Available online at http://www.ijabbr.com

International journal of Advanced Biological and Biomedical Research

Volume 1, Issue 7, 2013: 767-776

COMPRESSION OF BREAST CANCER IMAGES BY PRINCIPAL COMPONENT ANALYSIS

Monika Saraswat [1], A. K. Wadhwani [2] and Manish Dubey [3]

[1]Electrical, RGPV/MITS, Gwalior, M.P, India, [2]Electrical, RGPV/MITS, Gwalior, M.P, India, [3]Electrical, RGPV/MITS, Gwalior, M.P, India

ABSTRACT

The aim of this work, Early detection of the breast cancer and reduce the mortality rate. Breast cancer is the second most lethal cancer for women in the world today. Image compression is fundamental to the efficient and cost-effective use of digital medical imaging technology and applications. The conventional method for measuring quality and noise from the image by SNR and PSNR. Compressed the noise ratio from the cancer image and decomposed the PSNR value. The effective method Principal component analysis currently provides the most promising approach to high-quality image compression. Compressed the noise and improve the quality in the breast cancer image by PCA method.

Key Words: SNR, MSE, PSNR, Mammograms, PCA.

INTRODUCTION

Breast cancer is the most frequent cancer in women worldwide. The disease is curable if detected early enough. Screening is carried out on the basis of mammograms, which use x-ray images to reveal lumps in the breast. Calcium deposits can also indicate the existence of a tumor. However, the deposits are often only a few tenths of a millimeter in size and so deeply embedded in dense tissue that they are nearly undetectable in the images. The Digital Images to enhance the visual information which is a primary operation for almost all vision & image processing tasks in several areas such as Computer vision, biomedical image analysis and other fields [2]. PCA is a statistical procedure that uses an orthogonal property to transform to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables. The denoising phenomenon goal is to remove the noise while retaining the maximum possible the important signal or image features. At the time of acquisition and transmission the images are often corrupted by additive noise. The main aim of a denoising algorithm is to reduce the noise level, while preserving the image features. To achieve a good performance in this respect, a denoising algorithm has to adapt to image discontinuities. Generally the quality of image can be measured by the peak signal-to-noise ratio (PSNR). However, sometimes a denoised image with a high PSNR value does not have satisfactory visual quality [5].The objective of enhancement is to get finer details of an image and highlight the useful information. If a person acquires an image by using digital camera, the illumination such as fluorescent lamp in a room or sunlight in open air is kept to be uneven and

Corresponding Author E-mail: monimani52@gmail.com

uncontrollable. As a result the image degrades leaving excessively darker or brighter region in an image [4].

DATA ACQUISITION

Images of Mammogram are available in the Department of Electrical Engineering, MITS Gwalior. These images are available with the same specification (3000x4500 pixels with 16-bit pixel depth).

PRINCIPAL COMPONENT ANALYSES

The principle of dimensionality reduction with PCA is the representation of the dataset 'X' in terms of eigenvectors $e^i \in R^N$ of its covariance matrix. The eigenvectors oriented in the direction with the maximum variance of $\boldsymbol{X}$ in R^N carry the most relevant information of $\boldsymbol{X}$. These eigenvectors are called principal components [8]. Assume that n images in a set are originally represented in matrix form as $Ui \in R^{r \times c}$, i = 1,......,n, where r and c are, repetitively, the number of rows and columns of the matrix. In vectorized representation (matrix-to-vector alignment) each *Ui is* a N = r × c- dimensional vector *ai* computed by sequentially concatenating all of the lines of the matrix *Ui*. To compute the Principal Components the covariance matrix of *U* is formed and Eigen values, with the corresponding eigenvectors, are evaluated. The Eigen vectors forms a set of linearly independent vectors, i.e., the base $\{\varphi\}^n{}_{i=1}$ which consist of a new axis system [10].

