AN IMPROVED NEURAL IMAGE COMPRESSION APPROACH WITH CLUSTER BASED PRE-PROCESSING

Ashanta Ranjan Routray[1] and Flt.Lt.Dr. Munesh Chandra Adhikary[2]

Fakir Mohan University, Balasore-756019, Odisha, India

[1] ashan2r@yahoo.co.in, [2] mcadhikary@gmail.com

ABSTRACT

The convergence time for training back propagation neural network for image compression is slow as compared to other traditional image compression techniques. This article proposes a pre-processing technique i.e. Pre-processed Back propagation neural image compression (PBN) with an enhancement in performance measures like better convergence time with respect to decoded picture quality and compression ratios as compared to simple back-propagation based image compression and other image coding techniques for color images.

KEYWORDS

Back propagation, Bipolar sigmoid, Vector quantization

1. INTRODUCTION

Image compression standards address the problem of reducing the amount of data required to represent a digital color image. The JPEG committee released a new image-coding standard, JPEG2000 that serves the enhancement to the existing JPEG system. The JPEG2000 use wavelet transform where as JPEG use Discrete Cosine Transformation (DCT) [1].

Artificial Neural Networks (ANNs) have been applied to many problems related to image processing, and have demonstrated their dominance over traditional methods when dealing with noisy or partial data. Artificial neural networks are popular in function approximation, due to their ability to approximate complicated nonlinear functions [2]. The multi-layer perceptron (MLP) along with the back propagation (BP) learning algorithm is most frequently used neural network in practical situations.

2. RECENT IMAGE COMPRESSION STANDARDS

This section presents a review of recent image coding standards like JPEG, JPEG 2000 by Joint Photographic Experts Group, the name of the committee that created the JPEG standard and WebP by Google. These compression formats use different transformation methods and encoding techniques like Huffman coding and run length encoding. The performance of compression standard is measured in terms of reconstructed image quality and compression ratio.

2.1. JPEG Standard

JPEG standard supports both lossless and lossy compression of color images. There are several modes defined for JPEG, including baseline, lossless, progressive and hierarchical. The baseline mode is the most accepted one and supports lossy coding only. The lossless mode is not accepted but provides for lossless coding, although it does not support lossy [3].

DOI : 10.5121/ijcsit.2011.3614

In the baseline mode, the image is divided in 8x8 blocks and each block is transformed with DCT (Discrete Cosine Transform). Then transformed block coefficients of every block are quantized with a uniform scalar quantizer , followed by scanning in zig-zag manner. The DC coefficients of all blocks are coded separately, using entropy coding. Entropy Coding achieves additional compression with Huffman coding and run length coding.

The `baseline JPEG coder is the sequential encoding in its simplest form. Figure 1 and Figure 2 show the process of an encoder and decoder for gray scale images. Similarly, color image (RGB) compression can be performed by compressing separate component images .

