

# An Efficient method of image compression by merging IWPT transform coding with index vector Quantization through FNN

Gaurav Kumar Pandey

*Department of ECE  
Oriental Institute of Science  
and Technology, Raisen Road, Bhopal*

Rahul Dubey

*Department of ECE  
Oriental Institute of Science  
and Technology, Raisen Road Bhopal*

S.G. Kerhalkar

*Department of ECE  
Oriental Institute of Science  
and Technology, Raisen Road Bhopal*

**Abstract--** By the use of Neural Network, it was found that the reconstructed image has least image complexity and the image size has been reduced considerably by reducing the number of samples. This causes a remarkable increase in quality of the reconstructed image. A new quantization method is proposed in this paper. This method is useful for enhancement of compression quality when each kind of neural network is used to compress the image. Quantization, involved in image processing is achieved by compressing a range of values to a single quantum value. When the number of discrete symbols in a given stream is reduced, the stream becomes more compressible. For example, reducing the number of colors required to represent a digital image makes it possible to reduce its file size. This causes a remarkable enhancement in quality of the reconstructed image. For testing the proposed method we use IWPT transform coding and by merging it with the proposed quantization method a new compression algorithm is obtained. Then results of compression by the merged method are compared with some other transform methods. Compression time and complexity in the merged method is also better than JPEG and make it suitable for the systems with low processor and hardware implementation. Obtained results show that the proposed compression algorithm increases the compression quality of the images remarkably.

**Index Terms--**Artificial neural network, Feed Forward Neural Network (FNN), image compression, IWPT transform coding, quantization algorithm

## I. INTRODUCTION

The advent of digital cameras has drastically changed our priorities, one of the most common uses has been the storage, manipulation, and transfer of digital images. The files

that comprise these images, however, can be quite large and can quickly take up precious memory space on the computer's hard drive. In Image Compression, the researcher's aim is to

reduce the number of bits required to represent an image by removing the spatial and spectral redundancies. The recent growth of data intensive multimedia based applications have not only sustained the need for more efficient ways to encode the signals and images but also have made compression of

such signals central to storage and communication technology.

In recent days we are facing the increasing use of image in many parts of our life. satellites, cameras, medical equipments etc. all of these equipments use or produce image for different

purposes for using these images we have to save or transmit them, then because of the limitation in disk space and channel bandwidth We always need image compression for decreasing the size of data which must be save or transmit.

Generally there are several methods for image compression, in each case based on the conditions,

constraints and criteria we have to choose one of them. Some of these criteria are compression ratio, compression quality, compression time and ability for simple hardware implementation of image compression.

Artificial neural network has paved the way for overcoming these challenges. Parallel architecture, rapid compression and decompression, simple hardware implementation and ability for simultaneous noise canceling in noisy channels [15] that make it an ideal approach for real time systems which need fast compression, systems which can't use bulky Processors, commercial equipments are made as cheap as possible.

The image compression technique is mainly carried under these compression methods Predictive coding, transform coding and vector quantization, we can use neural network in all of these three categories.

The major problem that arises in image compression by neural network is the quality of the compressed or reconstructed image. In this paper we proposed a novel compression algorithm based on a new quantization approach. Experimental results show a great advantage in reconstructed images quality in comparison with other works in the field of compression with neural network[7].

In section [2] we explained the Integer Wavelet Packet Transform (IWPT) coding method. In section 3 the proposed method is introduced, in section 4 we have some experimental result with different images for comparing the proposed algorithm with previous works and in section 5 we explain the conclusions of this work.

