Image Compression Using Wavelet Transform and Vector-Quantization

Group_14
Anshul Mahajan[1], 300093480 || Rishavjeet Singh Sidhu[2], 300139600
[1,2]Master of Engineering in Electrical and Computer Engineering, University of Ottawa, Ottawa, Ontario, Canada

ABSTRACT

Image compression using various techniques is always evolving concept. Modifying various techniques to best fit different needs is what derives innovation in the field. One such modification is combining two lossy algorithms, Discrete Wavelet Transform and Vector-Quantization to process medicinal images. The question is to see how this algorithm performs universally for image compression. The effect of both Discrete Wavelet Transform (DWT) and Vector-Quantization (VQ) are discussed and how they merge into one algorithm is shown. Different preprocessing techniques can be added as per the applications of the algorithm.

Keywords: Discrete Wavelet Transform, Filter banks, Image Compression, Subbands, Vector Quantization.

1. INTRODUCTION

There's no doubt in the fact that with the advancement in technology more data is coming into play. This data either image, videos, text or audio needs to be stored somewhere and using a big chunk of space on it is not ideal. It is very important to use compression techniques on data such as image and videos in order to do further processing on them. In today's world, where image quality is increasing day by day it is very important to have efficient compression techniques which can make image transmission, uploading and downloading possible using less bits and less time.

The methodology of image compression is to encode the image with few bits by reducing the redundancy [2]. There are a lot of image compression techniques present, mainly divided in two types: lossy and lossless compression. There is a tradeoff between exact reconstruction and obtaining high compression ratio. Lossy techniques provide us with high compression ratio while sacrificing good reconstruction, on the other hand, lossless compression is useful for exact replica of the original image with less compression ratio. Some of the lossy compression techniques are DCT (Discrete Cosine Transform), DWT (Discreet Wavelet Transform) and Vector Quantization whereas RLE (Run Length Encoding), LZW (Lempel Ziv Welch), Arithmetic coding and Huffman encoding account for lossless compression techniques [6][10].

Some researchers in this field have achieved good compression using a hybrid algorithm combining Discrete Wavelet Transform and Vector Quantization. This paper follows the same approach for compression method. As human eye is less precise about the high-resolution details in an image, this can be used as basis of Discrete Wavelet Transform and Vector Quantization and thus wavelet coefficients containing this information can be encoded with low-order precision [2].

2. RELATED STUDIES

Research in the field of DWT and VQ is very interesting and booming right now. Using both these lossy compression techniques to get good compression ratio and high PSNR is what driving researchers in exploring and advancing this field. Some of the works in this field are:

1. Tejas S. Patel, Ravindra Modi, Keyur J. Patel [7], in 2013, proposed a method in which DWT is first applied to image to obtain the subbands. These subbands are then preprocessed before going through the VQ step in order

to obtain the compressed image. They were able to obtain good reconstruction of images at the cost of compression ratio compared to when regular DWT and VQ technique is used. Also, it had a big margin of error and preprocessing data needs to be sent to the decoder, which is also a drawback [7].

2. Aree and Jamal [8], in 2011, proposed a method for compression of medical images. Their algorithm includes converting image from RGB to YCbCr first and then applying Forward Discrete Wavelet Transform (FDWT) to get subbands. The LH and HL subbands were divided into non-overlapping 8x8 blocks. They applied Forward Discrete Cosine Transform and then it was quantized. At last Forward DPCM and entropy coding was done. They performed this technique of medical images like MRI. However, their algorithm produced blocking artifacts that can be seen in the recovered image. This not very good when it comes to medical image compression [8].

3. DISCRETE WAVELET TRANSFORM

3.1 Wavelet Theory

Discrete Wavelet Transform find many operations in the area of science and mathematics. In image processing DWT is mostly used for denoising the images. This is a wavelet transform technique that involves discretely sampled wavelets that give location and frequency information of the image. Wavelets are obtained by decomposing a 2D image by a wavelet function, such as Haar wavelets.

