Volume 2, Issue 11, November 2012 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Analysis Of Various Quality Metrics for Medical Image Processing

Ravi Kumar
Research scholler
Guru Nanak Dev Engineering College
Ludhiana, Punjab. India

Munish Rattan
Department of Electronics and Communication Engg.
Guru Nanak Dev Engineering College
Ludhiana, Punjab. India

Abstract*—*This paper presents the comparative analysis of various quality metrics for medical image processing. Measurement of image quality is important for many image processing applications. Image quality assessment is closely related to image similarity assessment in which quality is based on the differences (or similarity) between a degraded image and the original, unmodified image. Objective methods have been used for expressing the image quality because these are automatic and mathematically well defined procedure unlike subjective methods which are expansive, time consuming and observer dependent. In this work images have been subjected to various degrees of blur, noise, compression and contrast levels and quality has been measured in terms of well known metrics such as Mean Squared Error (MSE), Structural Similarity Index Metrics (SSIM) Peak Signal-to-Noise Ratio (PSNR), Maximum difference (MD), etc.

Index terms-Image quality analysis, Mean Square Error (MSE), Structural Similarity Index Metric(SSIM), Peak Signal to Noise Ratio(PSNR), Mean Absolute Error(MAE)

I. INTRODUCTION

Digital images are subject to a wide variety of distortions during acquisition, processing, compression, storage, transmission and reproduction, any of which may result in a degradation of visual quality [1]. Identifying the image quality measures that have highest sensitivity to these distortions would help systematic design of coding, communication and imaging systems and of improving or optimizing the picture quality for a desired quality of service at a minimum cost. For image quality measurement there are basically two approaches:-

1) Subjective measurements
2) Objective measurements.

Subjective measurements are the result of human experts providing their opinion of the image quality and objective measurements are performed with mathematical algorithms. For applications in which images are ultimately to be viewed by human beings is the only “correct” method of quantifying visual image quality i.e. through subjective evaluation. In practice, however, subjective evaluation is usually too inconvenient, time-consuming and expensive. The goal of research in objective image quality assessment is to develop quantitative measures that can automatically predict perceived image quality [2].
An objective image quality metric can play a variety of roles in image processing applications. It can be used to dynamically monitor and adjust image quality. It can be used to optimize algorithms and parameter settings of image processing systems. It can be used to benchmark image processing systems and algorithms [1, 3].

Objective image quality metrics can be classified according to the availability of an original (distortion-free) image, with which the distorted image is to be compared. Most existing approaches are known as full-reference, meaning that a complete reference image is assumed to be known. In many practical applications, however, the reference image is not available, and a no-reference or "blind" quality assessment approach is desirable. In a third type of method, the reference image is only partially available, in the form of a set of extracted features made available as side information to help evaluate the quality of the distorted image. This is referred to as reduced-reference quality assessment [4]. This new method focuses on full-reference image quality assessment.

The simplest and most widely used full-reference quality metric is the mean squared error (MSE), computed by averaging the squared intensity differences of distorted and reference image pixels, along with the related quantity of Peak Signal-to-Noise Ratio (PSNR). These are appealing because they are simple to calculate, have clear physical meanings, and are mathematically convenient in the context of optimization. But they are not very well matched to perceived visual quality. MSE and PSNR lack a critical feature: the ability to assess image similarity across distortion types. In the last three decades, a great deal of effort has gone into the development of quality assessment methods that take advantage of known characteristics of the human visual system (HVS) [1, 5]. The majority of the proposed perceptual quality assessment models have followed a strategy of modifying the MSE measure so that errors are penalized in accordance with their visibility.

Among many of objective measures of picture quality, based on computable distortion measures, listed in Table I [6,7].

Where $x(i, j)$ represents the original (reference) image and $y(i, j)$ represents the distorted (modified) image. Two Human visual systems (HVS) based image quality metrics are the universal image quality index (Q) [3] and the Structural Similarity Index (SSIM) [1, 5 , 8].

