

COMPARATIVE ANALYSIS: PCA AND JPEG

Shashank Kumar¹, Annpurna Singh², Manish Mishra³, Geetika Srivastava⁴

¹Department of Electronics and Communication Engineering, Amity University, Lucknow

²Department of Electronics and Communication Engineering, Amity University, Lucknow

³Department of Electronics and Communication Engineering, DDU Gorakhpur University, Gorakhpur

⁴Department of Electronics and Communication Engineering, Amity University, Lucknow

Abstract

This paper mainly address the image compression by using Principal component analysis (PCA) and JPEG. Image compression is the method of converting data file into smaller compact files for efficiency of storage and transmission. The main objective of image compression is to reduce redundancy of bits in the image in order to store and transmit data in an effective way. Image compression techniques deal with the problem of reduction in size of the file which results in efficient storage of image and bandwidth to transmit it. Image Compression helps in storage of more information in a given memory space. PCA is mathematical tool which is used for reducing the dimensionality of data .PCA extracts major variation in the data sets while removing other insignificant components. In this paper PCA is applied for image compression. Since PCA compression is lossy the compressed image is generally degraded as compared to original image. Also PCA is compared with most commonly used image compression technique which is JPEG. JPEG is also a lossy image compression technique which start by separating the image into 8*8 pixel groups and applying DCT transform individual 8*8 pixel groups .In this paper PCA and JPEG compression are applied to the image and various parameter like PSNR and MSE is calculated for both of the compression technique. The Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) are the two error parameter used to compare image compression quality

Keywords:- Image Compression, PCA, JPEG, DCT, PSNR, MSE

-----***-----

1. INTRODUCTION

Image Compression means reduction in the size of image in such a way that there is less or acceptable level of degradation in the quality of image. The reduced image requires less storage space as compared to the parent image. Thus there is efficient use of memory which is a very important resource for any processing unit. Image Compression is also required to transfer images rapidly over internet. As compressed image requires less bandwidth then original image. The image compression is entirely different than that of the compression of text files or programs. The compression of text files or images should be lossless as if it is lossy then the entire sense of information is lost or error in the program results in mal functioning of the program. However small loss of information introduced in the image does not results in noticeable change in the quality of the image. Therefore image compression can be lossy in nature. In this paper we are studying lossy image compression technique i.e. image compression using JPEG technique and image compression using principal component analysis. PCA is mathematical tool which is used for reducing the dimensionality of data .PCA extracts major variation in the data sets while removing other insignificant components. PCA is compared with most commonly used image compression technique which is JPEG. JPEG is also a lossy image compression technique which start by separating the image into 8*8 pixel groups and applying DCT transform individual 8*8 pixel groups .In this paper PCA and JPEG compression are applied to the image and various parameter like PSNR and MSE is calculated for both of the

compression technique. The Mean Square Error (MSE) and the Peak Signal to Noise Ratio (PSNR) are the two error parameter used to compare image compression quality

2. IMAGE COMPRESSION CONCEPTS:

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

Lossless Compression Method In Lossless Compression we can ideally extract original image from Compressed image with no distortion. Types of Lossless Compression Methods are as follows:

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

2. Huffman Coding: In this coding variable code length is used that is we use shortest code words for the most frequent symbols and the longest code words for less frequent symbols.

3. LZW Coding: LZW stands for Lempel–Ziv–Welch .The codes from 256-4095 are used to represent the sequence of bytes. This compression is mainly used to compress GIF images .It has find its application in compression of executable code, text files, also files which are similar or repeating themselves.

4. Area Coding: Area coding is a method or improved form of run length coding. In Area Coding rectangular area of similar characteristics is extracted.

Lossy Compression Method: Lossy Compression is compression algorithms in which after compression original data is always damaged. Types of Lossy Compression Methods are as follows:

1. Transformation Coding: In transformation coding the pixels of the image or information is transformed into another domain. For example in DCT the pixels are transformed into frequency domain. The coefficients of the DCT is computed. The coefficient has compaction property which is used to compress the image.