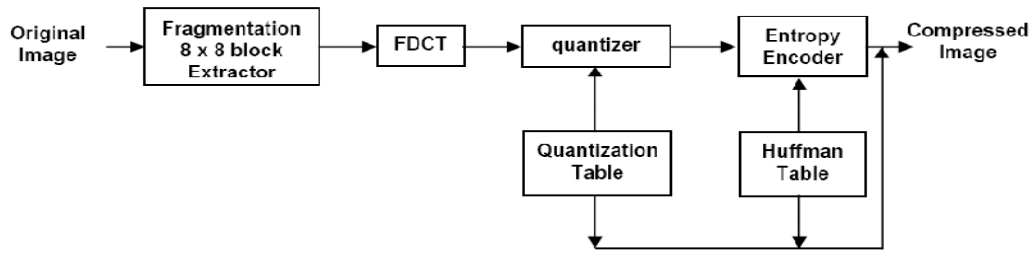


Figure 1. JPEG Encoder Block Diagram

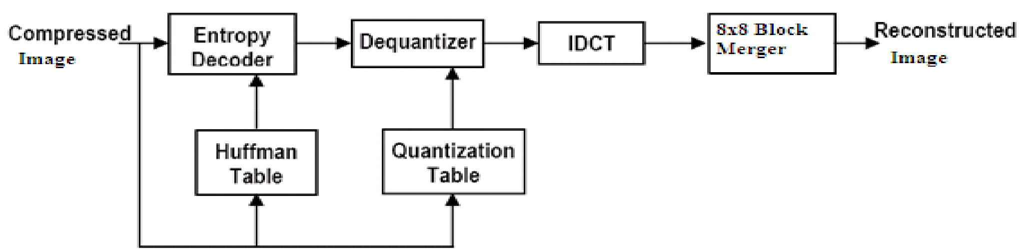


Figure 2. JPEG Decoder Block Diagram

2.2 JPEG 2000 Standard

The JPEG2000 working group hoped to create a standard which would address the faults of JPEG standard:

a. **Reduced low bit-rate compression**: JPEG offers outstanding rate-distortion performance in the average and high bit-rates, but at low bit-rates the biased distortion becomes unacceptable.
b. **Lossy and lossless compression**: There is currently no standard that can provide superior lossless and lossy compression in a single code-stream.
c. **Large image handling**: JPEG does not allow for the compression of images larger than 64K by 64K without tiling.
d. **Transmission in noisy environments**: JPEG was created before wireless communications became an everyday reality, therefore it does not acceptably handle such an error prone channel
e. **Computer-generated images**: JPEG was optimized for natural images and does not perform well on computer generated images.
f. **Compound documents**: JPEG shows poor performance when applied to bi-level (text) imagery.

Thus, the aim of the JPEG2000 working group is to develop a new image coding standard for deferent types of still images (bi-level, greyscale, color, multi component, hyper component), with different characteristics (natural, scientific, remote sensing, text rendered graphics, compound, etc) preferably within a unified and integrated system. This coding system is intended for low bit-rate applications and will exhibit rate-distortion and subjective image quality performance superior to existing standards.

As technology developed, it became clear that JPEG was not properly evolving to meet current needs. The widening of the application area for JPEG led to confusion among implementers and technologists, resulting in a standard that was more a list of components than an integrated whole. Afterwards, attempts to improve the standard were met with a naively sated marketplace. It was clear that improvement would only take place if a radical step forward was taken.

This coding system provide low bit-rate operation with rate-distortion and biased image quality performance superior to existing standards, without sacrifice performance at other points in the rate-distortion spectrum, incorporating at the same time many contemporary features [4]. The JPEG 2000 standard specify a lossy coding scheme which simply codes the difference between each pixel and the predict value for the pixel. The sequence of differences is encoded using Huffman or run length coding. Unfortunately, the massive size of the images for which lossy compression is required makes it necessary to have encoding methods that can support storage and progressive transmission of images.

The block diagram of the JPEG2000 encoder is illustrated in Figure 3 and 4. A discrete wavelet transform (DWT) is first applied on the source image data. The transformed coefficients are then quantized and coded using Huffman coding tree, before forming the output code stream (bit stream). At the decoder, the code stream is first decoded, dequantized and inverse discrete transformed, providing the reconstructed image data. The JPEG2000 can be both lossy and lossless [5]. This depends on the how wavelet transform and the quantization is applied.

