ISSN: 2321-7782 (Online)
Volume 3, Issue 5, May 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

An Improved Method for Joint Data Hiding and Compression of Multiband Images

Swathy Sebastian[1]
Electronics and communication
College of engineering Cherthala
Alappuzha, India

Abdulkareem V[2]
Electronics and communication
College of Engineering Cherthala
Alappuzha, India

Abstract: A novel and improved method for joint reversible data-hiding and compression scheme for digital multiband images is proposed in this work. The image compression techniques like Vector Quantization and Side Match Vector Quantization along with the image inpainting are used for data hiding and compression. In the proposed method, from the multiband image the user can select the bands to hide the secret data according to the size and nature of the data to be embedded. On the sender side, the bands are sub divided into equal size blocks and the number of blocks determines the hiding capacity.

Except for the blocks in the leftmost and topmost of the band, each of the other residual blocks will be embedded with secret data and compressed simultaneously by SMVQ or image inpainting, adaptively according to the current secret bit. VQ is also used for some complex blocks and the blocks in the leftmost and topmost of the band to control the visual distortion. The other bands which are not used for data hiding are also compressed with VQ. 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 the segmented sections. Only the index of the corresponding codeword is transmitted to the receiver, hence there is a huge reduction in the transmission rate. The proposed method is, a highly secure method for data hiding as the hidden data is retrieved with an infinite PSNR.

Keywords: Data-hiding, Vector quantization, Side match vector quantization, Image inpainting

I. INTRODUCTION

Nowadays, most digital content, especially digital images and videos, are converted into the compressed forms for transmission. Another important issue in an open network environment is how to transmit secret or private data securely.

Even though traditional cryptographic methods can encrypt the plaintext into the cipher text, the meaningless random data of the cipher text may also arouse the suspicion from the attacker.

Data hiding in an image involves embedding a large amount of secret information into a cover image with minimal perceptible degradation of image quality. However, the hiding capacity for secret data and the distortion of the cover image are a tradeoff since more hidden data always results in more degradation on the visual quality of the cover image. Moreover, when data hiding is implemented on the compressed domain of image, the hiding capacity and the visual quality of cover images can be further restricted.

With the definition of reversible data hiding, secret message can be hidden by schemes offered by algorithm designers. On the other hand, at the data extraction stage, both the original content and the hidden message should be perfectly extracted, and hence, how to design such schemes has become an interesting task.

Many well-accepted image compression algorithms have been proposed to counter this problem, such as VQ, SMVQ, JPEG, JPEG2000, and so on. One of the most commonly studied image compression techniques is VQ, which is an attractive choice because of its simplicity and cost effective implementation. Indeed, a variety of VQ techniques have been successfully applied in real applications such as speech and image coding. VQ not only has faster encode/decode time and a simpler framework than JPEG/JPEG2000 but it also requires limited information during decoding, and those advantages cost VQ a little low compression ratio and visual quality. VQ works best in applications in which the decoder has only limited information and a fast execution time is required.

Among the many image compression techniques that have been proposed for reversible data hiding techniques, Vector Quantization is one of the popular. Some other methods have also suggested in this field. Among them SMVQ is a recently developed one. It is modified version of VQ. Researches have also been done in this field using methods like histogram shifting, integer wavelet transform, difference expansion method etc. Some of them are explained here.

However, in all of the above mentioned schemes, data hiding is always conducted after image compression, which means the image compression process and the data hiding process are two independent modules on the server or sender side. Under this circumstance, the attacker may have the opportunity to intercept the compressed image without the watermark information embedded, and the two independent modules may cause a lower efficiency in applications. Also the implementation efficiency will be low. Therefore it will be better to design a method in which data hiding and image compression are done simultaneously and especially in a multi-band. So that the secret data will be more secure and the implementation efficiency will be higher.