PCA FOR IMAGE COMPRESSION

Using PCA for image compression also known as the Hotelling, or Karhunen and Leove(KL), transform. We have 10 images, each with N^2 pixels, we can form N^2 vectors, each with 10 dimensions. Each vector consists of all the intensity values from the same pixel from each picture [7]. This is different from the previous example because before 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 [11]. Now we perform the PCA on this set of data. We will get 10 eigenvectors because each vector is 10-dimensional. To compress the data, we can then choose to transform the data only using, say 5 of the eigenvectors. This gives us a final data set with only 5 dimensions, which has saved us ¼ of the space. However, when the original data is reproduced, the images have lost some of the information [6]. This compression technique is said to be lossy because the decompressed image is not exactly the same as the original, generally worse [9].

ERROR METRICS

Two of the error metrics used to compare the various image compression techniques is the Mean Square Error (MSE) given in equation 1, and the Peak Signal to Noise Ratio (PSNR) to achieve the desirable compression ratios. The MSE is the cumulative squared error between the compressed and the original image, whereas PSNR is a measure of the peak error [1]. A lower value for MSE means lesser error, and as seen from the inverse relation between the MSE and PSNR, this translates to a high value of PSNR. Logically, a higher value of PSNR is good because it means that the ratio of Signal to Noise is higher. Here, the 'signal' is the original image, and the 'noise' is the error in reconstruction [3].

MEAN SQUARE ERROR (MSE)

Mean square error is a criterion for an estimator the choice is the one that minimizes the sum of squared errors due to bias and due to variance. The average of the square of the difference between the desired response and the actual system output. In a loss function, MSE is called squared error loss. It can be estimated in one of many ways to quantify the difference between values implied by an estimate and the true quality being certificated. MSE is a risk function corresponding to the expected value of squared error. The MSE is the second moment of error and thus incorporates both the variance of the estimate and its bias.

$$MSE= (1/ (M\times N)) \times ((k(x, y))-(XX(x, y)))\ \hat{}2 \qquad (1)$$

Where m and n is the image size and k(x, y) is the input image and XX(x, y) is the retrieved image.

PEAK SIGNAL-TO-NOISE RATIO (PSNR)

The PSNR is most commonly used as a measure of quality of reconstruction of lossy compression codecs (e.g., for image compression). The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codecs it is used as an approximation to human perception of reconstruction quality, therefore in some cases one reconstruction may appear to be closer to the original than another, even though it has a lower PSNR (a higher PSNR would normally indicate that the reconstruction is of higher quality). One has to be extremely careful with the range of validity of this metric; it is only conclusively valid when it is used to compare results from the same codec (or codec type) and same content [12]. It is the ratio between the maximum possible power of a signal and the power of the corrupting noise, mathematically given in equation 2. Because many signals have a very wide dynamic range, PSNR is usually expressed in terms of the logarithmic decibel scale. The PSNR is most commonly used as a measure of quality of reconstruction in image compression etc. It is most easily defined via the mean squared error (MSE) which for two m×n monochrome images I and K where one of the images is considered noisy.

$$PSNR=10\times\log_{10} ((MAX_i\hat{}2)/MSE) \qquad (2)$$

Here, MAX_i is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. For colour images with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three. Typical values for the PSNR in Lossy image and video compression are between 30 and 50 dB, where higher is better. PSNR is computed by measuring the pixel difference between the original image and compressed image. Values for PSNR range between infinity for identical images, to 0 for images that have no commonality. PSNR decreases as the compression ratio increases for an image.

SIGNAL- TO –NOISE RATIO

Signal-to-noise ratio is defined as the power ratio between a signal (meaningful information) and the background noise (unwanted signal).

$$SNR\ (dB) = 10\log_{10} (P_{signal}/P_{noise}) \qquad (3)$$

Where 'P' is average power. Both signal and noise power must be measured at the same and equivalent points in a system, and within the same system bandwidth. If the signal and the noise are measured across the same impedance, then the SNR can be obtained by calculating the square of the amplitude ratio.

