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

# A Survey on Image Compression Methods with PCA & LDA

**Rakhi Seth[1], Sanjivani Shantaiya[2]**

[1] M. Tech. Scholar, Department of CSE, Disha Institute of Management and Technology, Raipur, C.G, India

[2]Assistant Professor, Department of CSE, Disha Institute of Management and Technology, Raipur, C.G, India

**Abstract:** ***This research paper presents extensive survey on various techniques of Image Compression using both PCA and LDA. LDA utilizes the label information in finding informative projection. It is used to find a linear combination of features which characterizes or separates two or more classes of objects or events. PCA is a statistical tool, which may increase the efficiency .In the study it may come as an aspect that it can be used as a combined approach in Image Compression also.***

**Keywords:** Image Compression, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Lossless Compression, Lossy Compression.

## 1. Introduction

Image is an array or matrix of square pixels arranged in the form of rows and columns. It is a virtual representation of these square pixels. An image is a picture which contains highly redundant and irrelevant information which is stored in an electronic form. There are three basic components of image formation:

1. The illumination.
2. The reflectance model of surface which are imaged.
3. The process of image formation.

Once the images are formed, the next process involves sampling and quantization of the analog image. Modern digital technology has made it possible to manipulate multidimensional signals. The goal of this manipulation is categorized into three categories:

1. Image Processing.
2. Image Analysis.
3. Image Understanding.

Our main focus is on Image Processing so here we discuss about Image Processing and their methods. 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 technique, the image is mostly presented in 2-D format. Image Processing are of two types:

1. Digital Image Processing.
2. Analog Image Processing.

Most commonly used image processing in these two image processing is **Digital Image Processing**. In Image processing there is one specific area through which size of data will be reduced which we call **Image Compression.** Image Compression 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.
2. Interpixel Redundancy.
3. Psycho visual Redundancy.

**Coding Redundancy** is a type of redundancy which comes when less than optimal code words are to be used in an image.

**Interpixel Redundancy** is a redundancy which results from the correlation between the two or more pixels of an image.

**Psycho visual Redundancy** comes due to data which is ignored by the human visual system.

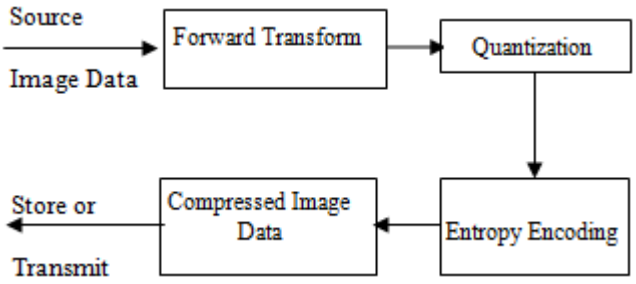


**Figure 2:** General Flow of Image Compression

Image Compression is categorized mainly into two:
1. Lossless Compression Method.
2. Lossy Compression Method.

**Lossless Compression Method**

Lossless Compression method is a class of compression algorithms that allows the original data to be perfectly reconstructed from the compressed data. Types of Lossless Compression Methods are as follows:

1. **Run Length Coding:** It is one of the simplest data compression methods which are based on the principle that the run of characters is replaced with the number of the same character and a single character.

2. **Huffman Coding:** It is based on the concept of a variable length code should use the shortest code words for the most likely symbols and the longest code words for the least likely symbols.
3. **LZW Coding:** It is the foremost technique for general purpose data compression due to its simplicity and versatility. LZW to compress text, executable code, and similar data files to about one-half their original size.
4. **Area Coding:** Area coding is a technique or enhanced form of run length coding. In the algorithms of area coding try to find rectangular regions with the same characteristics.