II. Proposed Scheme

The proposed scheme is an improved method for joint reversible data-hiding and compression scheme for digital multiband images. The image compression techniques like Vector Quantization and Side Match Vector Quantization along with the image inpainting are used for data hiding and compression. In the proposed method, the multiband image is used to hide the secret data. On the sender side, the image is separated into individual bands. From the multiband image the user can select the bands to hide the secret data according to the size and nature of the data to be embedded. The bands are then compressed according to the ascending order of their band number.

The bands used for hide secret data are sub divided into equal size blocks. The number of blocks and the correlation between neighboring blocks determines the hiding capacity. In the proposed method, rather than two separate modules, only a single module is used to realize the two functions, i.e., image compression and secret data embedding, simultaneously. The compression of bands is based mainly on the Side Match Vector Quantization (SMVQ) mechanism which is a modified form of VQ. Except for the blocks in the leftmost and topmost of the band, each of the other residual blocks will be embedded with secret data and compressed simultaneously by SMVQ or image inpainting adaptively according to the current secret bit. VQ is also used for some complex blocks and the blocks in the leftmost and topmost of the band to control the visual distortion and error diffusion caused by the progressive compression.

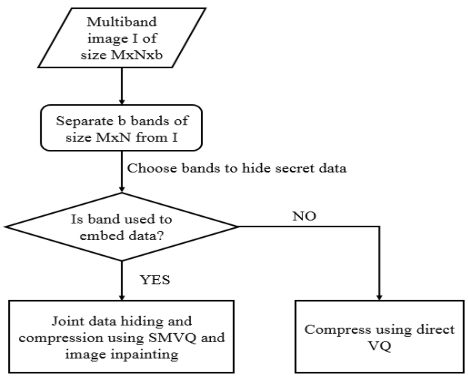

Fig. 2.1: Flowchart of the proposed method

The other bands which are not used for data hiding are also compressed with Vector Quantization. After segmenting the image compressed codes into a series of sections by the indicator bits, it is transmitted to the receiver. At the receiver end after receiving the secret embedded and compressed codes of the image, one can extract the embedded secret bits successfully during the image decompression. As an extension of VQ, SMVQ was developed to alleviate the block artifact of the decompressed image. As the correlation of neighboring blocks is considered and the indices of the subcodebooks are stored it increases the compression ratio considerably. In this work, the standard algorithm of SMVQ is modified to further achieve better decompression quality and to make it suitable for embedding secret bits. Also partial differential equation based image inpainting is used at the receiver for obtaining high quality decompressed image. Only the index of the corresponding codeword is transmitted to the receiver hence there is a huge reduction in the transmission rate. The proposed method is, a highly secure method for data hiding as the hidden data is retrieved with an infinite PSNR.

2.1 Separation of bands

In the proposed method, the multiband image is used to hide the secret data. Let us denote the original uncompressed image sized M x N x b as I where M, N and b represents the number of rows, columns and bands respectively. The user have the freedom to select the number and order of bands used to transmit secret data according to the size and secret nature of the embedding data. After deciding the bands to be used to hide data the image I is separated into b number of single band images $I_b = I(:,:,a)$, where a=1,2,..,b. Bands used to hide secret data are compressed with SMVQ and image inpainting adaptively according to the secret bit to be embedded. And the bands which are not used to hide secret data are compressed directly by Vector Quantization. In the proposed scheme the sender and the receiver both have the same codebook Ψ with W codewords and each codeword length is n^2.

2.2 Compression of Bands using Vector Quantization

The bands which are not used to hide secret data are compressed directly by using Vector Quantization. The individual band I_b is divided into non overlapping n x n blocks. VQ initially involves constructing a codebook. The elements in the codebook are called codewords. Generally, the LBG algorithm is employed to yield the desired codebook. With the generated codebook, each block in the band is encoded with the index of the nearest codeword, such that the total storage space for an image is minimized.

