IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

INVESTIGATION OF IMAGE COMPRESSION USING DIMENSIONALITY REDUCTION TECHNIQUES

Dr. H S Prasantha (0000-0003-4739-517X)
Professor
Department of Electronics & Communication Engineering
Nitte Meenakshi Institute of Technology, Bangalore, Karnataka, India

Abstract: Image compression aims at reducing the number of bits required to represent an image by removing the spatial and spectral redundancies as much as possible. The focus of the research work is only on still image compression. The lossy compression methods which give higher compression ratio are considered in the research work. The paper discusses the image compression using different Dimensionality reduction techniques such as SVD (Singular value Decomposition), PCA (Principal Component Analysis), ICA (Independent Component Analysis) on various images of any type and resolution.. The proposed work investigates the use of different Dimensionality reduction techniques to achieve compression. The comparison of different methods are done using different comparison parameters such as MSE, PSNR and Compression Ratio.

***Index Terms* – Singular Value Decomposition, compression ratio, Principal Component Analysis, Independent Component Analysis, block size, image compression**

I. INTRODUCTION

Image compression is a type of data compression applied to digital images, to reduce their cost for storage or transmission. Its aim is to reduce the number of bits required to represent an image by removing the spatial and spectral components to lowest possible level. The redundancy and irrelevancy that is present in the image will be reduced during image compression. The objective is to reduce the transmission and storage requirements of image data. The two types of compression which describes image compression are Lossless compression and Lossy compression. Image compression is very important for efficient transmission and storage of images. Demand for communication of multimedia data through the telecommunications network and accessing the multimedia data through Internet is growing explosively. With the use of digital cameras, requirements for storage, manipulation, and transfer of digital images, has grown explosively. These image files can be very large and can occupy a lot of memory. Also, suitability of the transforms is due to subjective quality of the decompressed images in terms of PSNR (Peak signal to Noise Ratio) and quality index, computation time and energy compaction property.

Image Compression is minimizing the size of an image without degrading the quality of the image to an unacceptable level. The reduction in file size allows more images to be stored in a given amount of disk or memory space. It also reduces the time required for images/video to be sent over the internet or downloaded from web pages. The SVD is a fundamental concept in science and engineering, and one of the most central problems in numerical linear algebra. It is also known as principal component analysis (PCA) in statistics and the Karhunen-Loeve (KL) or Hotelling expansion in pattern recognition. The beauty of the SVD is that it provides a robust method of storing larger images as smaller square ones. This is accomplished by representing the original image with each succeeding non-zero singular values. To reduce the storage size even further, one may approximate a “good enough” image with using even fewer singular values. SVD, one of the most useful tools of linear algebra is a factorization and approximation technique which effectively reduces any matrix into smaller invertible and square matrix. SVD is preferred over DCT, Haar and other transforms is due to the suitability of SVD even though matrix is not invertible and non-square. Also DCT, Haar transforms are linear whereas SVD is nonlinear transformation.

Real-world data, such as speech signals, digital photographs, or MRI scans, usually has a high dimensionality. In order to handle this data adequately, its dimensionality needs to be reduced. Dimensionality reduction is the transformation of high-dimensional data into a meaningful representation of reduced dimensionality. Ideally, the reduced representation should have a dimensionality that corresponds to the intrinsic dimensionality of the data. The intrinsic dimensionality of data is the minimum number of parameters needed to account for the observed properties of the data. Dimensionality reduction is important in many domains, since it mitigates the curse of dimensionality and other undesired properties of high-dimensional spaces. As a result, dimensionality reduction facilitates, among others, classification, visualization, and compression of high-dimensional data. Traditionally, dimensionality reduction was performed using linear techniques such as Principal Components Analysis (PCA), Independent Component Analysis (ICA) and factor analysis. However, these linear techniques cannot adequately handle complex nonlinear data. Therefore, in the last decade, a large number of nonlinear techniques for dimensionality reduction have been proposed. In contrast to the traditional linear techniques, the nonlinear techniques have the ability to deal with complex nonlinear data. In particular for real-world data, these nonlinear dimensionality reduction techniques may offer an advantage, because real-world data is likely to be highly nonlinear. Previous studies have shown that nonlinear techniques outperform their linear counterparts on complex artificial tasks.