**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. Types of Lossy Compression Methods are as follows:

1. **Transformation Coding:** DFT and DCT is a type of transforms which are used in changing the pixels of the original image into frequency domain coefficients. There are several properties in this type of coefficients. One is the compaction property. This is the basis for achieving the compression.
2. **Vector Quantization:** This is the method to develop a dictionary of Fixed- size vectors called code vectors. An image is divided into non-overlapping blocks and then for each value dictionary is determined as well as index is generated for the dictionary which is used as the encoding for an input original image.
3. **Fractal Coding:** Fractal coding introduces the idea of decomposition of an image into segments by using standard methods of image processing like color separation, edge detection and texture analysis. Each segment is stored in a library of fractals.
4. **Block Truncation Coding:** In this method, firstly the image is divided and then arrange a block of pixels and find a threshold and reconstruction values for each block then a bitmap of the block is derived and all those pixels got replaced which have the value greater than or equal to the threshold value by 1or 0.
5. **Sub band Coding:** The image got analysed as to produce the components which contain frequencies of well-defined blocks and sub bands. Quantization and coding is applied to each of the bands and then each sub bands will be designed separately.

## 2. Principal Component Analysis (PCA)

"Principal component analysis (PCA) is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set". In statistics, PCA is a method for simplifying a dataset by reducing multidimensional datasets to lower dimensions for analysis. PCA is a standard method commonly used for data reduction in statistical pattern recognition and signal processing. PCA is noted to be one of the most valuable results from applied linear algebra. It is used abundantly in all forms of analysis from neuroscience to computer graphics, because it is a simple non - parametric method of extracting relevant information from confusing datasets. PCA is also called the KARHUNEN-LOEVE Transform (KLT, named after Kari Karhunen & Michel Loeve) or the HOTELLING Transform. Its general objectives are:

1. Data reduction
2. Interpretation

## 3. Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis utilizes the label information in finding informative projection. It is used to find a linear combination of features which characterizes or separates two or more classes of objects or events. LDA seeks to reduce dimensionality while preserving as much of the class discriminatory information as possible. 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.

## 4. Related Work Using Image Compression with PCA and LDA

Hong-Bo Deng et al. (2005) proposes a facial expression recognition system based on Gabor feature using a novel local Gabor feature using a novel local Gabor filter bank. In order to evaluate the performance of the local Gabor filter bank, we first employed a two-stage feature compression method PCA plus LDA to select and compress the Gabor feature. They proves this, by showing the experimental results which shows both dimension reduction and good recognition performance they achieve the best average recognition rate of 97.33% [1].

A.Dwivedi et al. (2006) found that 2-Dimensional Principal Component works directly on the matrices and it deals with the high dimensional vector space data .They work on the principle of applying 2DPCA for color image compression. Method is tested on several standard images and found that the quality of reconstructed image is better than standard PCA based image compression. The other performance measures, such as computational time, compression ratio are ameliorated. A comparative study is made for color image compression using 2DPCA [2].

L.Vasa et al. (2009) provides an extended dynamic mesh compression technique based on PCA and works on the size of the basis cannot be neglected when considering the overall performance of a compression algorithm. In this paper they address this issue and provide the more sophisticated algorithm such as LPC and they proved this by providing a new approach which reduces the size of the basis by 90% with respect to direct encoding, which can lead to

approximately 25% increase in performance of compression algorithm [3].

Mamta Sharma al. (2010) studied about the source coding that it is the process of encoding information using fewer bits and also a use of uncoded representation for specific encoding schemes. Compression is a technology which reduces the number of bits required to store and/or transmit digital media. There are various techniques available for compression as she has analysed in her paper. The Huffman coding algorithm is compared with other compression techniques like Arithmetic, LZW and Run length Encoding [4].

Prabhakar Telagarapu et al. (2011) observed that Image Compression is a widely addressed research area. Many compression standards are in place. But still here there is a scope for high compression with quality reconstruction. The JPEG standard makes use DCT for compression. The introduction of the wavelets gave different dimensions to the compression. This paper aims at the analysis of compression using DCT and Wavelet transform by selecting proper threshold method [5].

Mridul kumar Mathur et al. (2012) studied carefully that image play an indispensable role in representing vital information and need to be saved for further use. In order to have efficient utilization of disk space and for better transmission rate of an image, compression is used and there are many algorithms is present for compression. In this paper compressing a BMP file which is on Delphi to implement Huffman coding especially in gray scale image [6].