RESULTS AND DISCUSSIONS

The effectiveness of compressed ratio of medical images. In this work I have taken 10 raw medical images samples of breast cancer. The images are very noisy in nature and they are not clearer. Medical images are raw images and they are compressed using a Conventional method, i.e., Signal-to-noise ratio and peak signal-to-noise ratio**.** So, after that, these images are a little bit smaller than a non-compressed images means original images but still retain all of the information and compressed easily. When we use this conventional method the compressed ratio of images are better. There is no change in retained energy, the compressed ratio, signal-to-noise ratio and peak signal-to-noise ratio are different of each images. The compressed ratio of all images in between 37.4122 to 120.4561, the minimum compressed ratio is 37.4122 of image 4 and the maximum compressed ratio is 120.4561 of image 8 in bold letter as shown in table 1. The table 1 shows that Calculation of CR, PSNR, MSE, SNR on the Digitized Mammogram Image**.** The next is find the SNR of all images the range SNR in between 15.7377 to 22.0441, the minimum SNR value is 15.7377 of image 1, and maximum value is 22.0441 of image 6. Similarly, the PSNR value of all medical images the range in between 96.7563 to 150.3133. The minimum PSNR value is 96.7563 of image 2 and maximum PSNR value is 150.3133 of image 1. But the qualities of all medical images are poorer. The next target is to improve the quality of medical images we use Principal component analysis method; the PCA work is to improve the quality of the image without any change in pixels. Matlab is a good programming toolbox package of version 7.8, provides functional software environment. The main goal of this package is to provide users with a set of integrated tools to create coding and simulate them easily. The original mammogram as shown in figure 1 and figure 2 is a gray image of original mammogram. The Salt & Pepper noise is scattered in whole part of the Image and disturbing it. The figure 3 shows that Salt & Pepper noise. After eliminating the noise in the image and convert it into colour map. The figure 4 shows the colour map image. Plot the intensities of the image as shown in figure 5.compressed the image in four sections 10, 25, 75 and 100 as shown in figure 6. After decompression of all medical images and the image becomes smooth and very clearer. The abnormality in the image is become clearer and the image becomes highest quality. By the use of PCA we find their Mean Square Error and peak signal-to-noise ratio, the MSE is minimum as compared to PSNR value. A higher PSNR normally indicate that the reconstruction is of higher quality. As we compared the PSNR values of conventional decompressed image and PSNR by the PCA method the PSNR values of conventional method is higher than PSNR values by PCA but we compared the Quality-wise the Images is compressed by PCA is of higher quality as compare to conventional method.

Table 1. Calculation of CR, PSNR, MSE, and SNR on the Digitized Mammogram Image

Medical Images	**CR (%)**	**RETAINED ENERGY**	**PSNR value of Decompressed Image**	**SNR**	**PSNR value by PCA method**	**MSE**
Image 1	69.6041	99.999	**150.3133**	**15.7377**	70.2185	0.006183
Image 2	52.7721	99.9997	96.7563	20.1098	62.6612	0.002315
Image 3	60.6671	99.9992	102.1213	18.6456	60.1197	0.063257
Image 4	**37.4122**	99.9992	113.3879	16.3188	80.6228	5.63E-04
Image 5	58.7008	99.9989	101.6903	17.7129	74.6661	0.002221
Image 6	63.3548	99.9993	**114.1775**	**22.0441**	**54.6709**	0.221814
Image 7	61.9099	99.9984	106.943	19.8983	67.8311	0.010714
Image 8	**120.4561**	99.9996	**120.4561**	21.2881	**95.6977**	1.75E-05
Image 9	81.7921	99.999	101.8405	16.8011	91.674	4.42E-05

CONCLUSIONS

This paper improves the quality of Breast cancer image by using Principal component analysis method. The image is clearly visible and also the proposed method is compared with other existing methods in terms of their image parameter SNR (signal-to-noise ratio) MSE (Mean Square Error) & PSNR (Peak Signal to Noise Ratio) of each image, but machine vision cannot adapt itself to the color changes. By the use of PCA method improves the quality and compressed the noise ratio in the image. The Cancer Image becomes high enhanced and we easily detect abnormalities in the images. This will help doctor to take or analyze the abnormal signs in the image and they take quickly action, patient have and according to which he/she can take necessary and appropriate treatment steps.

