International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Optimization of Image Compression Technique: Huffman Coding by Using PCDA

Rakhi Seth[1], Sanjivani Shantaiya[2]

[1, 2] Department of CSE,Chattisgarh Swami Vivekanand Technical University Bhilai, Disha Institute of Management and Technology, Raipur, C.G

Abstract*: Image Compression is a technique which removes the irrelevant data. Image Compression can be done by using PCA and LDA separately. In this paper, we proposed the new improved algorithm in which PCA and LDA is used as combined approach as PCDA with Huffman Coding by using this new algorithm we get the better Compression Ratio as well as Time taken for Compression is less.*

Keywords: Image Compression, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Huffman Coding, Compression Ratio

1. Introduction

An image is a picture which contains highly redundant and irrelevant information which is stored in an electronic form. An Image processing is any form of signal processing for which the input is an image such as a colour image, photograph, or any video frame; the output of image processing may be an image or a set of characteristics.

In Image processing there is one specific area through which size of data will be reduced which we call **Image Compression.** Image Compression[12] is defined as the technique through which one can remove the irrelevant and redundant data from image which help in storing and transmitting the data in more efficient manner. Compression is achieved by the removal of three types of basic data redundancies:

1. **Coding Redundancy**: is a type of redundancy which comes when less than optimal code words are to be used in an image.
2. **Interpixel Redundancy:** is a redundancy which results from the correlation between the two or more pixels of an image.
3. **Psycho visual Redundancy**: comes due to data which is ignored by the human visual system.

Image Compression is categorized mainly into two:

1. **Lossless Compression Method:** Lossless Compression method [6] is a class of compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. The lossless compression methods are:
   - Run Length Coding
   - Huffman Coding
   - LZW Coding
   - Area Coding
2. **Lossy Compression Method:** Lossy Compression method is a class of compression algorithms in which after compression original data is permanently loss and that's why we call this compression technique a lossy compression technique. The lossy compression method are:
   - Transformation Coding
   - Vector Quantization
   - Fractal Coding

2. PCDA Approach

PCDA is a combined approach of PCA with LDA and it works as follows:

2.1 Principal Component Analysis (PCA)

"Principal component analysis (PCA) [15] is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set". PCA is a way of identifying patterns in data and expressing the data in such a way as to highlight their similarities and differences. Once we have found the patterns in the data then we can compress the data by reducing the number of dimension. Algorithm of PCA is as follows:

Step 1: Calculate the mean of an image.

Step 2: The PCA is begun by correcting the image to that its column have zero means and unitary variances.

Image corrected by the Mean=image-mean of the image

Step 3: The covariance matrix is calculated by:

Covarianceof anImage=Image Corrected by the Mean × (Image Corrected by the Mean)^T

Step 4: Then the corresponding eigen values and eigen vectors are calculated.

Step 5: Then a matrix is obtained, that contain the list of eigenvectors of the covariance matrix.

Step 6: Final data are obtained by:

Finaldata=(MatrixObtainedinStep4)^T × (Image-Mean)^T

2.2 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis [1] easily handles the case where the with-in class frequencies are unequal and their performance has been examined on randomly generated test data.

LDA is also closely related to Principal Component Analysis (PCA) and factor analysis in that they both look for linear combinations of variables which best explain the data. LDA explicitly attempts to model the difference between the classes of data. PCA on the other hand does not take into account any difference in class, and factor analysis builds the feature combinations based on differences rather than similarities. Discriminant analysis is also different from factor analysis in that it is not an interdependence technique: a distinction between independent variables and dependent variables (also called criterion variables) must be made. For PCDA approach we have to use the algorithm of LDA viz:

Step 1: Find the mean of the resultant Principal Component Matrix.

Step 2: Average of the Mean.

Step 3: Calculate the variance.

Step 4: Find eigen vector.

Step 5: After finding the eigen vector, we have to multiply it by the resultant of Principal Component Analysis Matrix.