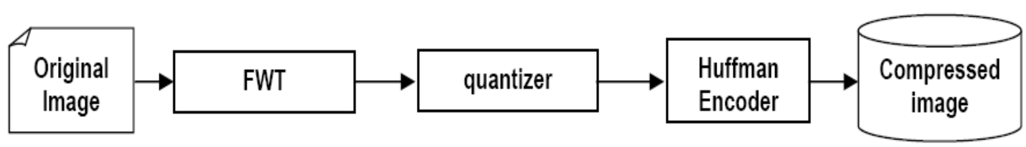


Figure 3. JPEG 2000 encoder

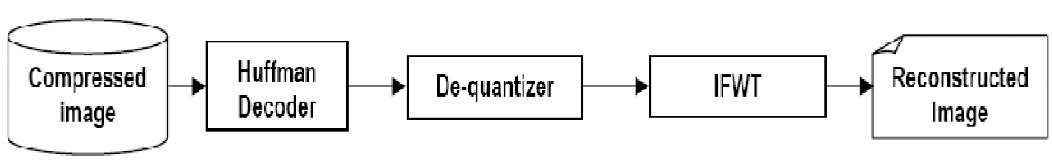


Figure 4. JPEG 2000 decode

2.3 WebP

Google has introduced an experimental new image format called WebP. The format is intended to reduce the file size of lossy images and the format delivers an average reduction in file size of 39 percent [6].

WebP still-image compression method uses VP8 video codec uses to compress individual frames [7]. It breaks the image into blocks (although 4-by-4 in size, rather than 8-by-8), but in place of JPEG's DCT. The encoder saves the predictions and the differences between the predictions and the real input blocks in the output file - if prediction is going well, as it should for most continuous-tone images like photos [8].

3. NEURAL NETWORKS

Neural-networks are getting applied to many fields from humanities to complex computational methods. The biological neural models are used in Artificial Neural Networks (ANNs). There is between 50 and 500 different types of neurons in our brain, they are mostly specialized cells based upon the basic neuron [9]. The Neuron consists of synapses, the soma, the axon and dendrites. Synapses are associations between neurons - they are not physical connections, but miniscule gaps that allow electric signals to move across neurons. These signals are then passed across to the soma which performs some operation and sends out its own electrical signal to the axon. The axon then distributes this signal to dendrites. Dendrites carry the information out to the various synapses, and the process repeats. Similar to biological neuron, basic artificial neuron has a certain number of inputs in input layer, each of which has a weight assigned to them. The net value of the neuron is then calculated. The net value is simply the weighted sum, the sum of all the inputs multiplied by their specific weight. Each neuron has its own unique threshold value, and it the net is greater than the threshold, the neuron outputs a 1, otherwise it outputs a 0. The output is then fed into all the neurons connected to it.

3.1 Image Compression with Back-Propagation Neural Networks

Apart from the existing traditional transform based technology on image compression represented by series of JPEG, MPEG, and H.26x standards, new technology such as neural networks and genetic algorithms are being developed to explore the future of image coding [10]. Many applications of neural networks to vector quantization have now become well established, and other aspects of neural network involvement in the area of image coding is to play significant roles in assisting with those traditional compression techniques [11].

3.1.1. Related Work with Basic Back-Propagation Neural Network

A neural network is an organized group of neurons. An artificial neuron is shown in Figure 5.Supervised learning learns based on the objective value or the desired outputs. During training the network tries to match the outputs with the desired target values [12]. This method has two types called auto-associative and hetero-associative. In auto-associative learning, the target values are the same as the input values, whereas in hetero-associative learning, the targets are generally different from the input values [13]. One of the most commonly used supervised Neural Net model is back propagation network that uses back propagation training. Back propagation algorithm is one of the well-known algorithms in neural networks [14].

Training the network is time consuming. It usually learns after several epochs, depending on size of the network. Therefore, large and complex network requires more training time compared to the smaller one. The network is trained for several epochs and stopped after reaching the maximum epoch. Minimum error tolerance is used provided that the differences between network output and known outcome are less than the specified value [15].

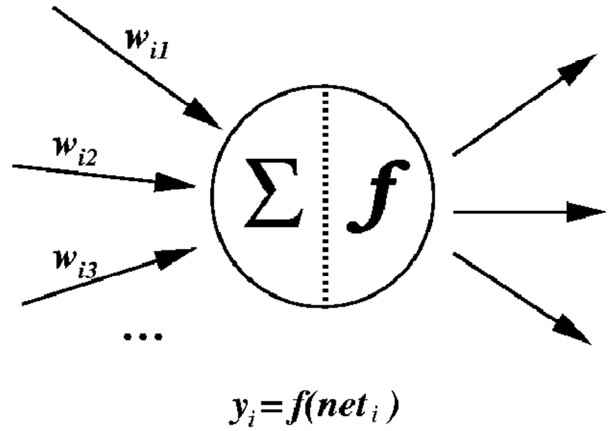


Figure 5. Artificial neuron

Depending upon the problem variety of Activation function is used for training. Using a nonlinear function which approximates a linear threshold allows a network to approximate nonlinear functions using only small number of nodes as shown in figure 5.