Dr.E.Kannan et al. (2012) proposed the idea for decreasing the communication bandwidth and save the transmitting power in the wireless endoscopy capsules; this paper presents a new near lossless image compression algorithm based on the Bayer format image. This algorithm can provide low average compression rate with high image quality for endoscopic images. The proposed technique produces a bit stream that results in a progressive and ultimately lossless reconstruction of an image similar to what one can obtain with a reversible wavelet codec. In addition, the proposed scheme provides near-lossless reconstruction with respect to a given bound after decoding of each layer of the successively refundable bit stream. Experimental results for both lossless and near-lossless cases indicate that the proposed compression scheme that innovatively combines lossless, near-lossless and progressive coding attributes, gives competitive performance in comparison to state-of-the-art compression schemes [7].

Ms.Pallavi M.Sune et al. (2013) found that Image compression is now essential for applications such as transmission and storage in database. They discuss about the image classification, wavelet compression and convert an image into an array using Delphi image control tool. They proposed an algorithm which is created in Delphi to implement Huffman coding and concluded that image compression is a valuable technique. They also concluded that Huffman coding is bet technique in lossless compression and Huffman Coding is best technique in lossless compression and Huffman work on the two passes to compress the file. The Wavelet Compression Engine was used in this study [8].

Athira B.Kaimal et al. (2013) forecasted performances of the various image compression techniques. These techniques are basically classified into two categories one is Lossy and other one is Lossless. Lossy compression technique mainly used to compress the multimedia data like audio, video especially in applications such as streaming media. Lossless compression technique is used for text and data files, such as bank records and text articles. In many cases it is advantageous to make a master lossless file that can then be used to compressed files for other purposes [9].

K.N Abdul Kader Nihal et al. (2014) have studied that there are several redundancies in an Image which can be removed by using various Image Compression techniques i.e. Lossless and Lossy, one can improve the compression technique by using PCA in two different approaches namely, Statistical Approach and PCA Neural Approach. PCA have been used mainly, removing Data reduction and Interpretation and after removing this from an image. They have studied that if we have 25 images each with N2 vectors and 25 dimensions. Now, PCA can be implemented on this set of data then only 20 eigen vectors out of 25 eigen vectors are used for compressing the data [10].

Maryam Imani et al. (2014) observed that PCA is one of the most conventional unsupervised feature extraction methods which extract features with the largest power. PCA discards the components of data with small variances may have useful information for discrimination between classes in classification process. They propose to apply the linear-discriminant analysis to those components of PCA which have small power. So they extract the information components for classification instead of discarding them. They proved this from the experimental results obtained by using two hyperspectral data that PCDA than both as an approach work more efficiently than using PCA and LDA as a separate approach. [11].

Er.Shruti Puniani et al. (2014) provide the brief review over the various image compression techniques like Huffman encoding, LZW coding ,VQ compression and some recently used new hybrid technique through which they do the critical analysis of various compression techniques as well as concluded that how the wavelet is the best technique in case of hybrid compression technique [12].

Sunita S Biswal et al. (2014) found that PCA is one of the statistical methods employed in image compression. Presented paper deals with four different types of PCA algorithms those are 2D-PCA, 3D-PCA, 2D- Kernel PCA, 3DKPCA. A comparative study is made for all four types of PCA based on their PSNR values. These algorithms are also tested on several standard test images. It has been found that quality of reconstructed image of 3DKPCA is better than other types of PCA based image compression [13].

Dr Sanjay kumar et al. (2014) studied that PCA, LDA, ICA are the methods which is used for the pattern extraction from

different kind of data type i.e. Images and data sets. In addition of that, these techniques are helpful in classifying the object based on their extracted pattern [14].

Neethu Mohan (2014) found that noise is one of the important problems in image processing applications in which most popular one is Blind noise level estimation. He proposes a system in which there is new noise level estimation and removal method, for estimating noise uses an approach called PCA in image blocks.PCA is a statistical technique which is frequently used in signal processing for reduction in dimension of data or for data correlation. The proposed idea work on the basis of finding variance estimate [15].

## 5. Conclusion

From the above discussion we conclude that, PCA is a statistical tool which is used to increase the efficiency and LDA is used as classifier for better image quality. There are several Image Compression methods which we want to improve by using PCA and LDA as a combined approach.

## 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/