3. Huffman Coding

Huffman Coding Technique [4] is a technique which works on both data and image for compression. It is a process which usually done in two passes. In first pass, a statistical model is going too built, and then in second pass the image data is encoded which is produced by that statistical model. The problem in Huffman Coding Technique is come during the first pass in which statistical model is built and applied on raw data through which the process become slow down and effect the efficiency and accuracy of the technique because all these depends on the statistical model so our main problem which we have to rectify in our project is these statistical model and do optimization in these Huffman encoding in terms of increasing the accuracy of this statistical model. Huffman Coding Technique is easy to implement and most popularly used lossless technique but there are certain other problem which arises due to the first pass i.e. this technique becomes relatively slow and other problems are like overhead due to Huffman tree. The algorithm for Huffman coding is:

Step 1: Read the image on to the workspace of the mat lab.
Step 2: Convert the given colour image into grey level image.
Step 3: Call a function which will find the symbols.
Step 4: Call a function which will calculate the probability of each symbol.

Step 5: Probability of symbols are arranged in decreasing order and lower probabilities are merged and this step is continued until only two probabilities are left and codes are assigned according to rule that; the highest probable symbol will have a shorter length code.

Step 6: Further Huffman encoding is performed i.e. mapping of the code words to the corresponding symbols will result in a compressed data.

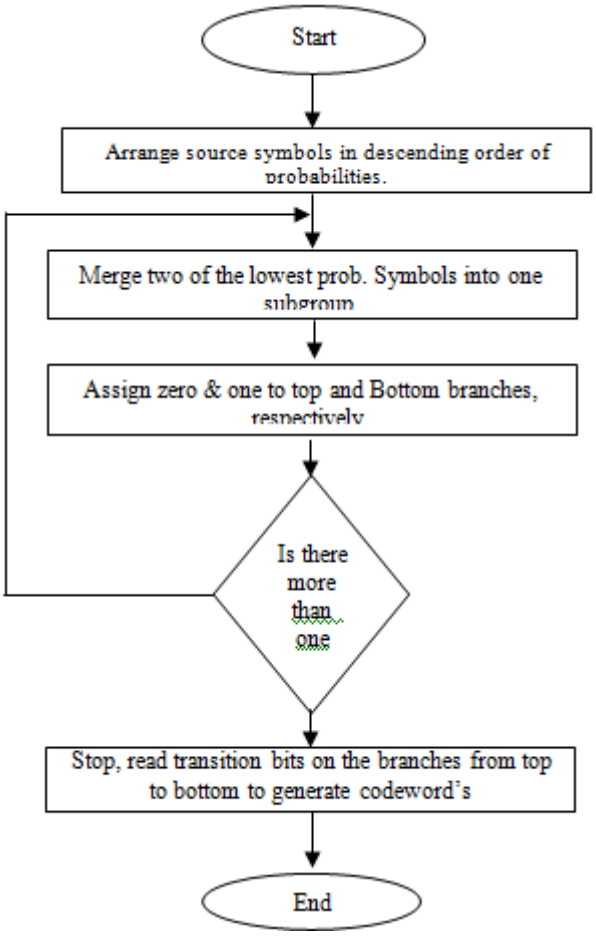

Figure 1: Flow Chart of Huffman Coding

4. Compression Ratio

Benchmark in image compression is the compression ratio. The compression ratio is used to measure the ability of data compression by the comparing the size of the image being compressed to the size of the original image. The greater the compression ratio means better quality compression we get.