We start with a mother wavelet. Mother wavelet must have a zero mean and a finite energy. This function undergoes scaling and translations to produce other functions. These functions generated with certain properties can be used as basis functions for different interests in signals, in our case images. The scaling and translations change the size and shape of the functions. Translations that don't change the shape but only size are generally called as dilations. We can obtain family of functions using mother wavelet and that function is given by equation 1 [9], in which 'a' represents scaling parameter and 'b' represents shifting parameter [9].

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\,\psi\left(\frac{t-b}{a}\right) \qquad (1)[9]$$

3.2 Wavelet Decomposition and Subbands

Discrete Wavelet Transform in image processing is an application of filter banks. Filter banks are 1-D in nature and the objects we want to analyze i.e. images are 2-D. So, we need to develop some technique to apply this 1-D filter banks first along the rows of the images and then along the columns or vice-versa.

In subband coding a filter bank with cascade of stages is used. Each stage is made up of Low Pass filters and High Pass filters. A widely used filter is Quadrature Mirror Filter (QMF). One example of a famous filter is Johnston filter that can be 8-tap, 16-tap or 32-tap. The difference between them is the value and number of coefficients they have. In general, they all satisfy this universal property that filters are symmetric. If h{n} is the impulse response of a Low Pass filter, then $(-1)h_{N-1-n}$ is the impulse response of High Pass filter [9]. However, filters having fewer taps are less efficient but also less computations are needed, whereas, filters with more taps are higher efficiency and more computations [9].

The process starts with taking an image. This image is processed through various filters, one Low Pass filter and one High Pass filter with specific transfer functions. Frequency response of both, Low Pass and High Pass filter are given in figure 1. Since we are splitting signal into two subbands and bandwidth in each subband is half, we get redundant samples. According to Nyquist theorem, we don't lose any information even if we use half samples from our image. This is achieved by applying a downsampler by 2 operation [9].

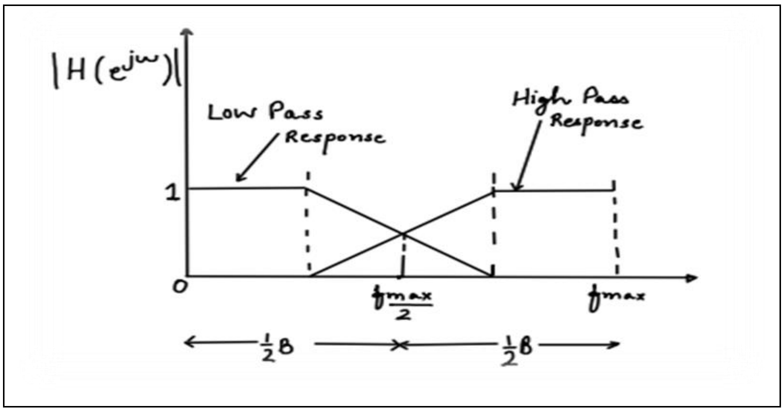

Figure 1. Response of High Pass and Low Pass Filters.

The generation of subbands is a multistep process and can be easily understood by Figure 2. It depicts the filters and where downsampling is done. In the figure ‘g’ depicts the high pass filter and ‘h’ depicts the low pass filter. High Pass filter is responsible for extracting information about the edges of input image and Low Pass filter is responsible for the approximation. The output is four subbands, Low-Low (LL), Low-High (LH), High-Low (HL) and High-High (HH) [9], all passed through a cascade of filters. The LL subband gives approximation of the input image as it passes through two low pass filters. The LH subband gives horizontal features of the image as it is processed through a low-pass filter against rows and high-pass filter against columns. The LH subband gives Vertical features as it passed through high-pass filter against rows and low-pass filter against columns. The HH subband contains the diagonal features as high-pass filter acts along both rows and columns. All this information in all subbands can be more easily understood in Figure 3.