After generating the codebook the similarity between each block and the codebook is measured and each block is mapped to a codeword which has least structural dissimilarity. To find the nearest codeword, the squared Euclidean distance between the current block and each codeword in the codebook is measured. The squared Euclidean distance between the block X = ($x_1\ x_2$ . . .x_n) and codeword C_j= (c_{j1}, c_{j2}. . .c_{jn}) where $C_j$the j^{th} codeword of the codebook, is given as,

D(X,C_j) = $\sum_i (x_i - c_{ji})^2$

The index j of the nearest codeword is saved for the current block. The process continues in the raster scanning order and all the indices are segmented into an array [index_a] where *a* varies from 1 to b.

2.3 Compression of Bands and Secret Data Embedding

The bands used to hide secret data are first divided into the non-overlapping n x n blocks. Denote all k divided blocks in raster scanning order as $B_{i,j}$, where k = $M * N/_{n^2}$, i = 1, 2 . . . M/n and j = 1, 2 . . . N/n. Before being embedded, the secret bits are scrambled by a secret key to ensure security. The blocks in the leftmost and topmost of the band I_b, i.e., $B_{i,1}$ for (i = 1, 2. . . M/n) and $B_{j,1}$ for (j = 2, 3. . . N/n), are encoded by VQ directly and are not used to embed secret bits. The residual blocks are encoded progressively in raster scanning order, and their encoded methods are related to the secret bits for embedding and the correlation between their neighboring blocks. The block diagram of the processing for each residual block is illustrated in Fig 3.2.

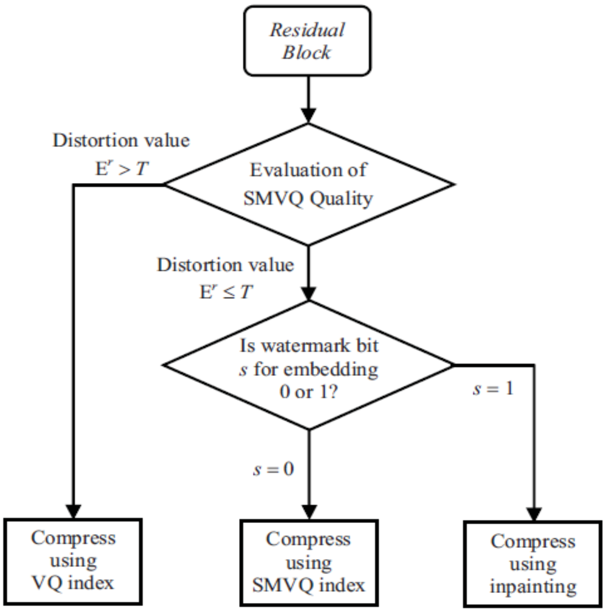

Fig 2.2 Flowchart of compression and secret data embedding for each residual block

Denote the current processing block as $B_{x,y}$ for (2 ≤ x ≤ M/n) (2 ≤ y ≤ N/n) and it's left and up blocks are $B_{x,y-1}$ and $B_{x-1,y}$ respectively. As shown in the fig 4.2. $c_{p,1}$ for (1 ≤ p ≤ n) and $c_{1,q}$ for (2 ≤ q ≤ n) represents the 2n - 1 pixels in the left and upper borders of $B_{x,y}$. The n pixels in the right border of $B_{x,y-1}$ and the n pixels in the bottom border of $B_{x-1,y}$ are denoted as $l_{p,n}$ and $u_{n,q}$ respectively. This 2n-1 pixels in the left and upper borders of $B_{x,y}$ are predicted by the neighboring pixels i.e,

$$c_{11}=\frac{l_{1,n}+u_{n,1}}{2}$$

$$c_{p,1}=l_{p,n}$$

$$c_{1,q}=u_{n,q}$$

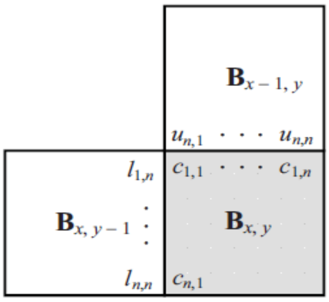

Fig 2.3 Prediction based on neighboring pixels

Instead of all n^2 pixels in $B_{x,y}$ only 2n-1 predicted pixels are used to search the codebook Ψ. After transforming all W codewords in the codebook Ψ into the n x n matrices, the mean square error (MSE) E_w is calculated between the 2n-1 predicted pixels in $B_{x,y}$ with the corresponding values of each transformed codeword C_w sized n x n.