## II. INTEGER WAVELET PACKET TRANSFORM CODING

One of the methods for image compression which place in the category of transform coding is IWPT. Lossless image compression plays a very important role in image processing. However, it develops very slowly and it is difficult to improve the efficiency of compression, since complete information and high fidelity are demanded. The lossless compression ratio of the conventional methods is around 2:1 and 3:1, which doesn't satisfy practical compression needs. So, it is necessary to develop more efficient lossless image compression methods. In the past decades, a lot of image compression methods using wavelet transforms have been proposed by many researchers. The wavelet transform is a popular transform due to its ability to offer useful spatio-frequency localization. However, if the spatio-frequency characteristics of a given signal do not match the spatio-frequency localizations offered by the wavelet, this may result in an inefficient decomposition. For example, for signals having high frequency stationary components, the wavelet transform is a bad fit. Arbitrary sub-band decomposition trees, which have been introduced in and as wavelet packets, conceptually represent an elegant generalization of wavelets.

In this section a review of the wavelet transform and of the filter bank scheme are given, and the wavelet packets transform are introduced. The block scheme of the single level wavelet transform is shown in Fig.1.

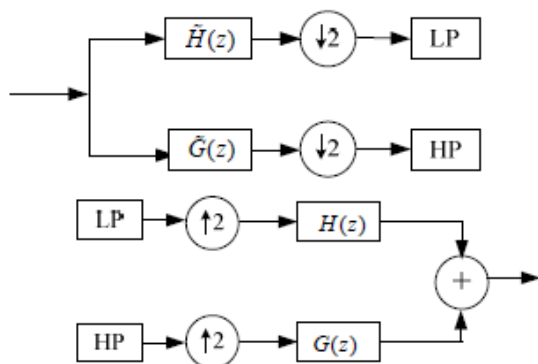


Fig.1 One level wavelet transforms

It has been proven that rounding off the output of each filter right before adding or subtracting yields a couple of perfect reconstruction forward and inverse IWT. It is straightforward to understand that the same procedure which leads to IWT from wavelet transform can be applied to the wavelet packets transform, yielding the IWPT[5]. The implementation follows the same scheme used for the IWT. The IWPT tree can be built iterating the single wavelet decomposition step on both the low-pass and high-pass branches, with rounding off in order to achieve the integer transforms. IWPT yields a representation which can be lossless, as it maps an integer valued sequence onto integer valued coefficients in the transformed domain; moreover, it allows for the selection of an adaptive representation, which

can match the variable characteristics of image better than the IWT.

## III. PROPOSED METHOD

### A. Overall View

In section 2 we came across Integer Wavelet Packet Transform method for transform coding, now we want to introduce the proposed quantization approach for modifying and improving the transform method in brief, in this approach by changing the pixels gray levels in sub images (which used as input for the neural network) statistically decrease the number of samples which must be reconstructed by neural network. This leads to a remarkable increment in the compression quality. first step in this method is splitting the image into equal sub images. In each sub image we have a set of gray levels based on the solidity of the image some of the gray levels can be similar then by aligning the gray levels in an increasing mode. Each set of similar gray levels is replaced with a number, these numbers begin with one for the minor values of gray levels, two for the next values and so on then each number will place at the same position in the 8x8 block as its gray level.

After replacing the gray levels we must make a lookup table to show each gray level's number. For this we need a packet for each sub image, size of this packet depends on the number of gray levels in the sub image. In this packet the gray levels are aligned increasingly in sequential bytes and byte number of each gray level indicates the gray levels number. In this paper we name this packet as gray levels number packet (GNP) this packet must be compressed and save or transmit with the main data.

For compressing the GNP we can use some lossless Methods like Huffman or LZW[12],[13] This algorithm must repeat in other blocks. Now we have 8x8 blocks which their gray levels are replaced with numbers and we name them (number blocks). Then we fed each of these blocks into a two layers neural network this network must be trained previously (with a train database). We can transmit or save output of the hidden layer plus the GNP (gray level's number packet) as the compressed data.