5. Result

Here, we are using different square images and then apply PCA on it, On the resultant Data we apply LDA. And at last we apply Huffman coding on it, the images that we got are:

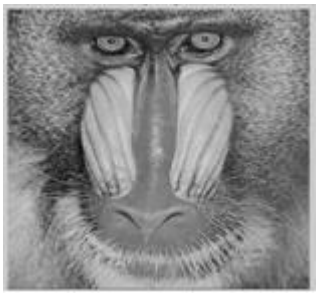

Figure 2: Original Image

Figure 3: After Applying PCA

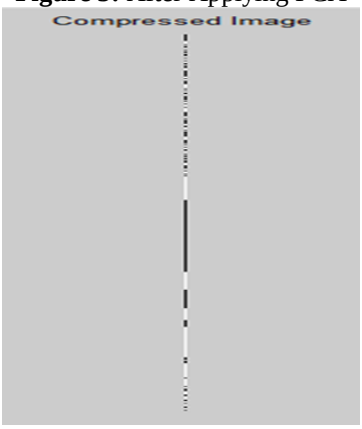

Figure 4: Compressed Image After Applying PCDA with Huffman Coding

The compression Ratio that we get on different images are,

Table 1: CR between When we apply Direct Huffman on the image and When we apply Huffman after applying PCDA

S.No.	Name of the Image	Compression Ratio	
		Huffman	PCDA+Huffman
1.	Lenna	1.58	2.46
2.	Baboon	0.577	2.46
3.	PeppersGray	1.83	2.41
4.	Cameraman	0.98	1.59
5.	Iris	1.2	1.59

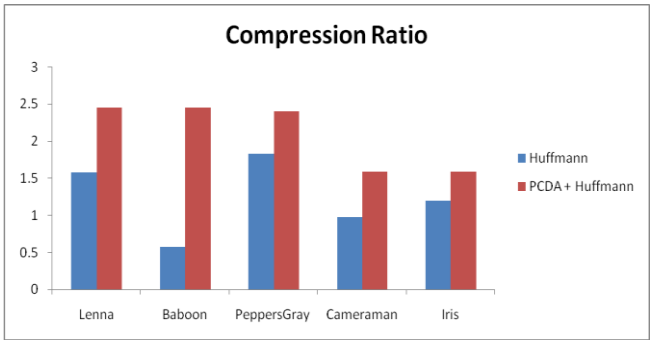

Figure 5: CR between When we apply Direct Huffman on the image and When we apply Huffman after applying PCDA

Table 2: Time Taken In Compression between When we apply Direct Huffman on the image and When we apply Huffman after applying PCDA

S.No.	Name of the Image	Time Taken for Compression in Seconds	
		Huffman	PCDA+Huffman
1.	Lenna	80.71	63.59
2.	Baboon	147.8	64.82
3.	PeppersGray	75.27	64.26
4.	Cameraman	49.01	13.38
5.	Iris	47.39	13.41

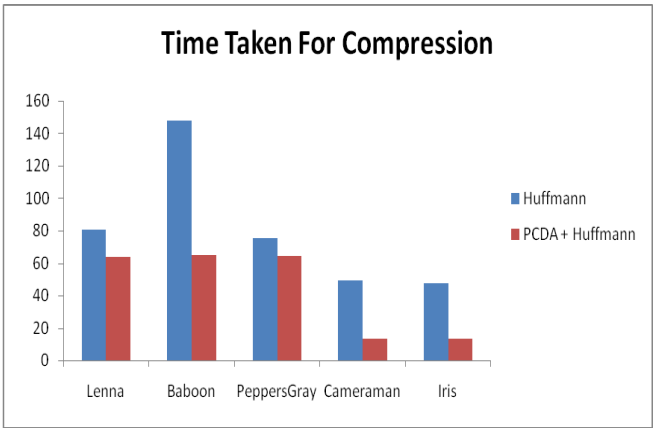

Figure 6: Time Taken for Compression between When we apply Direct Huffman on the image and When we apply Huffman after applying PCDA

6. Conclusion and Future Work

After implementing the proposed methodology We get the better results in terms of "Compression Ratio" as well as "Time taken For Compression" through which we conclude the above result of implementing PCDA in Huffman Coding makes Compression of an image much faster and gives Higher Compression Ratio. This work can be further extended for Decompression of an Image which may provide better results in terms of MSE (Mean Square Error) as well as higher PSNR (Peak Signal Noise Ratio). It can also be implemented in the Colour Images as well as its scope is also there with other Lossless and Lossy Compression Techniques.