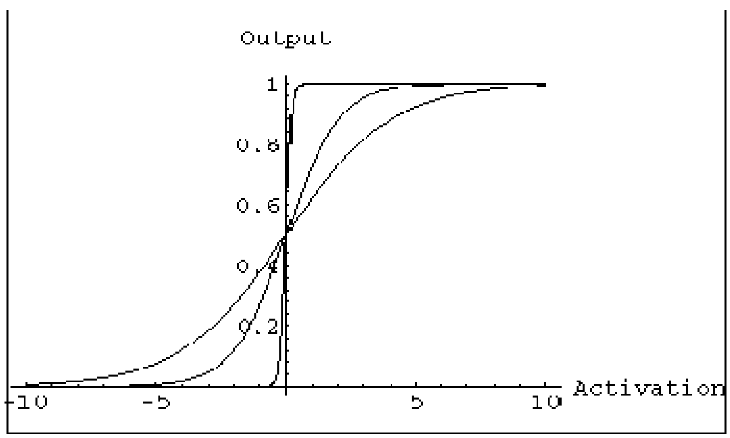


Figure 6. Sigmoid function

Learning/Training Neural Networks means adjustment of the weights of the connections such that the cost function is minimized.

Cost function:

$$\hat{C} = (\Sigma(x_i - x_i')^2)/N$$

Where x_i are desired output and x_i's are the output of the neural network.

Back Propagation Algorithm is summarized as follows.

Repeat:

1. **Choose training pair and copy it to input layer**
2. **Cycle that pattern through the net**
3. **Calculate error derivative between output activation and target output**
4. **Back propagate the summed product of the weights and errors in the output layer to calculate the error on the hidden units**
5. **Update weights according to the error on that unit**

Until error is low or the net settles

Three layers, one input layer, one output layer and one hidden layer, are designed. Input and output layer are fully connected to the hidden layer which is in-between. Compression is achieved by designing the value of *K*, the number of neurons at the hidden layer, less than that of neurons at both input and output layers. The input image is split up into blocks or vectors of 8×8, 4×4 or 16×16 pixels.

4. A NOVEL HYBRID NEURAL IMAGE COMPRESSION TECHNIQUE

Before the processing of image data the image are pre-processed to improve the rate of operation for the coding system. Under pre-processing clustering of the original image is carried out [16]. The term "clustering" refers to the partition of the original (source) image into no overlapping blocks, which are compressed independently, as though they were entirely distinct intensities. The efficiency of above technique is pattern dependant and images may contain a number of distinct intensities [17]. The neighbourhood pixels may have small difference of their pixel values. It is important that a Neural Net should converge quickly with high compression ratio. To achieve this, input sub image is pre-processed to minimize the difference in the gray levels of the neighbouring pixels to reduce computational complexity of Back Propagation Neural Image Compression. The steps of pre-processing are as follows.

Input : The set of N unique gray levels in image and the numbers of gray levels to be selected.

Output: Selected representative gray levels for image.

1. Initially all of the gray levels in the sub image are in same cluster, and its representative gray level is arbitrarily chosen.
2. Gray level that is farthest away from representative gray level (i.e max gray level) is taken as representative gray level of new cluster.
3. Gray levels which are closest to new cluster than to the existing one is moved to newly formed cluster.

Above steps are repeated until the number of clusters is equal to the number of selected gray levels.

The representative gray levels of each cluster are the selected gray levels. These selected gray levels are used to map the original sub image to pre-processed image [18].