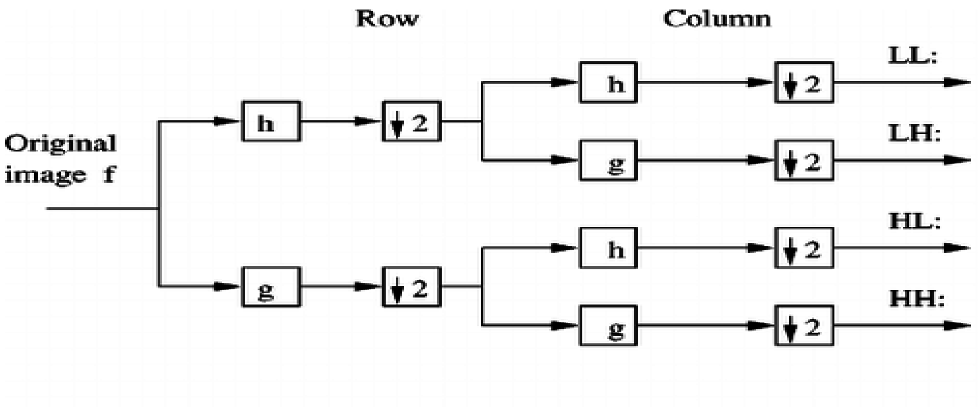

Figure 2. Generation of Subbands.

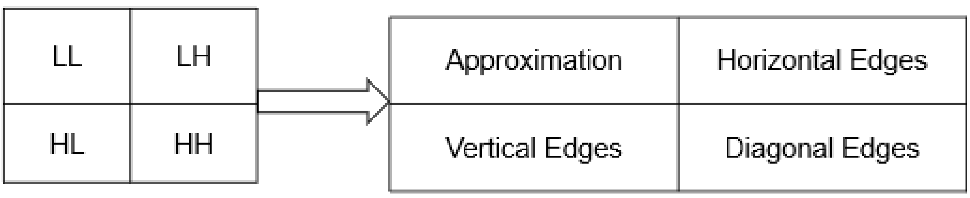

Figure 3. Information in Subbands.

3.3 Subband Coding Algorithm

The algorithm is divided basically into three steps:

1. **Analysis:** Source is processed through a filter bank and further subsampled. The Nyquist theorem justifies the subsampling, in this case downsampling [9].
2. **Coding:** This step aims at selecting compression technique and bit allocation to the subbands. As subbands contain different extent of information so, it makes sense to allocate bits using a measure of amount of information they contain. We can obtain 1,2 or 4 bits per sample data rate by encoding 4,2 or 1 subband respectively. By doing this we can significantly vary the compression ratio thus affecting the reconstructed image for better quality [9].
3. **Synthesis:** This takes place at the receiver side where decoder decodes the image and then output is upsampled by the same factor it was downsampled with. By upsampling we bring back the original rate by adding zeroes between samples. Then upsampled stream is passed through reconstruction filter bank. Adding every output from these filters provide the reconstructed output i.e., the reconstructed image [9].

3.4 Inverse DWT for Reconstruction

The Inverse Discrete wavelet transform reconstructs the image from sub-bands LL, LH, HL, HH. All the sub-bands are first up-sampled by 2 then low-pass and high-pass filter is applied first along the columns and then up-sampled again and then filters applied against rows to obtain the reconstructed image.