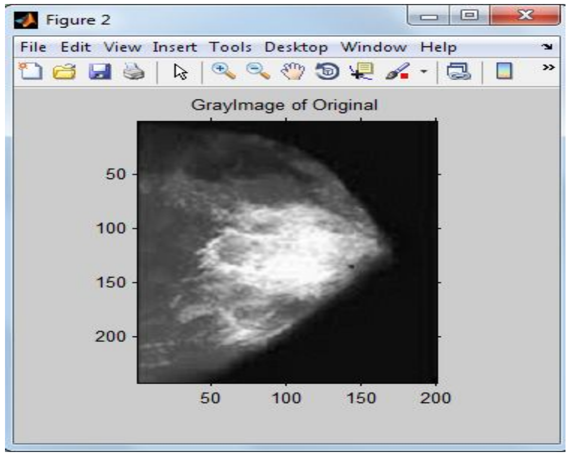

Figure 1 Original Mammogram

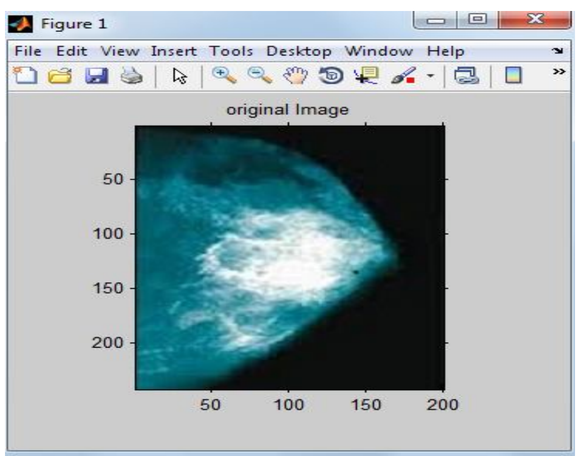

Figure 2 Gray Image of Original

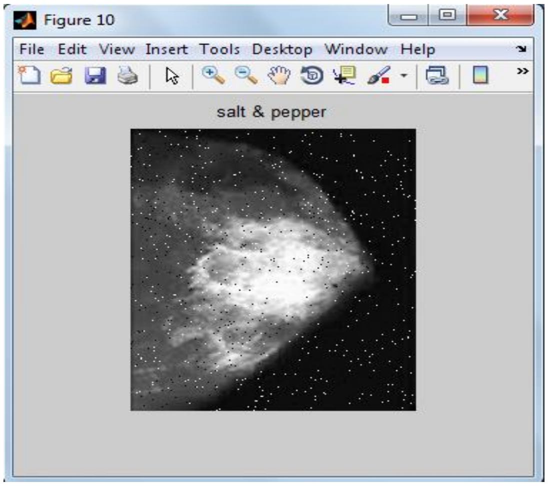

Figure 3 Salt & Pepper Noise

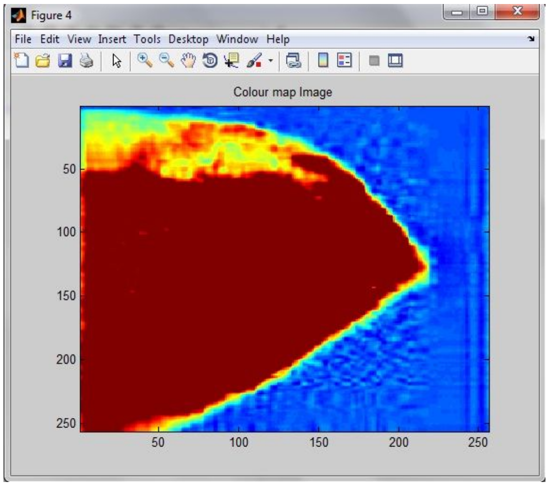

Figure 4 Colour Map Image

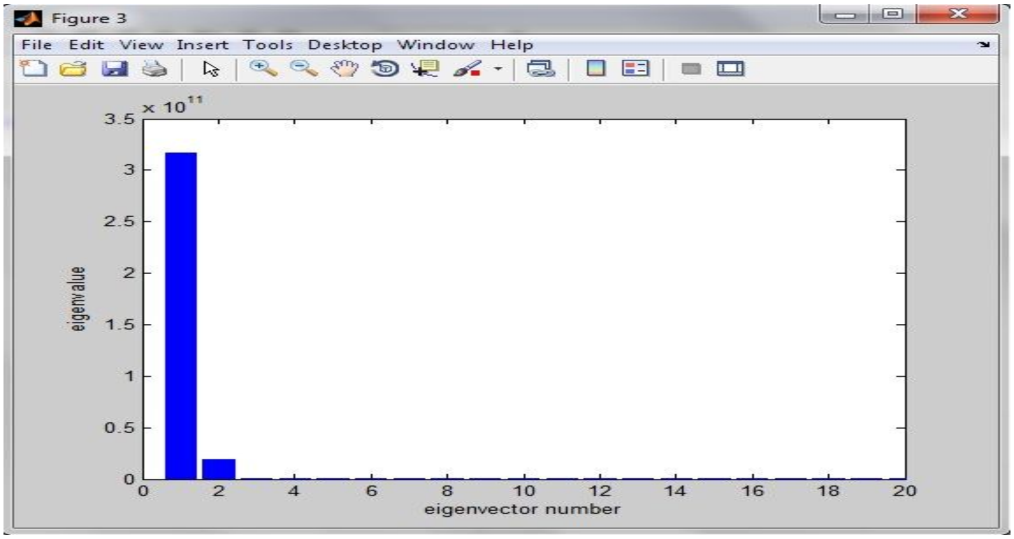

Figure 5 Plot Intensities of Image

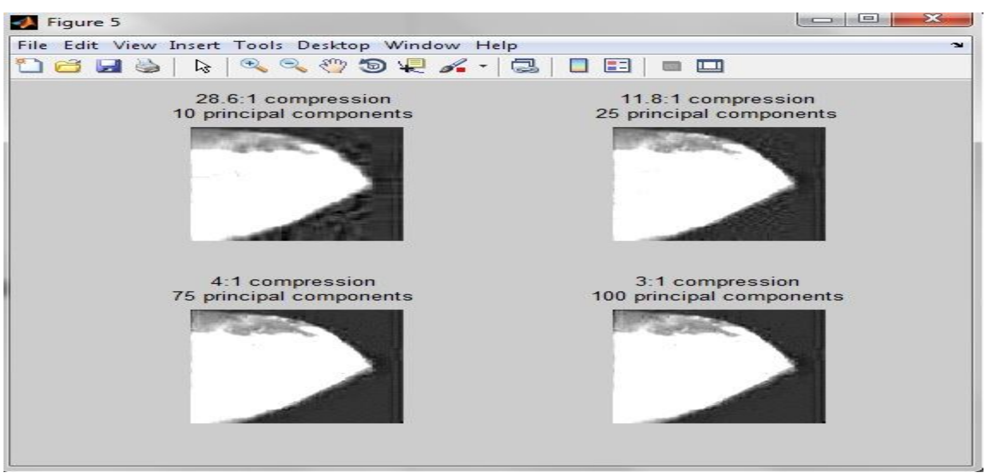

Figure 6 Compressed Image by PCA

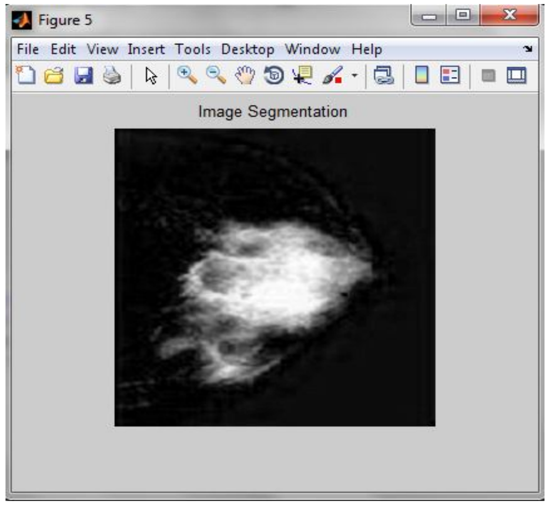

Figure 7 Segmented Abnormal Part in Image

ACKNOWLEDGEMENT

Authors are thankful to UGC NEW DEHI for providing financial assistance to this research paper under UGC major research project. Research grant sanctioned from UGC vide F-NO.37-505/2009(SR).dated 21.12.2009.

REFERENCES

[1]. Youhei Terai, Tomio Goto "Color Image Enhancement " Proceedings of IEEE, pp-392-393, Iligan City, 2009.