A simple example of 3X3 pixel sub image when applied to above proposed approach is given below.

Let the one dimensional representation of two dimensional sub image matrix A with 7 unique gray levels in image and the numbers of gray levels to be selected is 5.

A=[45,20,90, 30,27,20,43,90,85]

Sorted array with unique gray levels : 90 85 45 43 30 27 20 and the clustering as follows

✓	90 85	20 27 30 43 45				
✓	90 85	20 27 30			43 45	
✓	90 85	20	27 30	43 45		
✓	90	85	20		27 30	43 45

Representative gray levels are 90, 85,20,30,45

Pixel mapping by pre-processing sub image is B=[45,20,90,30,30,20,45,90,85]

The normalized pixel values of above 3X3 Sub image B will be the input to a node. Similarly for all 8X8 sub images of input image to be compressed will be pre-processed and the normalized pixel values will be input to nodes X1,X2,X3, ……,X64 which will help to converge the neural network at a comparatively faster rate than simple Back propagation based image coding[19].

Neural networks for image compression seem to be well suited to this particular function, as they have the ability to pre-process input patterns to produce simpler patterns with fewer components. This compressed information (stored in a hidden layer) preserves the full information obtained from the external environment [20].

The network used for image compression is divided into two parts as shown in Figure 3.3. The transmitter encodes and then transmits the output of hidden layer (only K values as compared to N of the original image). The receiver receives and decodes the K hidden outputs and generates the N outputs. Since the network is implementing an identity map, the reconstruction of the original image is achieved [21].

In order to get a higher compression, we quantized the hidden layer output into 8 or 16 distinct level, coded on 3 or 4 bits. In that way, the original N bytes are compressed to 3K/8 or 4K/8 bytes and getting a high compression rate [22]. To decompress a compressed image, the output layer of the network is used. The binary data is decoded back to real number and propagated through the N neurons layer to output, then converted from range (-1.0 , +1.0) to positive integer values between 0 and 255 (grey level). The neural network structure can be illustrated in Figure 3.3. Three layers, one input layer, one output layer and one hidden layer, are designed. Compression is achieved by designing the value of the number of neurons at the hidden layer, less than that of neurons at both input and output layers [23].

When training process is completed and the coupling weights are corrected and the test image is fed into the network and compressed image is obtained in the outputs of hidden layer. The outputs must be applied to the correct number of bits. The same number of total bits is used to represent input and hidden neurons, and then the Compression Ratio (CR) will be the ratio of number of input to hidden neurons. For example, to compress an image block of 8×8, 64 input and output neurons are required. In this case, if the number of hidden neurons is 16 (i.e. block image of size 4× 4), the compression ratio would be 64:16=4:1.

The equal number of neurons in the input layer and output layer of the network are fully connected to the hidden layer with weights initialized to small random values from -1 to +1.Each input unit receives input signal and sends it to the hidden unit by using activation function. The output of the hidden unit is will be calculated by bipolar sigmoid function and send the output signal from the hidden unit to the input of the output layer units. The output signal is computed by applying the activation function.

In the error propagation phase, the error of the difference between the training output and the desired output is calculated by using error tolerance. Based on the calculated error correction term the weights are updated by back propagation [24]. With updated weights again the error is calculated and back propagated. The training is continued until the error is less than tolerance. Larger value of learning rate may train network faster but may result in overshoot and vice-versa. In back-propagation of error proper value of learning rate is very important for stopping condition of neural net.