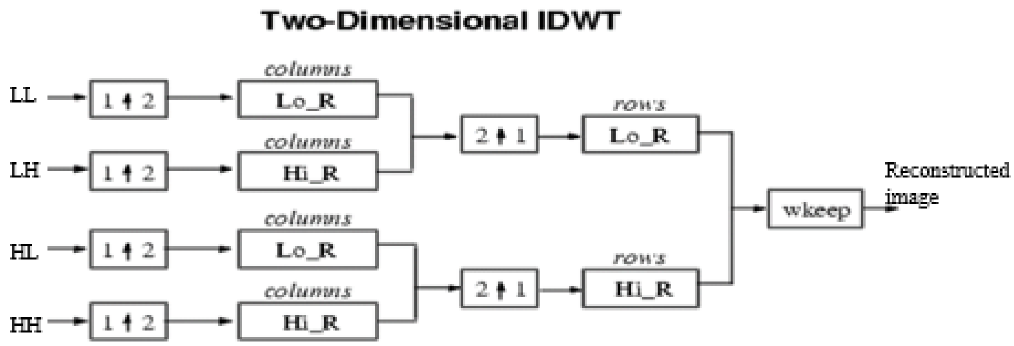

Figure 4. Reconstruction of Image using IDWT.

4. VECTOR QUANTIZATION

Vector Quantization works on the methodology that source outputs can be clubbed together in L-dimensional blocks and these blocks are known as Vectors. The quantization is performed on these vectors and then they undergo further processing to give the final output. One advantage of vector quantization over scalar is that it increases the optimality. The algorithm takes many inputs and maps into one output, so it is a lossy compression technique. It is defined on the methodology of “block quantization” or “pattern matching quantization” [9].

The process starts with grouping source output into vectors or blocks. We take block of L pixels of an image and define it as a vector of L-dimension. This vector is the input to our encoder. The encoder consists of a codebook and a binary index table. The codebook contains L dimensional vectors known as code vectors which are used to represent our clubbed source output vectors. The first vector is compared with the codebook to find the closest code-vector. The index value corresponding to that code-vector is then transferred to the decoder by the binary encoder. If the codebook is of size k, then our binary encoded output will be of $\lceil \log_2 k \rceil$ bits. On the decoder side, there is the same codebook and a lookup table. The lookup table receives the binary and then compares it with the index number in the decoder. The corresponding code-vector to the index value is sent further to obtain the reconstructed image. The whole process is represented by Figure 5 [9].

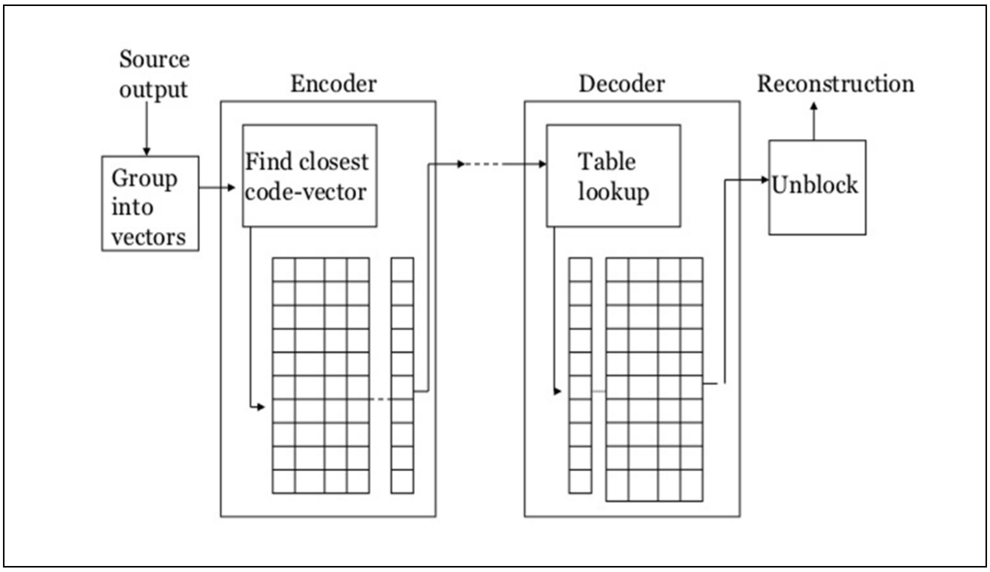

Figure 5. Vector Quantization Process - Encoder and Decoder

5. IMPLEMENTATION AND RESULTS

A hybrid compression algorithm consisting of Discrete Wavelet Transform and Vector Quantization is implemented. The algorithm consists of mainly preprocessing, wavelet decomposition, vector quantization and reconstruction.