[2]. Zhengya Xu, Hong Ren, Xinghuo "Color Image Enhancement by Virtual Histogram Approach" IEEEjournal consumer electronics, vol.-56, NO. 2, may 2010, pp.204-208. Subramanya A, "Image Compression Technique," Potentials IEEE, Vol. 20, Issue 1, pp 19-23, Feb2011.

[3]. David Jeff Jackson & Sidney Joel Hannah, " Comparative Analysis of image Compression Techniques," System Theory 1993, Proceedings SSST '93, 25th Southeastern Symposium,pp 513-517, 7 –9 March 1993.

[4]. Hong Zhang, Xiaofei Zhang & Shun Cao, " Analysis & Evaluation of Some Image Compression Techniques," High Performance Computing in AsiaPacific Region, 2000 Proceedings, 4th Int. Conference, vol. 2, pp 799-803,14-17 May, 2000.

[5]. Maryam Amirmazlaghani, Hamidreza Amindavar, "A Novelwavelet Domain Statistical Approach for Denoising SAR Images" ICIP, 2009.

[6]. Y.MuraliMohanBabu, M.V.Subramanyam & M.N.Giriprasad "Bayesian Denoising of SAR Image"
International Journal of Computer Science & Technology, Vol.2, Issue.1, PP-72-74, 2011,

[7]. C. Boncelet, "Image noise models," in Handbook of Image and Video Processing, A. C. Bovik, Ed.New York: Academic, 2000.

[8]. S. Mallat, "A Wavelet Tour of Signal Processing", 2nd ed. San Diego, CA: Academic, 1999.

[9]. J. K. Romberg, H. Choi, R. G. Baraniuk, "Bayesian tree-structured image modeling using waveletdomainhidden Markov models," IEEETrans. Image Process., Vol. 10, No. 7, pp. 1056–1068, Jul.2001.

[10]. K. John Peter , Dr K.Senthamarai Kannan , Dr S. Arumugan , G.Nagarajan, "Two-stage image denoising by Principal Component Analysis with Self Similarity pixel Strategy", International Journal of Computer Science and Network Security, VOL.11 No.5, May 2011, 296-301.

[11]. Ayan Chakrabarti and Keigo Hirakawa, "Effective separation of sparse and non-sparse image feature for denoising", ICASSP 2008, 857-860.

[12]. P. Cosman, R. Gray, "Evaluating Quality of Compressed Medical Images: SNR, Subjective Rating, and Diagnostic Accurac", in Proc. IEEE, Vol. 82, 1994, pp. 919-932.