TABLE 1
EXISTING MEASUE OF QUALITY METRICS

S. No.	TYPE	DESCRIPTION
1.	MSE (Mean Square Error)	$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j)-y(i,j))^2$
2.	PSNR (Peak Signal to Noise Ratio)	$PSNR = 10\log_{10}\frac{(2^n-1)^2}{\sqrt{MSE}}$
3.	AD (Average Difference)	$AD = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j)-y(i,j))$
4.	MD(Maximum Difference)	$MD = MAX\lvert x(i,j)-y(i,j)\rvert$
5.	MAE (Mean Absolute Error)	$MAE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\lvert((i,j)-y(i,j)\rvert$
6.	NK(Normalized Cross-Correlation)	$NK = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j)\times y(i,j))}{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j))^2}$
7.	SC (Structural Content)	$SC = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(y(i,j))^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j))^2}$
8.	IF(Image Fidelity)	$IF = 1-\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j)-y(i,j))^2}{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j))}$
9.	PMSE(Peak Mean Square Error)	$PMSE = \frac{1}{MN}\times\frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(x(i,j)-y(i,j))^2}{(MAX\,(x(i,j))^2}$
10.	SSIM(Structural Similarity Index Metrics)	$SSIM = \frac{(2\times\bar{x}\times\bar{y}+C1)(2\times\sigma_{xy}+C2)}{(\sigma_x^2+\sigma_y^2+C2)\times((\bar{x})^2+(\bar{y})^2+C1)}$

II. RECENT METHODS TO MEASURE THE QUALITY OF IMAGE

Universal image quality metric:

Let $x = \{ x_i \mid i= 1,2,3,..., N\}$, $y= \{ y_i \mid i= 1,2,3,..., N \}$ be the original and the test images, respectively. The proposed quality index is defined as: [2]

$$Q = \frac{4\times\sigma_{xy}\times\bar{x}\times\bar{y}}{(\sigma_x^2+\sigma_y^2)\times((\bar{x})^2+(\bar{y})^2)} \qquad \text{(i)}$$

$\bar{x}$, $\bar{y}$, σ_x^2 , σ_y^2 and σ_{xy} are given as:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N}x_i \qquad \text{(ii)}$$

$$\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i \qquad \text{(iii)}$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 \qquad \text{(iv)}$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y}) \qquad \text{(v)}$$

The dynamic range of Q is [-1, 1], with the best value achieved when *yi* = *xi*, *i* = 1,2,…,*n*. The index is computed for each window, leading to a quality map of the image. The overall quality index is the average of all the Q values in the quality map:

$$Q = \frac{1}{M}\sum_{j=1}^{M} Q_j \qquad \text{(vi)}$$

M = total number of windows.

SSIM (Structural similarity index metric): Q produces unstable results when either $((\bar{x})^2 + (\bar{y})^2)$ or $(\sigma_x^2 + \sigma_y^2)$ is very close to zero. In order to circumvent this problem, the measure has been generalized to the Structural Similarity Index (SSIM): [1]

$$SSIM = \frac{(2\times\bar{x}\times\bar{y} + C1)(2\times\sigma_{xy} + C2)}{(\sigma_x^2 + \sigma_y^2 + C2)\times((\bar{x})^2 + (\bar{y})^2 + C1)} \qquad \text{(vii)}$$

C1 and C2 are constant.

As in the case of Q, the overall image quality MSSIM is obtained by computing the average of SSIM values over all windows:

$$MSSIM = \frac{1}{M}\sum_{j=1}^{M} SSIM_j \qquad \text{(viii)}$$

M = total number of windows.

III. THE PROPOSED APPROACH

The evaluation of image quality is crucial for many image processing systems, such as those for compression, enhancement, transmission and reproduction. A great deal of effort has gone into designing quality assessment methods that take advantage of known characteristics of the HVS.

Natural image signals are highly structured. The most fundamental principle to image quality assessment is that the HVS is highly adapted to extract structural information from the visual scene, and therefore a measurement of structural similarity (or distortion) should provide a good approximation to perceptual image quality. Depending on how structural information and structural distortion are defined, there may be different ways to develop image quality assessment algorithms.

Following steps are taken in proposed approach:

1. Study the existing metrics.
2. Simulate the various quality metrics with MATLAB.
3. Execute the various quality metrics by adding noise, compression, blur and contrast applications with medical images.
4. Evaluate and compare the various quality metrics.