2. Vector Quantization: This is the technique which can develop a dictionary of Fixed- size vectors called code vectors. An image is distributed into non-overlapping blocks and then for each value dictionary is obtained as well as index is produced for the dictionary which is used as the encoding for an input original image.

3. Fractal Coding: In fractal coding image is divided into segments by using standard image processing technique like detection of edge, separation of colour, and analysis of text. These segment is used to generate the library of fractals.

4. Block Truncation Coding: Image is divided into blocks and quantizer is used to reduce the number of grey level in the image. This is done in a way that mean and standard deviation remain unaltered.

3. PRINCIPLE COMPONENT ANALYSIS (PCA)

PCA is process which is used for mainly minimize the no of variables or dimensions in Image compression. PCA computes most appropriate system of axis using the Singular Vector Decomposition. PCA is used because it computes most significant data characteristics with little loss of information.. Consider an example in which there are 25 images. The individual image has N^2 pixel. These N^2 pixel is converted into column vector .The final data is compressed say into vector of 20 dimension by applying PCA .Their is loss of information as the processed image has less information compared to the compressed image .Since there is a loss of information the PCA compressed is called lossy compression method. The compressed image is distorted version of the original image.

4. JPEG (TRANSFORM COMPRESSION)

JPEG stands for Joint Photographic Expert Group. JPEG is based on fact derived by fourier that low frequency content has more information or they are more significant as compared to the high frequency content. In JPEG image is broken into blocks of 8X8. These 8X8 bits are processed individually i.e this block or group is processed as 64 bytes. This 64 bytes are then converted into 2 to 20 bytes. This 2 to 20 bytes is a representation or approximation of the 8X8 group. 8X8 group is preferred over other group because 8X8 group is standardized group which an integrated circuit can process. The KLT transform or PCA transform has more compression ratio than JPEG transform but the process of KLT is more typical as compared to JPEG transform.

5. DISCRETE COSINE TRANSFORMS (DCT)

Image is broken into different frequency component using DCT. Quantization is the one of the most significant part of compression where less significant frequency term is discarded. This is the reason why we use the word "lossy". In decompression process most significant frequency components are used for the retrieval of the image. Since some of the frequency term is removed their is little distortion in the image. This distortion is compensated during the compression process. The following is a general overview of the JPEG process.

1. Pre processing of the image is done that is if the image is coloured we convert it into gray scale image. This image is partitioned into 8x8 blocks.
2. The heart of JPEG is DCT. In the partitioned 8x8 block DCT is applied
3. Quantization of block is done to compress the image.
4. Compressed block requires very little memory for storage.
5. Inverse Discrete Cosine Transform (IDCT) is used to decompress the image.

This equation represents the basic equation of DCT

$$F(u) = \left(\frac{2}{N}\right) \sum_{i=0}^{\frac{1}{2}N-1} A_i(i) \cos[\pi \cdot u/2 \cdot N(2i + 1)]$$

Equation:-Basic Equation Of DCT

PSNR AND MSE:- Image quality is compared using PSNR and MSE. MSE gives conglomerated squared error between the original image and the compressed image, whereas Peak error is computed using PSNR.

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_1^2}{MSE} \right)$$

$$= 20 \cdot \log_{10} \frac{(MAX)}{\sqrt{MSE}}$$

$$= 20 \cdot \log_{10}(MAX) - 10 \cdot \log_{10}(MSE)$$

For lossy compression PSNR is used to compare the image quality . For human visual point of view PSNR gives the approximation of the image quality. Higher PSNR indicates higher quality but this is not always true.