II. RESEARCH METHODOLOGY

Dimensionality reduction, or dimension reduction, is the transformation of data from a high-dimensional space into a low-dimensional space so that the low-dimensional representation retains some meaningful properties of the original data, ideally close to its intrinsic dimension. Working in high-dimensional spaces can be undesirable for many reasons; raw data are often sparse as a consequence of the curse of dimensionality, and analyzing the data is usually computationally intractable. Dimensionality reduction is common in fields that deal with large numbers of observations and/or large numbers of variables, such as signal processing, speech recognition, neuroinformatics, and bioinformatics. The different Dimensionality reduction techniques considered are SVD, PCA and ICA

2.1 SINGULAR VALUE DECOMPOSITION (SVD)

The SVD involves the decomposition of an image represented as matrix A in to U, S and V matrices where $UU^T = I, VV^T = I,$ I is a identity matrix, the columns of U are orthonormal eigenvectors of AA^T, the columns of V are orthonormal eigenvectors of A^TA and S is a diagonal matrix containing the square roots of Eigen values from U or V in descending order. The columns of U are the left singular vectors, S has singular values and is diagonal and V^T has rows that are the right singular vectors. The SVD represents an expansion of the original data in a coordinate system where the covariance matrix is diagonal. Decomposition of the image into U, S and V matrix is computationally very complex and for reconstruction multiplication of U, S and V matrix is also very complex. Hence method is proposed to minimize the complexity of the algorithm. A modified method proposed for preprocessing the SVD can be estimated in terms of the memory requirements and the computation time required in comparison with the SVD. The novel method of preprocessing reduces the computational complexity and also provides easy way of implementing SVD with reduced block size.

The following steps explain the algorithm for SVD

1. Let us consider a gray scale image of dimension M X N
2. Read the first image from set of images of equal sizes using a file pointer, file name, and read the image data i.e. im_data from file name.
3. Perform the SVD to image data which gives the value of U, S, V. U has size of MXM, S has size of MXN, and V has size of NXN.
4. Now depending upon the number of eigenvalues we are going to reconstruct the original image. If we fix a eigenvalue of 1, using the first singular values is image is reconstructed. Similarly for 4 eigenvalues, using the first four singular value images is reconstructed. It is repeated depending upon the number of eigenvalues we choose. If we select all the eigenvalues then the complete image is reconstructed.
5. After performing the SVD, calculate the MSE and PSNR

2.2 PRINCIPLE COMPONANT ANALYSIS (PCA)

Principal Component Analysis (PCA) is a statistical method usually used for reducing the dimensionality of data. "It involves replacing a group of series with a weighted average of those series, where the weights chosen so that the new vector (called the principal component or PC) explains as much of the variance of the original series as possible. This leaves a matrix of unexplained residuals, but this matrix can be reduced to a PC as well. In that case the original PC is called the first PC (PC1), and the PC of the residuals is called the second PC, or PC2. And there will be residuals from it too, yielding PC3, PC4, etc. The higher the number of the PC, the less important is the pattern it explains in the original data. PC1 is the dominant pattern; PC2 is the secondary pattern, etc. In many cases a large number of data series can be summarized with relatively few PCs."

Using PCA for image compression also knows as the Hotelling, or Karhunen and Leove (KL), transform. If we have 20 images, each with N^2 pixels, we can form N^2 vectors, each with 20 dimensions. Each vector consists of all the intensity values from the same pixel from each picture. This is different from the SVD because in case of SVD, we had a vector for image, and each item in that vector was a different pixel, whereas now we have a vector for each pixel, and each item in the vector is from a different image.
Now we perform the PCA on this set of data. We will get 20 eigenvectors because each vector is 20-dimensional. To compress the data, we can then choose to transform the data only using, say 15 of the eigenvectors. This gives us a final data set with only 15 dimensions, which has saved us 1/4 of the space. However, when the original data is reproduced, the images have lost some of the information. This compression technique is said to be lossy because the decompressed image is not exactly the same as the original.

