION

V. Experiments were conducted on a group of images to verify the effectiveness of the proposed scheme. In the experiment, the sizes of the divided non-overlapping image blocks were  $4 \times 4$ , i.e.,  $m=4, n=4$ . Accordingly, the length of each code word in the used SMVQ codebook was 16. The performances of the compression ratio and better peak signal to noise ratio for the proposed scheme were evaluated. Because the threshold  $T$  used in the procedure of the image compression and secret data embedding is closely related to the compression method for each residual block and also influences on the performance of the proposed scheme, testing for different image sizes conducted in the compression and secret embedding procedure. As shown the following experimental results.

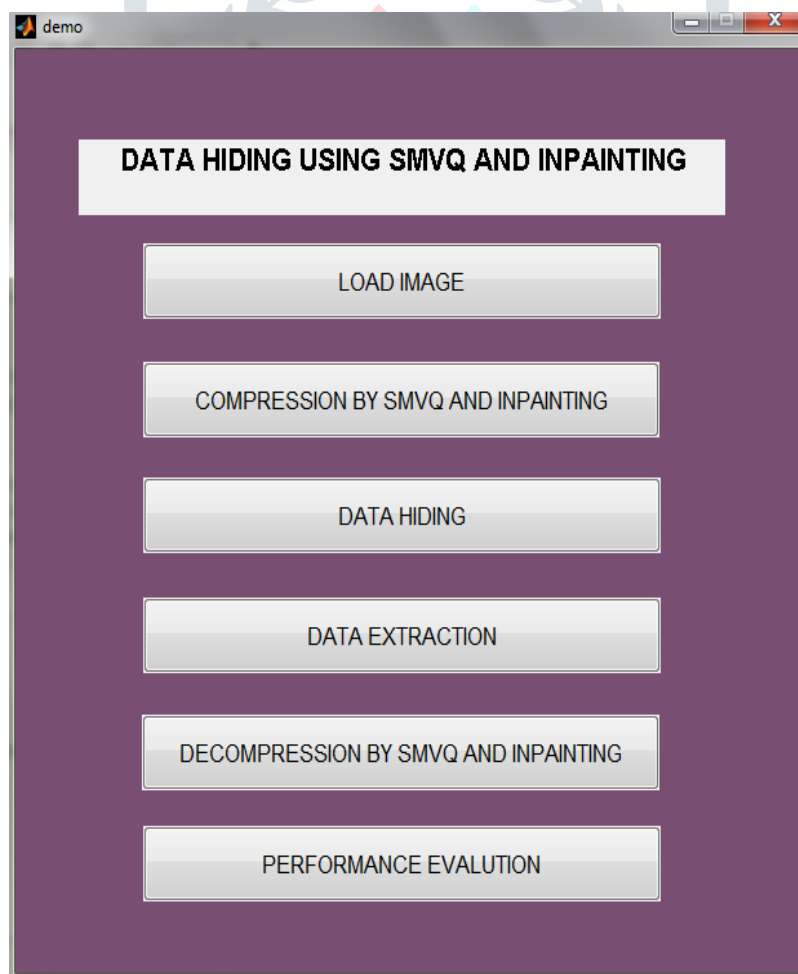


Figure5.1: Graphical user interface diagram



Figure (a): Boat Input image

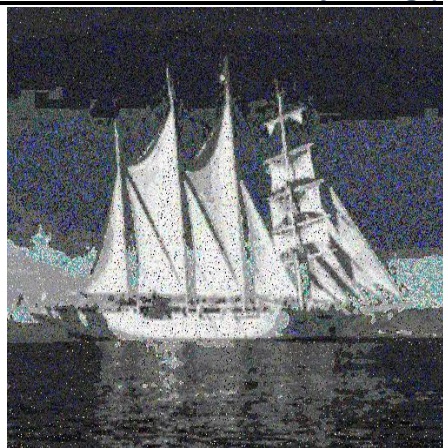
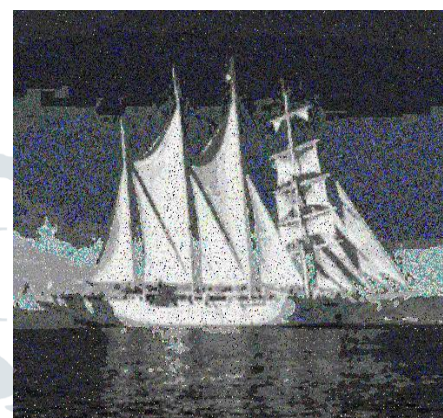


Figure (b): Compressed by SMVQ



Figure(c): secret image



Figure(d): compressed and encrypted image



Figure(e): Decrypted message



Figure (f): Decompressed image

TABLE: Performance evaluation of Boat image with different sizes

	W=256		W=512		W=1024	
	PSNR	CR	PSNR	CR	PSNR	CR
VQ(Existing Method)	30.06	16.00	30.12	14.12	30.11	12.80
SMVQ(Proposed method)	35.03	18	35.04	72	35.03	288

## I. CONCLUSION

In this work, we propose a joint data hiding and compress scheme by the method of SMVQ and image in painting. And the block with left part and top most part of image are used for the secret bits embed and compression at single time, the compression switches between the SMVQ and image in painting method that is according to bits. And also VQ is use for the complex block that control visul distortion and error diffusion. At receiver side the segmented compress code into series of section by its, and then the secret bits are extracted to index value in segment section, decompression is achieved by the method of SMVQ, VQ and image in painting. Further more that is propose scheme has 2 function of data hiding and compression into one module.