6. EXPERIMENTAL RESULT



Fig 1 ORIGINAL IMAGE FOR PCA COMPRESSION

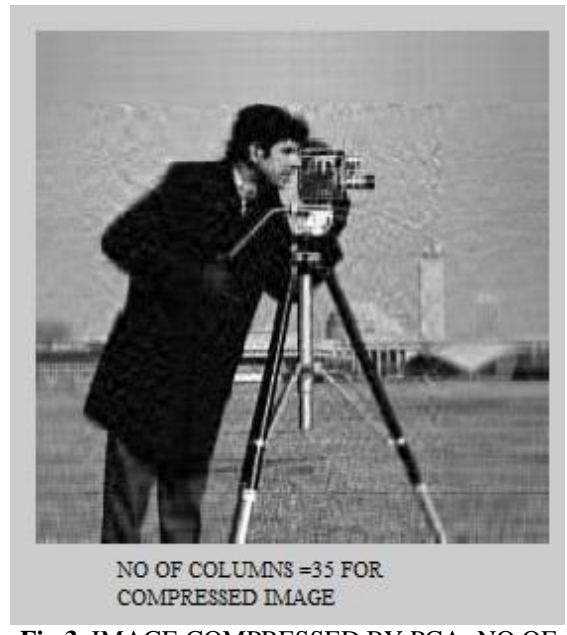


Fig 3 IMAGE COMPRESSED BY PCA :NO OF COLUMNS =35



Fig 2 IMAGE COMPRESSED BY PCA :NO OF COLUMNS =20



Fig 4 IMAGE COMPRESSED BY JPEG :QUALITY FACTOR=30



Fig 5 ORIGINAL IMAGE FOR JPEG COMPRESSION



Fig 6 IMAGE COMPRESSED BY JPEG :QUALITY FACTOR=15

7. RESULT AND DISCUSSION

In this paper PCA dimensionality reduction is applied to the image compression and parameters like PSNR and MSE has been defined. The PCA compression has been compared with the commonly used compression technique i.e JPEG. In PCA it was found that as the number of columns taken for the compression in PCA increases the perceptual quality of image increase. However the size of image also increases. In case of JPEG image it was noticed that the image perceptual quality improves as the quality factor increases..

8. CONCLUSION

In this paper we implemented image compression using PCA and JPEG .The parameter MSE and PSNR was defined for PCA and JPEG. The PCA uses dimensionality reduction for compression .PCA extracts major variation in the data

sets while removing other insignificant components and JPEG uses Discrete Cosine Transform for image compression..

REFERENCES

- [1] Subramanya A, "Image Compression Technique," Potentials IEEE, Vol. 20, Issue 1, pp. 19-23, Feb-March 2001.
- [2] 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.
- [3] Sunita S Biswal, Krishna Kalpita, Dipak R. Swain "Comparative Study on Image Compression Using Various Principal Component Analysis Algorithms" International Journal of Scientific and Reseach Publications, Volume 4, Issue 5, May 2014.
- [4] Dr. Sanjay Kumar, Er. Ankur Chauhan "A Surveyon Image Feature Selection Techniques" International Journal of Computer Science and Information Technologies, Vol.5 (5), 2014
- [5] Neethu Mohan "Removal of PCA5.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.
- [6] C.K. Li and H.Yuen, "A High Performance Image Compression Technique for Multimedia Applications," IEEE Transactions on Consumer Electronics, Vol. 42, no. 2, pp 239-243, 2 May 1996.
- [7] Anitha S, "2D image compression technique-A survey", International Journal of Scientific & Engineering Research Volume 2, Issue 7, pp 1-7 July-2011.
- [8] Ricardo L. de Queiroz "Processing JPEG-Compressed Images and Documents", IEEE Transactions On Image Processing, Vol. 7, No. 12, Pp1661-1667 december 1998.
- [9] Dr.B Eswara Reddy and K Venkata Narayana "A Lossless Image Compression Using Traditional and Lifting Based Wavelets", Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.2, and April 2012.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] Mamta Sharma "Compression Using Huffman Coding" IJCSNS International Journal of Computer

Science and Network Security, VOL.10 No.5, May 2010.

- [14] 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.