The following steps describe the PCA algorithm:

1. Let us consider a gray scale image of dimension M X N
2. Read the set of images and calculate the mean of all images.
3. Calculate the variance of all the images from the mean of images.
4. Find the eigenvector and eigenvalues of the variance matrix.
5. Arrange eigenvalues in descending order and their corresponding eigenvectors.
6. Calculate the feature vector i.e. principal component by multiplying image data and eigenvectors (arranged in descending order).
7. Now depending upon the number of eigenvalues (dominant) we are going to reconstruct the original image. If we select 1 eigenvalue, feature vector multiplied by transpose of eigenvector matrix and image is reconstructed. Similarly for 4 Eigen values, feature vector multiplied by transpose of Eigen vector matrix and image is reconstructed. It is repeated depending upon the number of eigenvalues we choose. If we select all the eigenvalues then the complete image is reconstructed.
8. After performing the PCA, calculate the MSE and PSNR

2.3 INDEPENDENT COMPONENT ANALYSIS (ICA)

Independent component analysis (ICA) presents a probabilistic image model in which an observed random vector x containing pixels from an image can be decomposed as X=AS, here S is a vector containing independent sources, which are linearly combined into the observations X through the basis function A_i where the superscript i denotes the i^{th} column of A. In non-orthogonal paradigm, collection of atoms (or basis functions) is termed as a dictionary which may be incomplete, complete or over complete. The terms basis functions and atoms are used interchangeably. Where A is an N X M unknown mixing matrix and S is a vector of independent sources. The standard goal of ICA is to infer (learn) A from a set of samples of the random vector X. To apply ICA to images, each sample of X usually contains the pixels in an image block

The following steps describe the ICA algorithm:

1. Let us consider a gray scale image of dimension M X N
2. Read the image, make the image 4x4 blocks and calculate the mean of all the blocks of images.
3. Calculate the variance of all the block of images from the mean of all the blocks of images.
4. Find the eigenvector and eigenvalues of the variance matrix.
5. Calculate the projection value by multiplying block of image data and eigenvector.
6. Initialize maximum projection value to negative infinity. Compare this maximum projection value with the projection value calculated and store the greater value in maximum projection value.
7. Find the maximum projection value in a block by above method and store that value in ICA projection and its index. Subtract the corresponding data from block of data. Find the next greater value in a block of data and store that value, Subtract the corresponding data from block of data repeats this operation for all the data in a block.
8. Repeat step 5 to step 7 for all the blocks of image and store the greater value first in a blocks of images and next greater value will be stored in second time continues this operation for all the values in a block.
9. Depending upon the number of eigenvalues reconstruct the image back. i.e. if the number of Eigen value is 1 then get the maximum Eigen vector from each of the block in an image and reconstruct the image. Similarly if the number of Eigen value is 4 then get the first four maximum value of Eigen vector from the each of block and reconstruct the image

III. IMPLEMENTATION DETAILS

Figure 1: Input image considered for experimentation (112X92 size)

Different image file formats of different resolutions are considered for the experimentation. The different file formats such a TIFF, JPEG, PNG, BMP, etc. are considered for experimentation. Extensive experimentation is done using different file formats of different resolutions and by using SVD, PCA and ICA. Sample result obtained after applying different Dimensionality reduction techniques such as SVD, PCA and ICA are displayed for further discussions and analysis. The different dimensionality reduction techniques are compared with respect to RANK are compared in terms of MSE and PSNR. Figure 1 shows the different input image considered for experimentation

The comparison parameters considered are MSE and PSNR are given by

Mean Square Error (MSE): The Mean Square Error measures the difference between the frames which is usually applied to Human Visual System. It is based on pixel-pixel comparison of the image frames.

$$d(X,Y) = \frac{\sum_{i-1}^{m,n}{}_{j-1}(X_{i,j}-Y_{i,j})^2}{mn} \qquad (1)$$

Peak Signal to Noise Ratio (PSNR): PSNR is measured on a logarithmic scale and depends on the mean squared error (MSE) of between an original and an impaired image or video frame, relative to $(2^n-1)^2$ (the square of the highest-possible signal value in the image, where n is the number of bits per image sample).

$$PSNR_{db} = 10log_{10}\left(\frac{(2^n-1)^2}{MSE}\right) \qquad (2)$$

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

Table 1: Result obtained by applying SVD for the input images for different Ranks