5.1 Image Preprocessing

Images may have different types of noises in it. Such as medical images especially ultrasound scans have salt and pepper noises [1]. In such a case, taking care of this challenge first and removing this noise using a median filter from the scans will account for the preprocessing stage. Different types of preprocessing can be used before decomposition into wavelets depending upon the application and images. For implementation in image compression, we need to preserve edges and for further decomposing into wavelets we are converting our color image into grayscale. This accounts as our preprocessing stage.

5.2 Generation of Wavelet Coefficients

Multiple levels of discrete wavelet decomposition can be achieved. In a single level of wavelet decomposition, we obtain four coefficients as we know according to the theory, LL, LH, HL, HH. This is done by simply applying DWT to the filtered image to generate subbands. For second level decomposition, we must apply DWT to either of the four subbands previously obtained. The most fit choice is to decompose LL as it contains the maximum information. So, in a 2-level decomposition we apply DWT to LL subband obtained by 1-level decomposition to obtain seven subbands in total (three from 1-level decomposition, LH, HL, HH and four from 2-level decomposition, LL1, LH2, HL2, HH2). In a 3-level decomposition, again LL subband obtained from 2nd level decomposition undergoes further DWT to provide ten subbands in total (three from 1-level decomposition, LH, HL, HH, three from 2-level decomposition, LH2, HL2, HH2 and four from 3-level decomposition, LL2, LH2, HL2, HH2). The 1-level, 2-level and 3-level decomposition results provided in Figure 6 and Figure 7. A threshold is applied to these coefficients. The coefficients below this threshold are replaced with zero values and coefficients above this threshold are used. The decomposed images can be further reconstructed using Inverse Discrete Wavelet Transform.

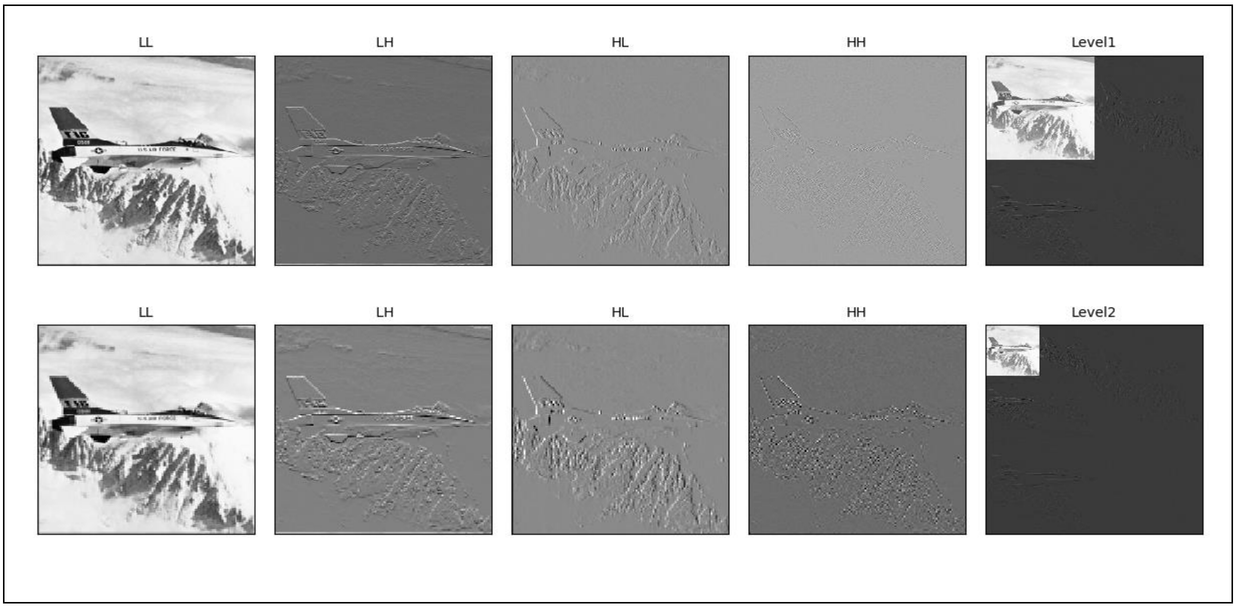

Figure 6. Level-1 and Level-2 decomposition of Test image showing individual subbands and combined subbands.