$$E_w = \Sigma(c_{p1} - c_{p1}^{w})^2 + \Sigma(c_{1q} - c_{1q}^{w})^2$$

The R codewords with the smallest MSEs, are selected to generate one subcodebook Θ_{xy} for the block B_{xy} (R < W). Assume the codeword indexed λ has the smallest MSE, E_r with all n^2 pixels in B_{xy} (0 ≤ λ≤ R-1). If E_r > T, it implies that the current residual block B_{xy} locates in a relatively complex region and it has lower correlation with its neighboring blocks. In order to achieve better decompression quality, the standard, block independent VQ with codebook Ψ is used to compress the blockB_{xy}, and no secret bits are embedded.

Otherwise, if $E^r \leq$ T, it implies that the current residual block B_{xy} locates in a relatively smooth region and it has higher correlation with its neighboring blocks. So SMVQ or image inpainting is adaptively utilized to compress the block B_{xy} according to the secret bit s for embedding. If s=0, SMVQ is selected else if s=1, Image inpainting is used.

Also note that, if VQ is adopted, an indicator bit, 0, should be added before the compressed code of the VQ index for B_{xy}. If not, the indicator bit 1, is added as the prefix of the compressed code for B_{xy} to indicate that secret bit is embedded.

As for the block B_{xy}, if its $E^r \leq$ T and the current secret bit s for embedding is 0, SMVQ is utilized to conduct compression, which means that the index value λ occupying $log_2 R$ bits is used to represent the block B_{xy} in the compressed code. Because the codeword number R in subcodebook Θ_{xy} is less than the codeword number W of the original codebook Ψ the length of the compressed code for B_{xy} using SMVQ must be shorter than using VQ. On the other hand, if $E^r \leq$ T and the current secret bit s for embedding is 1, the image inpainting technique is used.

There are mainly three classes of the image inpainting methods, i.e., partial differential equation (PDE) based methods, interpolation-based methods, and patch-based methods. The PDE-based inpainting methods often propagate the available information of gray values automatically from surrounding areas into region Ω to be inpainted along a specific direction. There are several mathematical physics models that can be used for PDE-based inpainting, such as the fluid dynamics model and the heat transfer model. Different PDE models correspond to the different methods of information propagation. Image inpainting can recover the image structural information effectively when the processed region is not too large. Evidently, if $E^r \leq$ T it implies that B_{xy} locates in a relatively smooth region. Thus, it is suitable to conduct image inpainting in the compression for B_{xy} under this condition.

In this scheme a PDE based image inpainting method using the fluid dynamics model is adopted. Denote B_ as the region including the current block B_{xy} that needs compression by inpainting and the available neighboring region of B_{xy}. Let $B_{\chi}(\varepsilon,\eta)$ be the gray value of B_{χ} in the coordinate (ε,). The Laplacian $\Delta B_{\chi}(\varepsilon,\eta)$ is used as a smoothness measure of the region B_{χ}. By analogizing the inpainting 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 isophote directions.

The field of isophote is defined as:

$$\nabla^{\perp} B_{\chi}(\varepsilon,\eta)=(-\frac{\partial i}{\partial \varepsilon}+\frac{\partial j}{\partial \eta})\, B_{\chi}(\varepsilon,\eta)$$

where i and j are unit directional vectors. Clearly, variations in image gray values are minimal along the isophote directions. Having finished the inpainting process, $\nabla^{\perp} B_{\chi}(\varepsilon,\eta)$ should be normal to the gradient of the smoothness $\Delta B_{\chi}(\varepsilon,\eta)$:

$$\nabla[\Delta B_{\chi}(\varepsilon,\eta)\ ].\ \nabla^{\perp} B_{\chi}(\varepsilon,\eta)=0$$