### B. Image Reconstruction

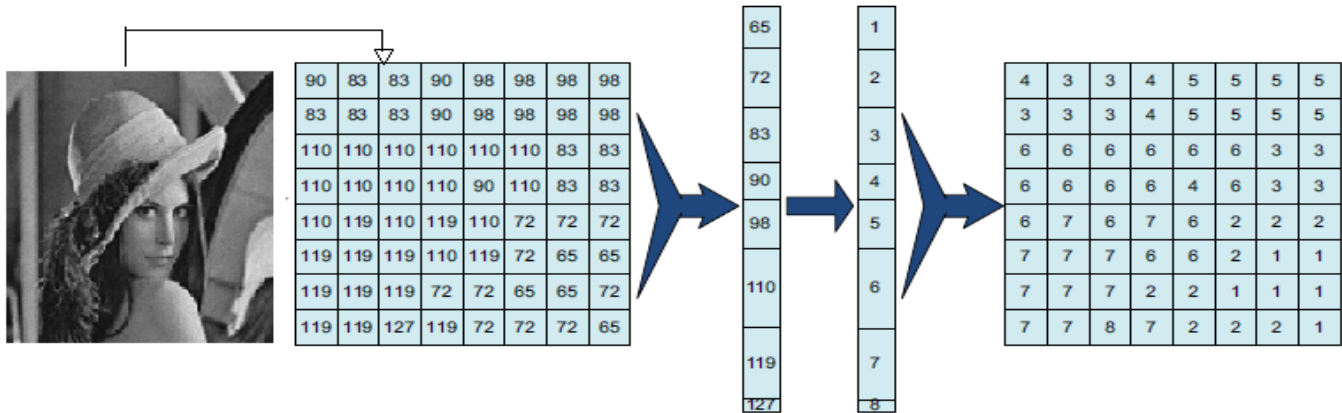
For reconstructing the image we will use output layer of the neural network, outputs of the hidden layer which has been saved as the compressed data will fed into the second part of the network and the outputs in an ideal network must be same as the original number block. Then by using the GNP we can change this number block to a grayscale block which can be exhibited as an image.

### C. Benefits of the proposed method

As we discussed in this method we replace the gray levels of each block with the numbers and this numbers repeat in other blocks for other gray levels then the duty of the neural network in this method is reconstructing these limited numbers. Based on the robustness of the image number of the gray levels in each block is different but in a normal image average of this number is ten .then we can see that in

this method the number of samples that must be reconstruct in each block by the network decrease from 25664 samples to 1064 samples this reduction in the number of samples causes a great improvement in the performance of the neural network by decreasing number of neurons in the hidden layer and increasing the quality of decompressed image. There are several approaches for estimating number of neurons in the hidden layer which cause the best approximation and avoid over-fitting. For proving the fact that with decreasing the

number of samples reconstructed by the network we can decrease hidden neurons we use one of these papers. Many authors proposed a method based on comparing SNRF of the error signal ( $SNRF_E$ ) and SNRF of the white Gaussian noise ( $SNRF_{WGN}$ )



**Fig.2 Replacement trend for converting gray level block at the left to number block at the right**

$$SNRF_E = \frac{E_S}{E_N} = \frac{C(e_i, e_{i-1})}{C(e_i, e_i) - C(e_i, e_{i-1})} \quad (1)$$

$$SNRF_{WGN} = \frac{C(n_i, n_{i-1})}{C(n_i, n_i) - C(n_i, n_{i-1})} \quad (2)$$

Where  $E_s$  and  $E_N$  represent energy of signal and noise,  $C$  is the correlation calculation ( $e_i$ ) is the error in the train sample number  $i$

$$e_i = s_i + n_i \quad (i = 1, 2, \dots, N) \quad (3)$$

Where  $N$  represents number of train samples ( $n$ ) is the noise factor in error and ( $s$ ) is the approximation error For choosing the optimum number of hidden layers neurons we must increase the number of hidden neurons until the  $SNRF_E$  become less than the threshold represents in (4) and then stop.

$$Th_{SNRF-WGN}(N) = \frac{1.7}{\sqrt{N}} \quad (4)$$

From (4) we can conclude that by decreasing the number of train samples ( $N$ )  $Th_{SNRF-WGN}(N)$  increase and it hence decreases the number of hidden layers neurons and vice versa.