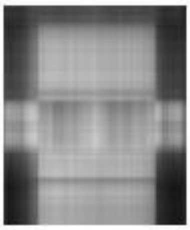
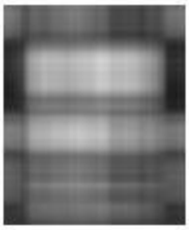
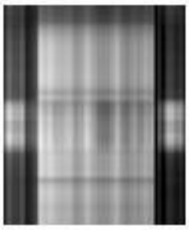
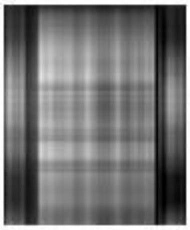
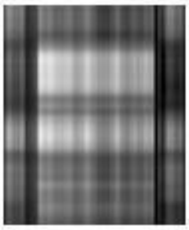
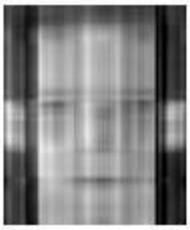
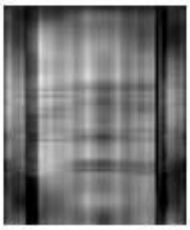
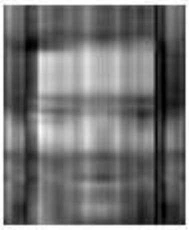
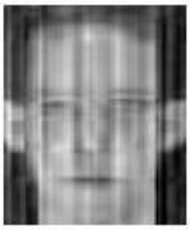
Reconstructed image for the singular value of 2		
		
MSE = 401, PSNR = 22	MSE = 637, PSNR = 20	MSE = 456, PSNR = 21.5
Reconstructed image for the singular value of 4		
		
MSE = 205, PSNR = 25	MSE = 395, PSNR = 22	MSE = 243, PSNR = 24.3
Reconstructed image for the singular value of 8		
		
MSE = 100, PSNR = 28.1	MSE = 221, PSNR = 24.7	MSE = 98.9, PSNR = 28.1
Reconstructed image for the singular value of 10		
		
MSE = 79, PSNR = 29.1	MSE = 179, PSNR = 25.6	MSE = 70.8, PSNR = 29.6

Table 2: Result obtained by applying PCA for the input images for different Eigen Values

Reconstructed image for the Eigen value of 2		
		
MSE = 1212, PSNR = 17	MSE = 1467, PSNR = 16.5	MSE = 1321, PSNR = 17
Reconstructed image for the Eigen value of 4		

		
MSE = 1034, PSNR = 18	MSE = 1227, PSNR = 17.2	MSE = 1034, PSNR = 18
Reconstructed image for the Eigen value of 8		
		
MSE = 893, PSNR = 21	MSE = 651, PSNR = 20	MSE = 511, PSNR = 21
Reconstructed image for the Eigen value of 10		
		
MSE = 443, PSNR = 21.7	MSE = 573, PSNR = 20.5	MSE = 466, PSNR = 21.4

In the proposed work for experimentation, the different image file formats such as JPEG, TIFF, PNG and BMP Formats of various sizes such as 128 X 128, 102 X 102 & 64 X 64 for image compression are chosen. For image compression dimensionality reduction algorithms like SVD, PCA and ICA are investigated. The quality measures such as MSE, PSNR, & CR etc. are used for performance consolidation. The three lossy techniques Singular Value Decomposition, Principal Component Analysis and Independent Component Analysis give better error resilience, scalability. Singular Value Decomposition and Principal Component Analysis avoids blocking artifacts, from the results of Mean Square Error and Peak Signal to Noise Ratio we can say that Singular Value Decomposition performs better than Principal Component Analysis and Independent Component Analysis and also Independent Component Analysis performs better then Principal Component Analysis. The table 1, 2, 3 shows the variation in MSE and PSNR with respect to rank (singular value or Eigen value) and table 4 shows the comparison of SVD, PCA and ICA with respect to MSE, PSNR and Compression Ratio. The objective of the proposed work is not to state which technique is better but to understand the apply of different techniques to achieve Image compression.