The scalar product in the above equation indicates projection of the smoothness change onto the direction of isophote. If we let the projection value be equal to the change of image gray values with respect to time t, the following PDE can be acquired

$$\frac{\partial}{\partial t} B_{\chi}(\varepsilon,\eta)=\nabla[\Delta B_{\chi}(\varepsilon,\eta)\ ].\ \nabla^{\perp} B_{\chi}(\varepsilon,\eta)$$

Information propagation of this inpainting model finishes until the gray values in B_{xy} reach stable state. Hence, the structural and geometric information of the block B_{xy} can be recovered effectively without serious blurring on edges. Consequently, when s = 1, in order to indicate that block B_{xy} is processed by inpainting and differentiate from the index λ produced by SMVQ, the index value R occupying log_2(R + 1) bits is used as the compressed code of B_{xy} (R > λ). For simplicity, assume that log_2(R + 1) is an integer and log_2R ≤ log_2(R + 1).

After the current block B_{xy} is processed, the following block in raster-scanning order repeats the above procedure. Each processed block should be substituted with its corresponding decompressed result, i.e., VQ codeword, SMVQ codeword, or inpainting result, for the success of progressive mechanism. The whole procedure of band compression and secret data embedding finishes until all residual blocks are processed. Then, the compressed codes of all blocks in the band are concatenated and transmitted to the receiver side.

2.4 Decompression of Multiband Image and Extraction of Hidden Data

At receiver end, after receiving the compressed codes, the receiver conducts the decompression process to obtain the decoded image that is visually similar to the original uncompressed multiband image, and the embedded secret bits can be extracted during the decompression. The decoder receives the index code cell, index that was transmitted and segments each index array as, $index_i$ where i varies from 1, 2. . . .b.

Then the classifier check the size of each index array. The index array corresponding to bands which contain hidden data that is the bands coded with SMVQ and image inpainting has the size equal to the number of blocks multiplied with the number of digits contained in its binary equivalent. And the bands which are not embedded with secret data contains an extra indicator bit. So the size of the index array corresponding to bands coded with VQ is greater than the other index arrays. According to the decision, the bands are decompressed from their corresponding index arrays. And all the individual bands are then combined according to their band number to obtain multiband cover image. On the other side, from the decoded secret bits the hidden data is retrieved.

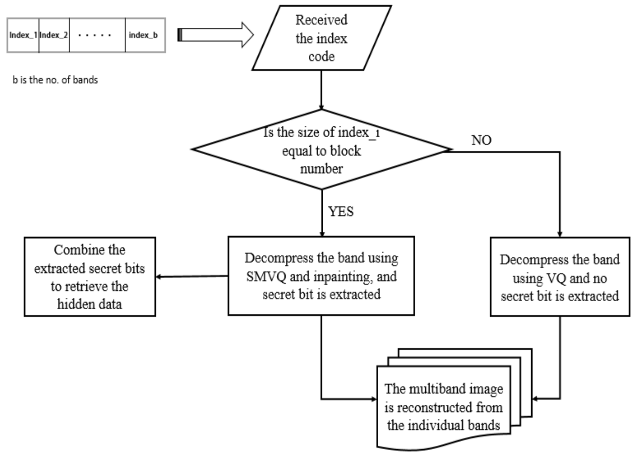

Fig 2.4 Flowchart of decompression of multiband image and secret data extraction

2.4.1 Decompression and Secret Data Extraction from Individual Bands

After identifying the index array corresponding to the bands embedded with secret data, the secret bits are extracted from the received codes by the corresponding decompression methods. The leftmost and topmost blocks of the band need to be used in the decompression for other residual blocks, they should be first decompressed by their VQ indices retrieved from the compressed codes. Each VQ index of these pre-decompressed blocks occupies log2W bits. And the residual blocks are processed block by block in raster-scanning order.

To conduct the decompression and secret bit extraction of each residual block, the compressed codes are segmented into a series of sections adaptively according to the indicator bits. Explicitly, if the current indicator bit in the compressed codes is 0, this indicator bit and the following log_2W bits are segmented as a section, which means this section corresponds to a VQ compressed block with no embedded secret bit. The decimal value of the last log_2W bits in this section is exactly the VQ index that can be used directly to recover the block.