References

[1] Hong-Bo Deng, Lian-Wen Jin, Li-Xin Zhen, Jian-Cheng Huang "A New Facial Expression Recognition Method Based on Local Gabor Filter Bank and PCA plus LDA" International Journal of Information Technology Vol. 11 No. 11, pp.86-96, 2005.

[2] Ashutosh Dwivedi, Arvind Tolambiya, Prabhanjan Kandula, N Subhash Chandra Bose, Ashiwani Kumar, Prem K Kalra "Color Image Compression Using 2-Dimensional Principal Component Analysis (2DPCA)" Proc. of ASID,pp.488-491,October 2006.

[3] L. Vasa and V. Skala "COBRA: Compression of the Basis for PCA Represented Animations" COMPUTER GRAPHICS forum Volume 28, number 6 pp. 1529–1540, 2009.

[4] Mamta Sharma "Compression Using Huffman Coding" IJCSNS International Journal of Computer Science and Network Security, VOL.10 No.5, May 2010.
[5] Prabhakar.Telagarapu, V.Jagan Naveen, A.Lakshmi.Prasanthi, G.Vijaya Santhi "Image Compression Using DCT and Wavelet Transformations" International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 4, No. 3, September 2011.
[6] Mridul Kumar Mathur, Seema Loonker, Dr. Dheeraj Saxena "Lossless Huffman Coding Technique for Image Compression and Reconstruction Using Binary Trees" IJCTA Vol 3 (1), 76-79, Jan-Feb 2012.
[7] Dr.E.KANNAN, G. Murugan "Lossless Image Compression Algorithm for Transmitting Over Low Bandwidth Line" International Journal of Advanced Research in Computer Science and Software Engineering Volume 2, Issue 2, February 2012.
[8] Ms. Pallavi M. Sune Prof. Vijaya K. Shandilya "Image Compression Techniques based On Wavelet and Huffman Coding" International Journal of Advanced Research in Computer Science and Software Engineering Volume 3, Issue 4, April 2013.
[9] Athira B. Kaimal, S. Manimurugan, C.S.C .Devadass "Image Compression Techniques: A Survey" International Journal of Engineering Inventions Volume 2, Issue 4, PP: 26-28, February 2013.
[10] K.N. Abdul Kader Nihal1, Dr. A.R. Mohamed Shanavas "A Survey and Study of Image Compression Methods" IOSR Journal of Computer Engineering (IOSR-JCE), Volume 16, Issue 4, PP 11-16, Ver. V (Jul – Aug. 2014)
[11] Maryam Imani, Hassan Ghassemian "Principal Component Discriminant Analysis for Feature Extraction and Classification of Hyperspectral Images" IEEE, 2014.
[12] Er. Shruti Puniani, Er. Nishi Madaan "Various Image Compression Techniques: A Review" International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3 Issue 4, April 2014.
[13] Sunita S Biswal, Krishna Kalpita, Dipak R. Swain "Comparative Study on Image Compression Using Various Principal Component Analysis Algorithms" International Journal of Scientific and Research Publications, Volume 4, Issue 5, May 2014.
[14] Dr. Sanjay Kumar, Er. Ankur Chauhan "A Survey on Image Feature Selection Techniques" International Journal of Computer Science and Information Technologies, Vol.5 (5), 2014.
[15] Neethu Mohan "Removal of PCA Based Estimated Noise in Processed Images" International Journal of Science and Research (IJSR) Volume 3 Issue 10, pp. 897-899, October 2014.
[16] https://liorpachter.wordpress.com/tag/probabilistic-pca/