Table 3: Result obtained by applying ICA for the input images for different Eigen Values

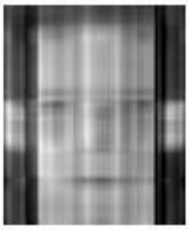
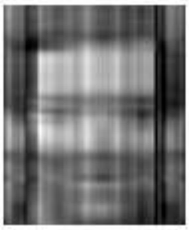
Reconstructed image for the Eigen value of 2		
		
MSE = 150, PSNR = 26.3	MSE = 246, PSNR = 24.2	MSE = 132, PSNR = 27
Reconstructed image for the Eigen value of 4		
		
MSE = 95, PSNR = 28.3	MSE = 145, PSNR = 26.5	MSE = 79.6, PSNR = 29

Table 4: Comparison Chart for SVD, PCA and ICA

SVD:

Number of Singular Value	**2**	**4**	**8**	**10**
MSE	401.0227	205.35	100.63	78.9083
PSNR	22.0991	25.0015	28.103	29.15
Compression Ratio	22.06	11.03	5.5	4.4

PCA:

Number of Eigen Value	**2**	**4**	**8**	**10**
MSE	1212.7	1034.2	893.16	442.91
PSNR	17.29	17.98	21.06	21.667
Compression Ratio	23	11.5	5.7	4.6

ICA:

Number of Eigen Value	**1**	**2**	**3**	**4**
MSE	202.76	150.61	118.66	94.715
PSNR	25.06	26.35	27.38	28.33
Compression Ratio	7.11	3.55	2.56	1.63

V. SUMMARY AND CONCLUSIONS

The overall aim of the paper is to understand the usage of different dimensionality reduction techniques such as SVD, PCA and ICA to achieve image compression. The different techniques are applied with variation in rank and parameters such as MSE and PSNR are noted. The dimensionality reduction techniques are having very good energy compaction property but the limitation is computational complexity. The computational complexity can be reduced by using fast computational procedures.

The computational complexity can be reduced by dividing the image into blocks as a preprocessing stage and the dimensionality reduction techniques are applied for each block.

REFERENCES

1. H. S. Prasantha, H. L. Shashidhara, and K. N. Balasubramanya Murthy. Image compression using SVD. In Proceedings of the International Conference on Computational Intelligence and Multimedia Applications, pages 143–145. IEEE Computer Society, 2007.
1. Abhiram Ranade, *et. al.*, "A Variation on SVD Based Image Compression", Image Vision Computation", 25(6), 20017, pp.771-777, 2007
2. Messaoudi, *et. al.*, "A Good Compromise between Images Quality and Compression Rate of the DCT Technique", GVIP Conference, CICC, Cairo, Egypt., 2005
3. Dr. Edel Garcia, "Singular Value Decomposition (SVD)" A Fast Track Tutorial Published on September 11 2006.
4. Niranjan Damera Venkata, "Image Quality Assessment Based on a Degradation Model", IEEE transactions on image processing, Vol. 9, No. 4, 2000.
5. Satu Jumisko, *et .al.*, "Evaluation of Subjective Video Quality of Mobile Devices", MULTIMEDIA '05, Proceedings of the 13th annual ACM international conference on Multimedia, 2005
6. Gunasheela K S, H S Prasantha, "Satellite image compressiondetailed survey of the algorithms", Proceedings of ICCR in LNNS Springer, 2017, vol. 14, pp. 187-198.
7. Raghavendra.M.J, Prasantha.H.S and S.Sandya, "Image Compression Using Hybrid Combinations of DCT SVD and RLE", International Journal of Computer Techniques, Volume 2 Issue 5-2015.
8. Gunasheela K S, H S Prasantha, "Compressive sensing for image compression: survey of algorithms", Proceedings of Emerging Research in Computing, Information, Communication and Applications, ERCICA, Springer publication, Bengaluru, 2018
9. K N Shruthi, B M Shashank, Y. SaiKrishna Saketh, H.S Prasantha and S. Sandya, "Comparison Analysis Of A Biomedical Image For Compression Using Various Transform Coding Techniques", IEEE, pp. 297-303, 2016
10. Raghavendra.M.J, Prasantha.H.S and S.Sandya, "DCT SVD Based Hybrid Transform Coding for Image Compression", International Journal of Recent and Innovative Trends in computing and communication. 2015
11. Prasantha H.S., Shashidhara H.L, Balasubra manya Murthy, K N: Comparative analysis of different interpolation schemes in image processing. In: International Conference on Advanced Communication Systems (ICACS), India, pp. 17–24 (2007)
12. Prasantha H.S, Shashidhara H.L and K.N.Balasubramanya Murthy, "Image Scaling comparison using Universal Image Quality Index", International Conference on Advances in computing, control and Telecommunication Technologies (ICACCTT), India, pp.859-863, Dec 2009.