Otherwise, if the current indicator bit is 1, this indicator bit and the following log_2(R+1) bits are then segmented as a section, which means this section corresponds to an SMVQ or inpainting compressed block. Denote the decimal value of the last log_2(R+1) bits in this section as λ'. Under this circumstance, if λ' is equal to R, it implies that the residual block corresponding to this section was compressed by inpainting and that the embedded secret bit in this block is 1. Otherwise, if $\lambda' \, \varepsilon \, [0, R-1]$, it implies that the block corresponding to this section was compressed by SMVQ and that the embedded secret bit is 0.

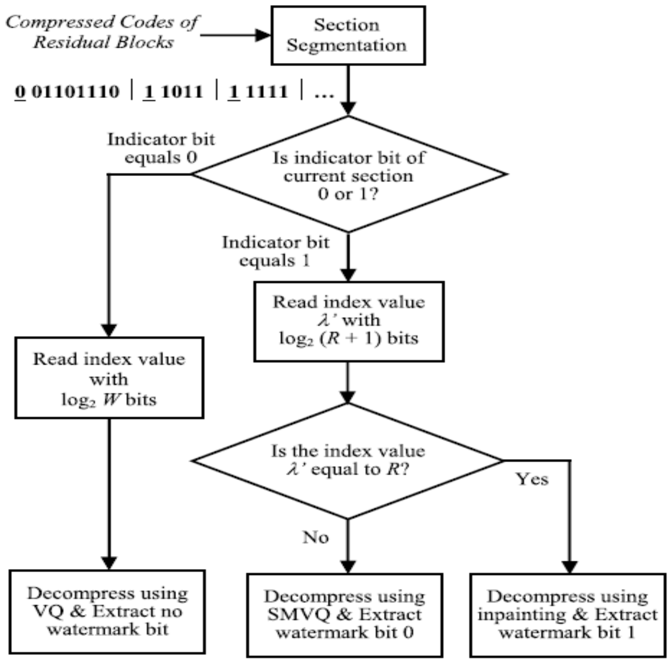

Fig 2.5 Flowchart of decompression and secret bit extraction for each residual block.

If the current segmented section corresponds to an inpainting compressed block the available information of its neighboring decompressed blocks are utilized to conduct recovery by the same inpainting technique used in the compression process. If the current segmented section corresponds to an SMVQ compressed block SMVQ index value, i.e. λ', is used to recover this block with the assistance of its left and upper decompressed blocks. Using the same prediction method used in the compression side, the 2n-1 pixels in the left and upper borders are estimated by the neighboring pixels in its left and upper decompressed blocks. Similarly, the MSEs are calculated between these 2n-1 predicted pixels in the block with the corresponding values of all W codewords in the codebook Ψ. Then, the R codewords in Ψ with the smallest MSEs are chosen to generate a subcodebook. Finally, the codeword indexed λ', in the generated subcodebook is used to recover the block.

After all the segmented sections in the compressed codes complete the above described procedure, the embedded secret bits can be extracted correctly, and the decompressed band I_b can be obtained successfully. Due to the decoding of the compressed codes, the decompressed band I_b doesn't contain the embedded secret bits any longer. Note that the process of secret bit extraction can also be conducted independently, which means that the receiver can obtain all embedded bits by simply segmenting and analyzing the compressed codes without the decoding.

The index array of VQ compressed bands have an extra indicator bit to indicate that the compressed code do not contain secret bit. So the size of those index arrays are always greater than the number of blocks get compressed. If the received index array has size greater than the number of block number, then those indices are used to decompress using VQ decompression method and no secret data are retrieved. From the received index array select each index and find out the corresponding codeword from the codebook. This codeword will be the nearest codeword for the corresponding block. So the block is constructed using the codeword. Both the number of elements in the block and the length of the codeword will be the same. In the same way the entire band is reconstructed in the raster scanning order. Since the decompression is done with generated codeword, the reconstructed band will be different from the original band as the VQ method is a lossy image compression method. So we couldn't get an infinite PSNR.