The proposed quantization method by decreasing the number of samples which must be reconstructed from 25664 to 1064 decrease the number of train samples ( $N$ ) which are

essential for an optimum training of the neural network and as we demonstrate this cause decrease number of neurons in the hidden layer and also the training time hence more compression plus the better quality in the reconstructed image.

#### IV. EXPERIMENTAL RESULTS

For training the network with proposed method we use Lena image For testing the training image are in the size of  $512 \times 512$  and they are chosen in different solidity and correlation for testing this approach. For testing we use peak signal-to-noise ratio (PSNR). We calculate this criterion on the test images by using three methods. Jpeg2000, Jpeg and Proposed method we also have results of two previous works in the field of compression by neural network on Ref (1) and Ref (2) in the table 1 we compare the result of these three methods in equal compression ratio as we can see the results of the proposed method is fairly higher than other compression methods by neural network and its results are very close to the jpeg. The reconstructed image of Lena with proposed method is shown in Fig. 3 (compression ratio is 6).

In Fig. 4 we compare results of Ref (2), Jpeg and the proposed method in different compression ratios. For calculating the compression ratio (CR) in the proposed method we use (5)

$$CR = \frac{P \times b_p}{\sum_{i=1}^B (H \times b_H + b_{GNP_i})} \quad (5)$$

P is number of pixels in the original image,  $b_p$  is number of bits which we use for saving gray level of one pixel, B is number of blocks in the image, H is number of neurons in the hidden layer,  $b_H$  is number of bits for saving output of each neuron in the hidden layer and  $b_{GNP_i}$  is the number of bits which we use for saving GNP of block i. For simulating the method we use sigmoid as the transfer function of the hidden layer and pure linear as the transfer function of the output layer. For training the network Back Propagation algorithm is used. In table I we compare the compression time of jpeg2000, jpeg and the proposed method in different compression ratios. For simulating and comparing the compression time we performed all respective calculations in MATLAB software. All of the methods are simulated by Pentium Dual-Core 2.5 GHZ processor.

It's observed that the compression time in the proposed method is so better than these two popular methods. It must be

noticed that for simulating Jpeg and Jpeg 2000 compression methods we use ready codes. But in the proposed method the code can be improved for reducing the computation complexity.

## V. CONNCLUSION

A new quantization method for image compression by neural network is proposed. Then by merging this method with IWPT transform coding a new image compression algorithm is obtained and its results are estimated. As we can see in table I, when the compression ratio is 6, average PSNR of the proposed method is 38.2 db better from the classic method. This method is also 7.35 db better from Ref (2) and its results are 1.22 db under Jpeg results.



Fig. 3 (a) Original image (b) Reconstructed image with the proposed method(Compression Ratio=6)

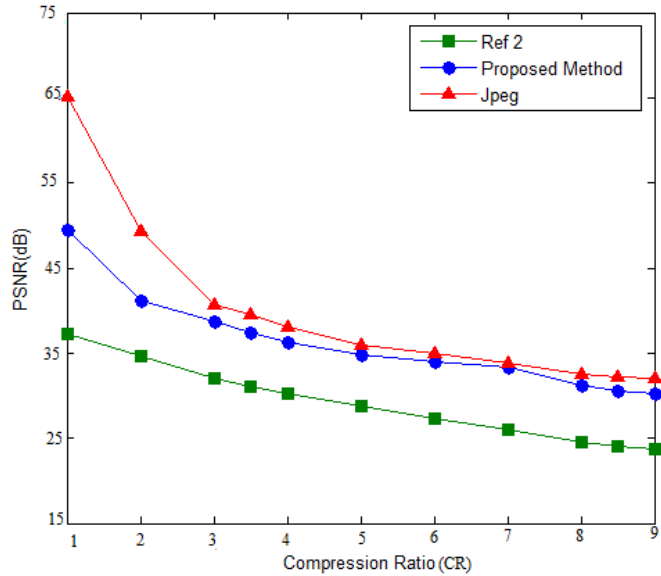


Fig.4 Compression ratio Vs PSNR in Ref[2], Jpeg and Proposed method

In table I we compare the compression time of the proposed method with the jpeg 2000 and jpeg as we can see the average compression time of the proposed method is 67 ms faster than jpeg and 805 ms faster than jpeg 2000