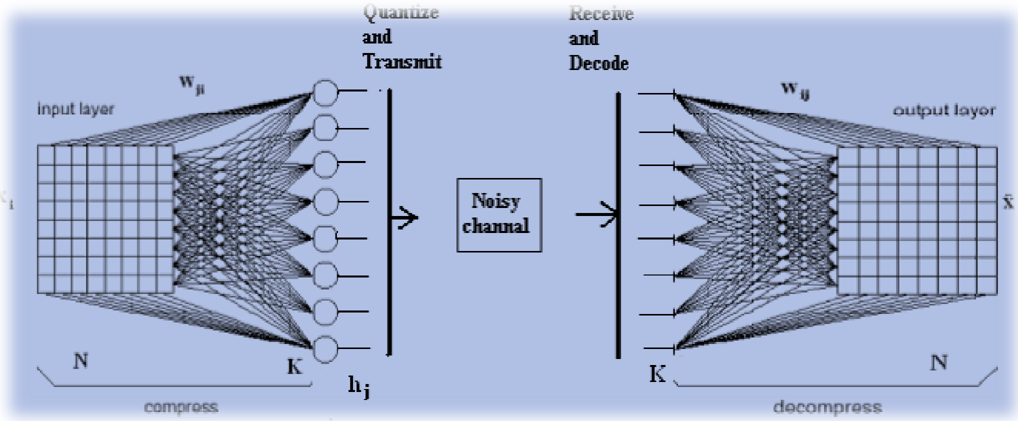


Figure 7. A neural network compression/ decompression pair.

5. Simulation Results

The following conditions are used for computer simulations. The images used consists of 256X256(=65536) pixels, each of which has its color values represented by 24 bits. The desired quality of the decompressed image and the required compression ratios can be achieved by fixing the error tolerance to any level, considering the time of simulation. Mapping the image by our pre-processing approach has helped the Back- Propagation Neural Network to converge easily compared to previous Techniques.

To compare the various image compression techniques, we have used the Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) as error metrics. The MSE is the cumulative squared error between the compressed and the original image. The PSNR is a measure of the peak error [25].

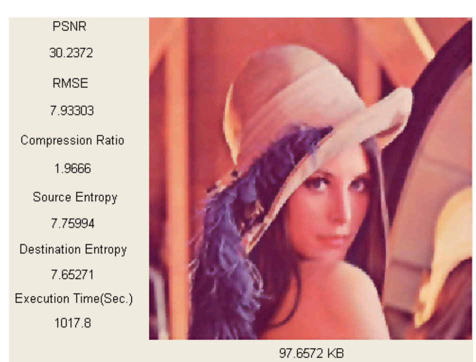
$$\mathrm{MSE} = \frac{1}{\mathrm{MXN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f'(x,y) - f(x,y)]^2 \tag{1}$$

Where $f'(x,y)$ is the compressed image and $f(x,y)$ is the original image with MXN pixels

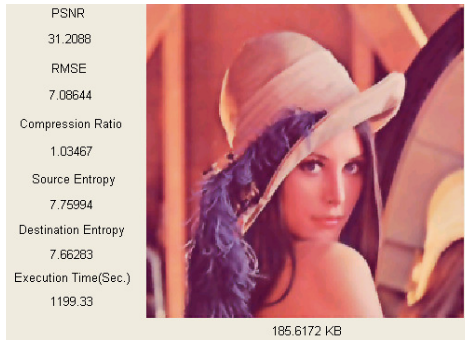
$$\mathrm{PSNR} = 10 \log_{10} \left(\frac{255X255}{\mathrm{MSE}} \right) \text{ dB} \tag{2}$$

The Quality of reconstructed image and the Convergence time of the Back propagation Neural Network were compared for both the conditions; image without pre-processing by clustering and with clustering. The PSNR of compressed image without pre-processing is around 20dB and the convergence time is 340seconds.