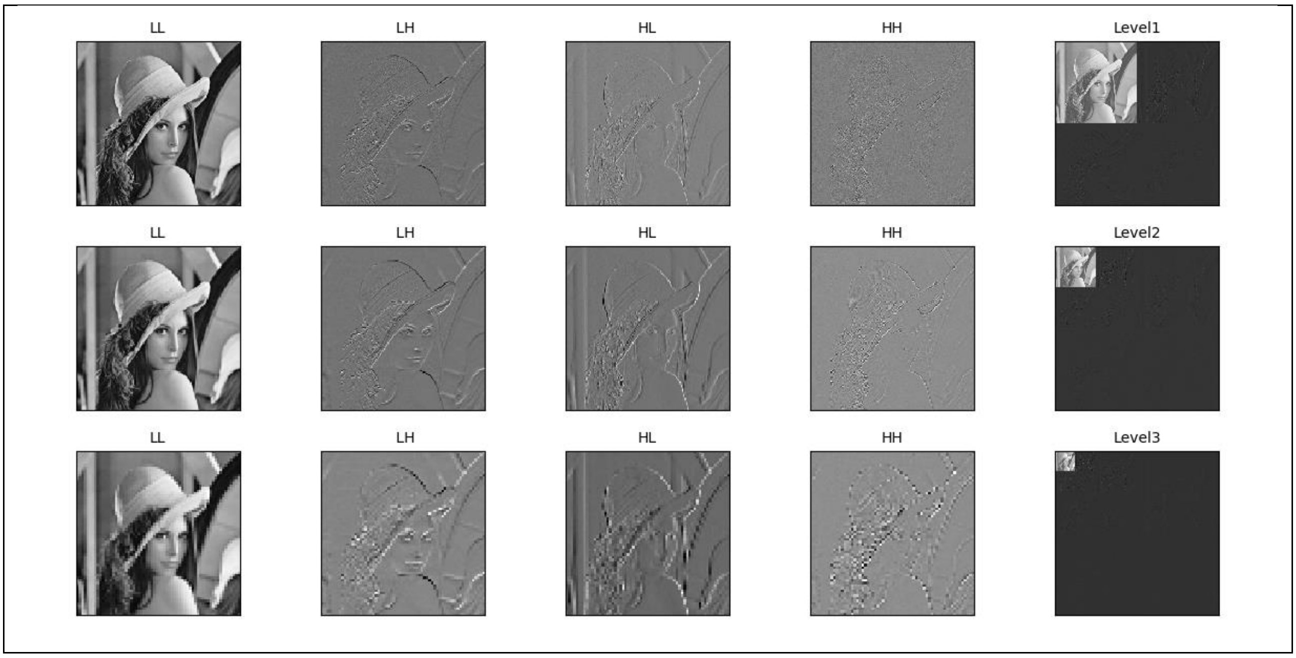

Figure 7. Level-1, Level-2 and Level-3 decomposition of Test image showing individual subbands and combined subbands.

5.3 Vector Quantization

Next step is to vector quantize the coefficients. Vector quantization is applied on every sub-band except the most significant one, and that is, LL subband. It holds the most significant information of the image, so we need to reserve it for further reconstruction. The codebook generation is a complex process and is done using the k-means algorithm. The result of the vector quantization can be encoded using binary encoder. The compressed image can be stored or transmitted. The results for solely Vector Quantization step are shown in Figure 7. Table 2 depicts the compression ratio which is calculated using the original and the compressed image.