IV. RESULTS & COMPARISON

Comparative evaluation of various quality metrics has been done in MATLAB with three different kinds of images which are commonly used in areas of medical image processing:

1. MRI image (noise removal and compression)
2. X-Ray image (for blur applications)
3. Ultra sound image (contrast application)

MRI image has been chosen for application of noise. These kinds of images have small size details which are susceptible to random noise. Gussian noise has been chosen for adding noise to the image. In Gaussian Noise, power is directly proportional to the variance. So, different values of noise variances are taken for the purpose of adding different level of noise powers to the image. The images corresponding to different noise powers are shown in Figure 1 and values of various quality metrics is shown in Table II.

As can be seen from Table II, all quality metrics acknowledge the changes to the noise powers variation. MD appears to be the most sensitive metric whereas SC & PMSE being the least sensitive to the variation in noise power. From Table it is clear that SSIM give results with greater accuracy.

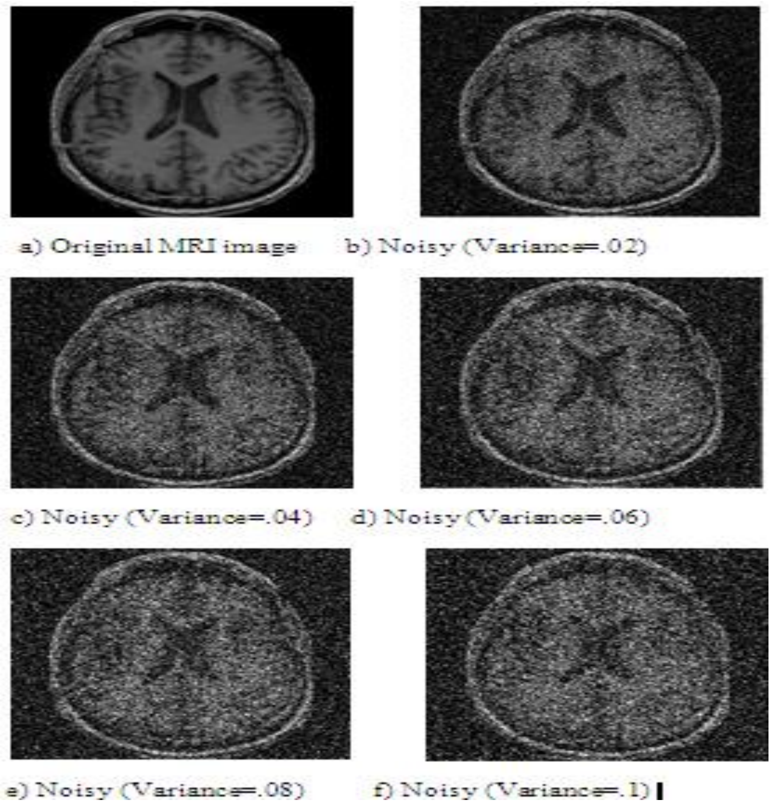

Figure 1: Different MRI denoising images with variance

TABLE II
COMPARISON OF VARIOUS QUALITY METRICS FOR NOISY MRI IMAGE

Variance	**MSE**	**PSNR**	**AD**	**MD**	**MAE**	**NK**	SC	**IF**	PMSE	**SSIM**
0.02	**58.4048**	**39.2986**	**-6.6495**	**153**	**22.4545**	**1.0146**	1.3225	**.7066**	.0011	**.2046**
0.04	**63.9898**	**39.1002**	**-10.0501**	**190**	**30.5875**	**1.0263**	1.5972	**.4554**	.0012	**0.137**
0.06	**65.2490**	**39.0579**	**-13.766**	**239**	**36.6935**	**1.0598**	1.9023	**.2172**	.0013	**.1070**
0.08	**66.5353**	**39.0155**	**-16.6531**	**255**	**41.4679**	**1.0829**	2.1662	**0.0084**	.0013	**.0881**
0.1	**68.3621**	**38.9567**	**-18.5784**	**255**	**45.1500**	**1.0900**	2.3725	**-0.1925**	.0013	**.0749**

Table III shows different values of quality metrics for MRI compression image. The compression depend upon the block processing means image is bigger we can use smaller block of image at a time .To compress the image, the first step is to subdivide the input image into non overlapping pixel blocks of size 8×8. They are subsequently processed left to right, top to bottom. The two-dimensional DCT is computed for each block. The DCT coefficients are then quantized, coded, and transmitted. By applying block processing image signal firstly converted from space domain to frequency domain, then we use mask as a frequency domain filter. By multiplying frequency domain filter to frequency domain signal and we get compression signal which will depend upon the no. of elements in the filter. Table III shows that by decreasing no. of elements are less in filter percentage of compression is changes.