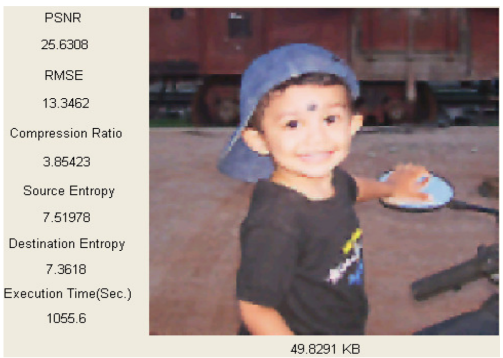
(a)



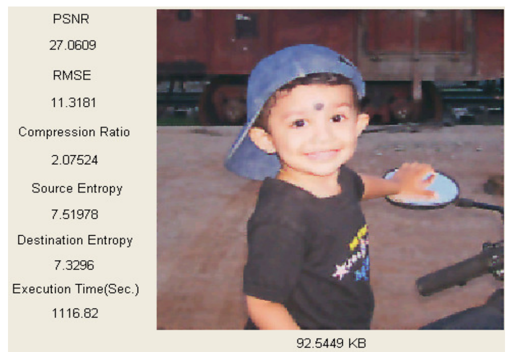
(b)



(c)



(d)



(e)

(f)

Figure 8. Pre-processed and Compressed images with
(a) (d) number of hidden layers = 4
(b) (e) number of hidden layers = 8
(c) (f) number of hidden layers = 16

6. CONCLUSION

With pre-processing PSNR value is increased slightly but the convergence time for training simulation has been reduced as compared to simple back propagation neural image compression. Experimental results of the proposed pre-processing technique i.e. pre-processed back propagation neural image compression (PBNN) showed an enhancement in performance measures with respect to decoded picture quality and compression ratios, compared to the existing JPEG and JPEG2000 as shown Table 1. The compression ratio of PBNN is better than JPEG and PSNR is better than JPEG 2000.

Standard Name	Lena		Lee	
	C.R.	PSNR	C.R	PSNR
JPEG	10.22	40.63	10.16	38.54
JPEG 2000	12.09	34.11	12.43	32.30
PBNN	11.64	24.74	11.84	23.33

Table 1: Comparisons of PBNN with JPEG and JPEG 2000

Back-propagation neural network algorithm is an iterative technique for learning the relationship between an input and output. This algorithm has been successfully used in many real-world applications; however, it suffers from slow convergence problems. A pre-processing technique is presented for image data compression for improved convergence as shown in Table 2. The quality of images using this technique is better as compared to simple back propagation neural image coding technique as shown in Figure 8.

No. of Iteration	No. of Block size	No. of Hidden layers	PSNR	RMSE	Compression Ratio	Execution Time(Sec)
100	4	4	28.23	9.97	3.78	527
		8	29.79	8.34	1.96	616
		16	30.32	7.85	1.02	734
	8	4	24.74	14.9	11.64	149
		8	26.58	12.04	6.2	165
		16	27.91	10.33	3.3	193
150	4	4	28.26	9.9	3.8	818
		8	30.12	8.02	1.94	854
		16	30.85	7.38	1.04	951
	8	4	24.84	14.7	11.66	202
		8	26.8	11.75	6.25	231
		16	28.36	9.83	3.32	267
200	4	4	28.33	9.84	3.88	1123
		8	30.23	7.93	1.97	1017
		16	31.2	7.08	1.03	1199
	8	4	24.87	14.68	11.6	275
		8	26.96	11.65	6.32	289
		16	28.61	9.56	3.37	317

Table 2: PBPN Training Results

Type	Average PSNR obtained	Average Compression Percentage
JPEG 2000	39.49	27.67
JPEG	39.36	22.37
PBNN	29.32	24.21

Table 3. Performance Comparisons

REFERENCES

[1] Rafael C. Gonzalez, Richard Eugene Woods ,Digital Image Processing PHI,2008 Pankaj N. Topiwala , Wavelet image and video compression Springer, 1998

[2] B. Carpentieri, M. J. Weinberger, and G. Seroussi. Lossless compression of continuous tone images. Proceedings of the IEEE, 88(11):1797–1809, November 2000.

[3] Dr. Weidong Kou Springer , Digital image compression: algorithms and standards, 1995 Subhasia Saha, "Image Compression – from DCT to Wavelets: A review", ACM Cross words students magazine, Vol.6, No.3, Spring 2000.

[4] J. Jiang. Image compression with neural networks – a survey. Signal Processing: Image Communication, 14, 1999.