TABLE 1 COMPARISSON OF THE COMPRESSION TIME IN JPEG2000, JPEG AND PROPOSED METHOD

Compression Ratio (CR)	JPEG2000	JPEG	Proposed Method
6	903ms	118ms	49ms
8	840ms	105ms	37ms
10	796ms	102ms	35ms

From the results we can conclude that using this method in comparison with jpeg can remarkably reduce the computation complexity and it also can give us close compression quality to jpeg method then this method can be perfect when we need rapid and real time compression or when we can't use a high speed processor or when we have some constraints in using processor's time for compression. It also seems that implementing this method in FPGA or integrated circuit is simpler than jpeg.

## REFERENCES

- [1] Liying Ma and K. Khorasani, "Application of Adaptive Constructive Neural Networks to Image Compression" *IEEE Transactions on Neural Networks*, vol. 13, no. 5, september 2002
- [2] S. Anna Durai, and E. Anna Saro, "Image Compression with Back- Propagation Neural Network using Cumulative Distribution Function" *International Journal of Applied Science, Engineering and Technology*,3;4 2007
- [3] Robert D. Dony, Simon Haykin, "Neural Networ Approaches to Image Compression" Proceedings of the IEEE, vol. 83, no. 2, february 1995.
- [4] W. Cottrell and P. Munro, "Principal components analysis of images via back propagation," in SPIE Vol. 1001 *Visual Communications and Image Processing* , pp. 1070–1077, 1988.
- [5] A. R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers", Technical report, Department of Mathematics, Princeton University, 1996.,
- [6] W. Cottrell, P. Munro, and D. Zipser, "Learning internal representations from gray-scale images: An example of extensional programming," in *Ninth Annual Conf. of the Cognitive Society*, July 16-18 1987, pp. 462– 473.
- [7] D. Anthony, E. Hines, D. Taylor, and J. Barham, "A study of data compression using neural networks and principal component analysis," in *Colloquium on Biomedical Applications of Digital Signal Processing*, 1989, pp. 1–5.
- [8] E. Oja, "A simplified neuron model as a principal component analyzer," *J. Math. Biology*, vol. 15, pp. 267– 273, 1982.
- [9] I. Diamantaras, "Principal component learning networks and applications," *Ph.D. dissertation, Princeton University*, October 1992.
- [10] A. Namphol, M. Arozullah, and S. Chin, "Higher-order data compression with neural networks," in Proc. *Int. Joint Conf. on Neural Networks '91*, 1991, pp. I 55–59.
- [11] D. DeMers and G. W. Cottrell, "Non-linear dimensionality reduction," in *Advances in Neural Information Processing Systems 5*, 1993.
- [12] David A. Huffman, "A Method for the Construction of Minimum-Redundancy Codes" *Proceedings of the I.R.E.* 1952
- [13] Jacob Ziv, Abraham Lempel "Compression of Individual Sequences via Variable-Rate Coding" *IEEE Transactions on Information Theory*, Vol. IT 24, No.5, September 1978
- [14] Yinyin Liu, Janusz A. Starzyk, Zhen Zhu, "optimizing number of hidden neurons in neural networks" *IASTED international conference on Artificial intelligence and applications* PP: 121 – 126, 2007
- [15] Khaled Masmoudi, Marc Antonini, Pierre Kornprobst, Laurent Perrinet, "A novel bio-inspired static image compression scheme for noisy data transmission over low-bandwidth channels" *ICASSP 2010*, 2010