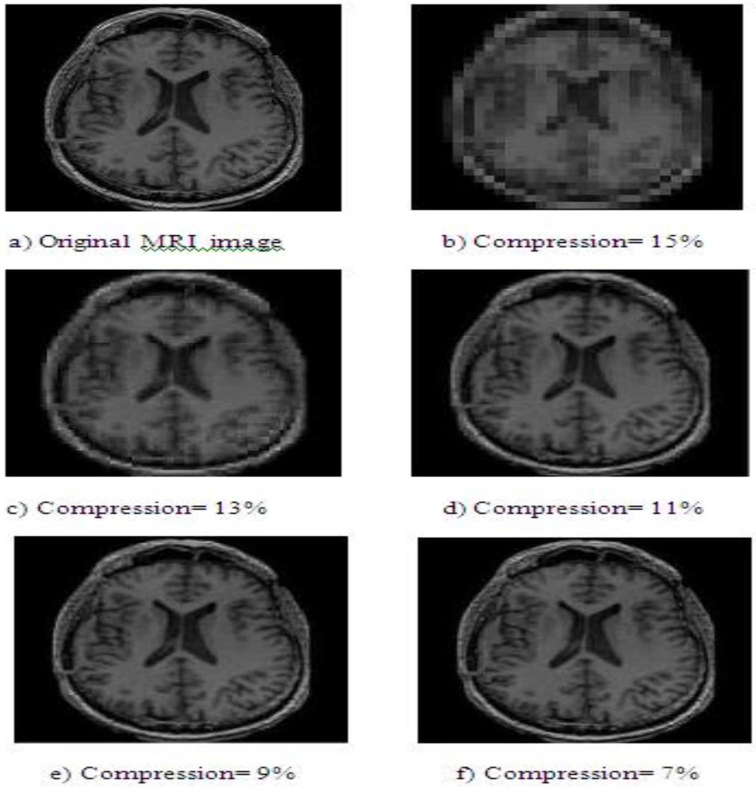

Figure 2:Different MRI images with compression percentage

TABLE III
COMPARISON OF VARIOUS QUALITY METRICS FOR COMPRESSED MRI IMAGE

Compression%	**MSE**	PSNR	AD	**MD**	MAE	NK	SC	IF	**PMSE**	**SSIM**
15	**41.1901**	40.0568	.1255	**162**	10.7485	.8843	.8835	.8852	**7.9236E-04**	**.5804**
13	**29.0606**	40.8143	-.1179	**152**	7.0392	.9443	.9426	.9459	**5.5903E-04**	**.7564**
11	**20.7899**	41.5415	-.0835	**137**	5.0232	.9681	.9667	.9694	**3.9993E-04**	**.8579**
9	**15.5008**	42.1790	-.0326	**114**	3.7552	.9806	.9796	.9817	**2.9818E-04**	**.9184**
7	**11.0426**	42.9154	-.0148	**92**	2.7915	.9875	.9865	.9885	**2.1242E-04**	**.9518**

The images corresponding to different percentage of compression have been shown in Figure 2 and values of various quality metrics are shown in Table III. From the Table III it is clear that most of the quality metrics acknowledge the change in percentage of compression except AD, NK, SC and IF. So MSE, PSNR, MAE, PMSE, MD and SSIM are more significant whereas AD, NK, SC and IF have little significance in image compression applications. From table it is clear that SSIM give results with greater accuracy and high performance.

Blurring operation on an image causes fading of sharp edges in the image. This operation is performed by passing the image through a low pass filter because sharp edge information in an image corresponds to high frequency contents which get suppressed by low pass filter. X-Ray image is commonly used in blur kind of image processing. We have chosen a 2-D linear spatial Gaussian filter with different values of variances for this purpose.