## REFERENCES

- [1]. National Institute of Standards & Technology, —Announcing the Advanced Encryption Standard (AES),|| Federal Information Processing Standards Publication, vol. 197, no. 1, 2001.
- [2]. R. L. Rivest, A. Shamir and L. Adleman, —A Method for Obtaining Digital Signatures and Public-Key Cryptosystems,|| Communications of the ACM, vol. 21, no. 2, pp. 120-126, 1978.
- [3]. P. C. Su and C. C. Kuo, —Steganography in JPEG2000 Compressed Images,|| IEEE Transactions on Consumer Electronics, vol. 49, no. 4, pp. 824-832, 2003.
- [4]. H. W. Tseng and C. C. Chang, —High Capacity Data Hiding in JPEG- Compressed Images,|| Informatica, vol. 15, no. 1, pp. 127-142,2004.
- [5]. S.Y. C. Hu, —High-Capacity Image Hiding Scheme Based on Vector Quantization,|| Pattern Recognition, vol. 39, no. 9, pp. 1715-1724, 2006
- [6]. C. C. Lee, W. H. Ku and S. Y. Huang, —A New Steganographic Scheme Based on Vector Quantization and Search- Order Coding,|| IET Image Processing, vol. 3, no. 4, pp. 243-248, 2009.
- [7]. C. C. Chen and C. C. Chang, —High Capacity SMVQ Based Hiding Scheme Using Adaptive Index,|| Signal Processing, vol. 90, no. 7, pp.2141-2149, 2010. www.ijird.com June, 2014 Vol 3 Issue 6
- [8]. L. S. Chen and J. C. Lin, —Steganography Scheme Based on Side Match Vector Quantization,|| Optical Engineering, vol. 49, no. 3, pp.0370081-0370087, 2010.
- [9]. W. J.Wang, C. T. Huang and S. J. Wang, —VQ Applications in Steganographic Data Hiding Upon Multimedia Images,|| IEEE Systems Journal, vol. 5, no. 4, pp. 528-537, 2011
- [10] H. W. Tseng and C. C. Chang, “High Capacity Data Hiding inJPEG-Compressed Images,” Informatica, vol. 15, no. 1.