Image 1 – Original

Image 1 – Compressed

Image 2 – Original

Image 2 - Compressed

Image 3 – Original

Image 3 - Compressed

Figure 8. Test Images compressed by Vector Quantization

Table 1. Compression Ratio of test images using Vector Quantization

IMAGE	ORIGINAL SIZE	COMPRESSED SIZE	COMPRESSION RATIO
Image 1	68 KB	21 KB	3.23
Image 2	49 KB	16 KB	3.06
Image 3	50 KB	17 KB	2.94

5.4 Performance of Hybrid Algorithm using DWT-VQ

The quality of reconstructed image can be measured by Peak Signal to Noise Ratio (PSNR) [1]. The performance of our hybrid algorithm is shown in Figure 9 and Figure 10, using the original test image and the compressed test image. The PSNR of various test images is shown in Table 2. The optimal PSNR obtained for our compression technique is 45.15 when performed on various test images.

Figure 9. Original (left) and Reconstructed (right) test image using DWT-VQ

Figure 10. Original (left) and Reconstructed (right) test image using DWT-VQ

Table 2. PSNR of various test images using DWT-VQ hybrid algorithm

SR. NO.	TEST IMAGE	PSNR
1.	Image 1	45.15
2.	Image 2	43.78
3.	Image 3	33.8
4.	Image 4	28.8
5.	Image 5	29.76
6.	Image 6	42.54
7.	Image 7	34.35
8.	Image 8	44.15

6. CONCLUSION

A combined technique based on Discrete Wavelet Transform and Vector Quantization is presented. This technique is not optimal but can be manipulated for specific applications for better image compression. Comparing to vector quantization and DWT alone, the hybrid algorithm produces better quality reconstructed image. The PSNR obtained by this hybrid algorithm is around 45, which is quite good and imposes that it can be used for good quality image reconstruction. The various level decomposition can be used to get different results for compression. There is potential in this field for research regarding its applicability for different applications.

REFERENCES

[1] Paul Nii Tackie Ammah, Ebenezer Owusu, "Robust medical image compression based on wavelet transform and vector Quantization", Informatics in Medicine Unlocked 15 (2019) 100183.

[2] Amir Averbuch, Danny Lazar and Moshe Israeli. "Image Compression Using Wavelet Transform and Multiresolution Decomposition", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 5, NO. 1, JANUARY 1996.

[3] Binit Amin, Patel Amrutbhai, "Vector Quantization based Lossy Image Compression using Wavelets – A Review", International Journal of Innovative Research in Science, Engineering and Technology, (An ISO 3297: 2007 Certified Organization), Vol. 3, Issue 3, March 2014.

[4] DAVID A. HUFFMAN, ASSOCIATE, IRE, "A Method for the Construction of Minimum-Redundancy Codes*", PROCEEDINGS OF THE I.R.E., September.

[5] Rishav Chatterjee, "Image Compression and Resizing using Vector Quantization and Other Efficient Algorithms", ISSN 2321 3361 © 2017 IJESC, Research Article, Volume 7 Issue No.6

[6] M. Mozammel Hoque Chowdhury and Amina Khatun, "Image Compression Using Discrete Wavelet Transform", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 1, July 2012, ISSN (Online): 1694-0814.

[7] Tejas S. Patel, Ravindra Modi, Keyur J. Patel, "Image Compression Using DWT and Vector Quantization", International Journal of Innovative Research in Computer and Communication Engineering, *Vol. 1, Issue 3, May 2013.*

[8] Aree Ali Mohammed, Jamal Ali Hussein, “Efficient Hybrid Transform Scheme for Medical Image Compression”, International Journal of Computer Applications (0975 – 8887) Volume 27– No.7, August 2011.
[9] Khalid Sayood, [Introduction to Data Compression], Morgan Kaufmann, 225 Wyman Street, Waltham, MA 02451, USA.
[10] Mr. Shubham Yadav, Mrs. Shikha Singh, “A Review on image Compression Techniques”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 4 Issue 9, September 2015.