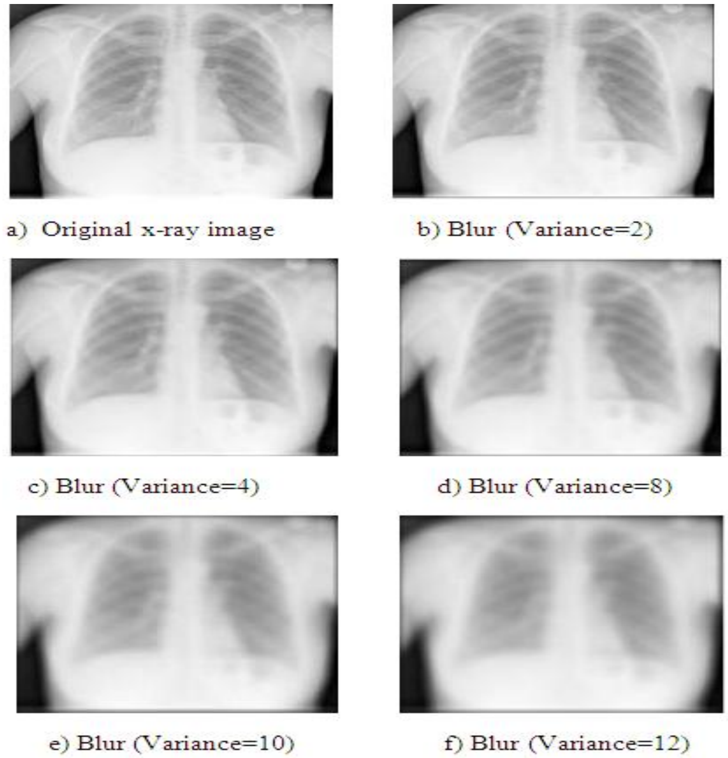

Figure 3: Different x-ray blur images with variance

TABLE IV
COMPARISON OF VARIOUS QUALITY METRICS FOR BLURRED X-RAY IMAGE

Variance	**MSE**	PSNR	**AD**	MD	**MAE**	NK	SC	**IF**	PMSE	**SSIM**
0.02	**58.4048**	39.2986	**-6.6495**	153	**22.4545**	1.0146	1.3225	**.7066**	.0011	**.2046**
0.04	**63.9898**	39.1002	**-10.0501**	190	**30.5875**	1.0263	1.5972	**.4554**	.0012	**0.137**
0.06	**65.2490**	39.0579	**-13.766**	239	**36.6935**	1.0598	1.9023	**.2172**	.0013	**.1070**
0.08	**66.5353**	39.0155	**-16.6531**	255	**41.4679**	1.0829	2.1662	**0.0084**	.0013	**.0881**
0.1	**68.3621**	38.9567	**-18.5784**	255	**45.1500**	1.0900	2.3725	**-0.1925**	.0013	**.0749**

The images corresponding to different degrees of blur have been shown in Figure 3 and values of various quality metrics is shown in Table IV. From the table it is clear that most of the quality metrics acknowledge the change in degree of blur except MD, NK, SC and PMSE. So MSE, PSNR, AD, MAE, IF and SSIM are more significant whereas MD, NK, SC and PMSE have little significance in image blur applications. From Table IV it is clear that SSIM give results with greater consistency.

Contrast of an image shows the distribution of pixel values of an image from minimum to maximum values. A low contrast image has this distribution confined in a small range whereas a high contrast image has pixel values distribution over the full range from minimum to maximum possible pixel value. We have chosen an ultrasound image for variation in its contrast. Contrast of the image has been varied with the help of 'imadjust' function which is inbuilt in MATLAB. It changes the confinement range of pixel of output image as per specification in its arguments.

a) Original ultrasound image

b) Contrast (Stretching limits=0.1-0.9)

c) Contrast (Stretching limits=0.2-0.8)

d) Contrast (Stretching limits=0.3-0.7)

e) Contrast (Stretching limits=0.4-0.6)

Figure 4: Different ultrasound contrast images with stretching limits

TABLE V
COMPARISON OF VARIOUS QUALITY METRICS FOR DIFFERENT CONTRASTS OF ULTRASOUND IMAGE

Stretch limits	**MSE**	**PSNR**	**AD**	**MD**	**MAE**	NK	SC	IF	**PMSE**	**SSIM**
0.1-0.9	**.0029**	**60.8271**	**-31.3847**	**46**	**31.3858**	1.3495	1.8778	.8211	**5.1432E-08**	**.6849**
0.2-0.8	**.2227**	**51.3925**	**-41.5054**	**51**	**41.5541**	1.3108	1.9393	1.0000	**3.7097E-06**	**.6110**
0.3-0.7	**1.9779**	**46.6498**	**-52.1700**	**77**	**52.5986**	1.2891	2.1547	.4235	**3.5213E-05**	**.5225**
0.4-0.6	**10.6019**	**43.0039**	**-62.4974**	**102**	**64.9350**	1.2501	2.4378	1.0000	**1.7662E-04**	**.3716**