[5] M. H. Hassoun. Fundamentals of Artificial Neural Networks. MIT Press, 1995.

[6] http://webp.codeplex.com/

[7] http://en.wikipedia.org/wiki/WebP

[8] http://code.google.com/speed/webp/

[9] Anil K. Jain ,"Fundamentals of Digital Image Processing", Prentice-Hall of India, India, 2001.

[10] Al Bovik , Handbook of image and video processing, Academic Press, 2005

[11] Khalid Sayood. "Introduction to Data Compression", second edition, Academic press, 2000.

[12] Andrew B. Wattson, " Image Compression Using the Discrete Cosine Transform", NASA Ames Research Center, Mathematical Journal, Vol.4, No.1, pp. 81-88, 1994.

[13] Z.Fan and R.D.Queiroz. Maximum likelihood estimation of JPEG quantization table in the identification of Bitmap Compression. IEEE, 948 951,2000.

[14] T. Cover and J. Thomas, Elements of Information Theory. NewYork: John Wiley & Sons, Inc., 1991.

[15] W. Fourati and M. S. Bouhlel. A novel approach to improve the performance of JPEG 2000. GVIP journal, Volume 5, Issue 5,pp.1-9,2005.

[16] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," Signal Processing: Image Communication, Vol. 17, Elsevier Science B.V., pp. 3-48, 2002.

[17] C. S. Burrus, R. A. Gopinath, and H. Guo, Introduction to Wavelets and Wavelet Transforms: A Primer. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[18] Z. Xiong, K. Ramchandran, M. T. Orchard, and Y.-Q. Zhang, "A comparative study of DCT— and wavelet-based image coding," IEEE Transactions on Circuits and Systems for Video Technology, vol. 9, no. 5, pp. 692–695, 1999.

[19] R. R. Coifman and M. V. Wickerhauser, "Entropy based algorithms for best basis selection," IEEE Transactions on Information Theory, vol. 32, pp. 712–718, Mar. 1992.

[20] D S. Taubman and M. W. Marcellin. JPEG 2000: Image Compression Fundamentals, Standards and Practice. Kluwer International Series in Engineering and Computer Science. Kluwer Academic Publishers.

[21] M. H. Hassoun, Fundamentals of Artificial Neural Networks, MIT Press, Cambridge, MA, 1995.

[22] Amerijckx, C., Velerysen, M., Thissen, P. and Legat, J.D. (1998). Image compression by self-organized Kohonen map.IEEE Transactions on Neural Networks, 9 (3), 503–507.

[23] S. Panchanathan, T. H. Yeap, and B. Pilache, "Neural network for image compression," in Applications of Artificial Neural Networks III, vol. 1709 of Proceedings of SPIE, pp. 376–385, Orlando, Fla, USA, April 1992.

[24] Haykin, S. (2003). Neural Networks: A Comprehensive Foundation. New Delhi: Prentice Hall India

[25] Ashanta Ranjan Routray and Munesh Chandra Adhikary. "Image Compression Based on Wavelet and Quantization with Optimal Huffman Code. International Journal of Computer Applications (ISSN 0975 – 8887) 5(2):6–9, August 2010. Published By Foundation of Computer Science, USA"

Authors

Ashanta Ranjan Routray is a Lecturer in Fakir Mohan University, Balasore, Orissa. He has completed his degree in Electrical engineering in 1999, received his M.E. degree in Computer Science and Engineering from Government College of Engineering, Tirunelveli, Tamilnadu in 2001 and submitted PhD in June 2011. His research interest includes Information Theory and Image Coding. He has already published many research papers in referred National and International journals.

Flt. Lt. Dr. Munesh Chandra Adhikary is a Reader and Head of PG Department of APAB, Fakir Mohan University, Balasore, India. He has receiced his PhD degree in Computer Science from Utkal University, Orissa. His research interest includes Neural Networks, fuzzy logic and image processing. He has already published many research papers in referred National and International journals.