After decompressing all the bands, the multiband image can be obtained by combining all the individual bands in the ascending order of their band number. On the other side, secret bits decoded from different bands can be combined to form the transmitted hidden data. Since the secret data was not transmitted but the bands of the cover image were coded adaptively according to the binary equivalent of the secret data, so on the receiver side the user can extract the exactly same secret data without any distortion. Hence can retrieve the hidden data with an infinite PSNR. This is one of the main advantage of the proposed method

III. Conclusion

The proposed work is an improved and secure method for joint reversible data-hiding and compression scheme for digital multiband images. From the multiband image the user can select the bands to hide the secret data according to the size and nature of the data to be embedded.

The experimental results show that the proposed scheme has a very good performances for hiding capacity, compression ratio, and decompression quality. Also these factors can be further increased by the proper selection of the host image. Furthermore, the proposed scheme can integrate the two functions of data hiding and image compression into a single module seamlessly. The major advantage of this method is that the transmitter send only the index and the codebook to the receiver. So it saves a lot of memory and also there is a huge reduction in the transmission rate. By using multiband image more data can be embedded in a single cover image. The proposed method is, a highly secure method for data hiding as the hidden data is retrieved with an infinite PSNR. And also provides a satisfactory PSNR in the range of 32 to 35 for the host image. So it can be concluded that the proposed method is a very good option for data hiding and image compression. Due to its high security and efficiency the proposed method has wide applications in different areas like military, law enforcement, medical field etc...

References

1. Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantization design," *IEEE Trans. Commun.*, vol. 28, no. 1, pp. 84–95, Jan. 1980.
2. K. Sayood, "Introduction to Data Compression", Morgan Kaufmann, 1996, pp. 87-94.J
3. T. Kim, "Side match and overlap match vector quantizers for images," IEEE Trans. Image Process., vol. 1, no. 2, pp. 170-185, Apr. 1992.
4. M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting", in Proc. 27th Int. Conf. Comput. Graph. Interact. Tech., New Orleans, LA, USA, Jul. 2000, pp. 417-424.
5. W. C. Du andW. J. Hsu, "Adaptive data hiding based on VQ compressed images," IEE Proc. Vis., Image Signal Process., vol. 150, no. 4,pp. 233-238, Aug. 2003.
6. C. C. Chang and W. C. Wu, "Hiding secret data adaptively in vector quantisation index tables," IEE Proc. Vis., Image Signal Process., vol. 153, no. 5, pp. 589-597, Oct. 2006.
7. S. C. Shie and S. D. Lin, "Data hiding based on compressed VQ indices of images," Comput. Standards Inter., vol. 31, no. 6, pp. 1143-1149, 2009.
8. C.C. Chang, T.S. Nguyen, C.C. Lin, "A reversible data hiding scheme for VQ indices using locally adaptive coding", J. Vis. Commun. Image Representation 22 664-672, 2011.
9. C.-C. Chang and W.-C. Wu, "A steganographic method for hiding secret data using side match vector quantization," IEICE Trans. Inf. Syst., vol. E88-D, no. 9, pp. 2159-2167, 2005.
10. C. C. Chang, W. L. Tai, and C. C. Lin, "A reversible data hiding scheme based on side match vector quantization," IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 10, pp. 1301-1308, Oct. 2006.
11. C. C. Chen and C. C. Chang, "High capacity SMVQ-based hiding scheme using adaptive index", Signal Process., vol. 90, no. 7, pp. 2141-2149, Apr 2010.
12. Jiann-Der Lee, Member, IEEE, Yaw-Hwang Chiou, and Jing-Ming Guo, Senior Member, IEEE "Reversible Data Hiding Based on Histogram Modification of SMVQ Indices" IEEE Trans. Information Forensics and security, vol. 5, no. 4, pp. 600-612, December 2010.