From the Table V it is clear that most of the quality metrics acknowledge the change in degree of contrast except NK and IF. So MSE, PSNR, AD, MD, MAE, PMSE and SSIM are more significant whereas, NK, IF and SC have little significance in ultrasound image contrast applications. As can be seen from the table SSIM has high accuracy of results and grater values of SSIM indicate greater image similarity.

V. CONCLUSION

In this work comparative analysis of various quality metrics has been presented. Three different kinds of medical application images have been processed for different levels of noise powers, compression, blur and contrast. All image quality metrics described in this paper have been computed for all four types of image processing operations. Output results have been analyzed for comparison of various quality metrics. These result shows that only a subset of quality metrics is suited for a particular type of image processing operation except the SSIM metric which is capable of expressing image quality irrespective of the type of operation. Other quality metrics are simple but not the accurate for all kinds of image processing operations. SSIM has high performance, works accurately and provides a good approximation of quality measurement however with relatively larger computational time.

REFERENCES

[1] Z.Wang, A.C Bovik, H.R Sheikh, and E.P Simoncelli , "*Image Quality Assessment: From Error Visibility to Structural Similarity,*"IEEE Transactions of Image Processing, vol. 13, pp. 1-12, April 2004.

[2] Zhou Wang and Alan C. Bovik, "*A Universal Image Quality Index*," IEEE Signal Processing Letters,vol. 9, no. 3, pp. 81-84, Mar. 2002.

[3] Eric Silva, Karen A. Panetta, Sos S. Agaian, *"Quantify similarity with measurement of enhancement by entropy*," Proceedings: Mobile Multimedia/Image Processing for Security Applications, SPIE Security Symposium 2007, Vol. 6579, April 2007.

[4] Z. Wang, A. C. Bovik, and L. Lu, "*Why is image quality assessment so difficult*," in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, vol. 4, Orlando, FL, May 2002, pp. 3313–3316.

[5] Z. Wang, H. R. Sheikh, and A. C. Bovik, "*Objective video quality assessment*," in The Handbook of Video Databases: Design and Applications, B. Furht and O. Marques, Eds. Boca Raton, FL: CRC Press,2003.

[6] Alan C. Brooks, Xiaonan Zhao, and Thrasyvoulos N. Pappas, "*Structural Similarity Quality Metrics in a Coding Context: Exploring the Space of Realistic Distortions*", IEEE Transactions on image processing, Vol. 17, No. 8, August 2008.

[7] A. M. Eskicioglu and P. S. Fisher, "*Image quality measures and their performance*,"IEEE Trans. Commun., vol. 43, pp. 2959–2965, Dec. 1995.

[8] E. Peli, "Contrast in complex images," J. Opt. Soc. Amer. A, vol. 7, pp. 2032–2040, 1990.

[9] Sonja Grgic , Mislav Grgic,and Marta Mrak, "*Reliability Of Objective Picture Quality Measures*", Journal of Electrical Engineering, Vol. 55, No. 1-2,3rd October 2004

[10] Gonzalez and Woods, "*Digital Image Processing using MATLAB*", Prentice Hall, New Delhi, India, Year 2006.

[11] N. Ramyashree, P. Pavithra, T. V Shruthi, and Dr.Jharna Maunder, "*Enhancement of Aerial and Medical Image using Multi resolution pyramid*,"IEEE Int. Conf. vol.1 ,August 2010.

[12] M.N.Nobi and M.A. Yusuf, "*A new method to remove noise in magnetic resonance and ultrasound images*," J.Sci.Res.3 (1), pp.81-89, 2011.

[13] Basant Kumar, S.P Sing, Anand Mohan ,and Animesh Anand, " *Performance of Quality Metrics for Compressed medical Images Through Mean Score Prediction*," J.Med.Imaging Health Inf. Vol.2,2012.

