IET-UK International Conference on Information and Communication Technology in Electrical Sciences (ICTES 2007), Dr. M.G.R. University, Chennai, Tamil Nadu, India. Dec. 20-22, 2007. pp. 737-742.

IMAGE QUALITY ASSESSMENT IN RETINAL IMAGE COMPRESSION SYSTEMS

S. R. Nirmala[1], S. Dandapat[2] and P. K. Bora[3]

[1] ECE Department, Indian Institute of Technology Guwahati, India , nirmala@iitg.ernet.in

Keywords: Image quality, retinal image, diagnostic features, LMSE, image compression.

Abstract

This paper presents a simple method to assess the quality of compressed retinal images. It is based on the idea that the retinal image should possess some common features which helps to define the quality of the image. Retinal blood vessels are important diagnostic features in ophthalmological images. Changes in retinal vessel structure helps in detecting the Diabetic Retinopathy(DR). The 2-D discrete laplacian operator is used to detect the vessel structure. The laplacian error map shows spatial error distribution across an image. The overall image quality is given by laplacian mean square error (LMSE). The experimental result shows that laplacian error map localizes the error in a better way compared to structural similarity index (SSIM) map. Similarly, the objective error measure LMSE performs well in discriminating the images with different quality. The image quality is also evaluated using mean square error (MSE),peak signal to noise ratio (PSNR) and mean structural similarity index (MSSIM).

1 Introduction

The use of digital imaging techniques is very common in highly specialized hospitals. Good quality retinal images can be used in ophthalmology to diagnose eye diseases. Ophthalmologists produce more digital images, which requires large memory size for storage and take more time for transmission in telemedicine system. It would be advantageous to compress digital retinal images to reduce storage size and transmission time. There are two types of compression methods described as lossless and lossy. Lossless or reversible compression techniques (Huffman, Lempel-Ziv or Arithmetic coding) are data preserving; that is, the decompressed image is exactly the same as the original image. However, it achieves a relatively low compression ratio. Conversely, lossy or irreversible compression of an image can significantly reduce the storage requirements and transmission time, but the decompressed image may not be same as the original. The choice of a suitable compression level is crucial in the case of medical images [1]. Higher compression may distort the essential diagnostic features in medical images. This may degrade the accuracy of detection and classification of clinical features in computer assisted diagnosis. Hence the assessment of quality of reconstructed images in lossy medical image compression methods becomes indispensable.

Automatic image quality assessment (IQA) is a major issue in retinal image compression system for evaluating the lossy compression methods employed in image transmission [2]. A much less frequently studied issue is having to decide whether a particular image is suitable for diagnosis purposes. This is especially important in telemedicine application, where a remote operator is in-charge of receiving the images,which will eventually be analyzed by an ophthalmologist. The operator may not be familiar with the criteria of making the assessment of images and hence some form of automated support should be provided.

Generally, the quality of reconstructed images in lossy compression method is evaluated using subjective and objective measures. The most frequently used subjective measure is the Mean Opinion Score (MOS), obtained from a group of human subjects. In this test, original and reconstructed images are displayed before the experts to evaluate the preservation of diagnostic features in the reconstructed image [3]. Subjective measure is more appropriate when human observer is the ultimate receiver. The main merit of this method is the true reflection of the perceived image quality. But the test has the drawbacks of being inconvenient, time consuming, expensive and observer responses can vary significantly.

These problems have resulted in an extensive research to develop objective measures for image quality assessment. Various objective measures have been proposed to quantify the distortion. The most widely used objective measures are mean square error (MSE) and peak signal to noise ratio (PSNR). The objective measures exploit the difference between corresponding pixel values in original and reconstructed images, which is squared and the average is taken over all pixels in the image. These measures are mathematically simple to compute and independent of viewing conditions but they do not correlate well with the viewer's response. Hence same MSE does not mean that two images are of same subjective quality. Because of the above mentioned shortcomings of objective measures, many other quality measures that attempt to incorporate properties of the human visual system (HVS) such as sensitivity to edges, local smoothness, have been developed [4]-[7]. A useful property of HVS based quality measure would be the ability at better prediction of subjective quality. But in practice, it is difficult to understand the HVS system completely and it introduces increased complexity in the measurement. A new, recently developed image quality measure called structural similarity index (SSIM) works on the principle that HVS is sensitive to structural information in an image [8],[9]. SSIM considers the luminance , contrast and structural content of the image and then combines all the three to give a quality index.

Automated image quality assessment has the potential to assist in the early detection of eye diseases. Retinal image is used to examine the retina of the eye. The main features of fundus retinal image are optic disc, fovea and blood vessels. These features play important role in the diagnosis of retinal diseases. The changes in blood vessel pattern helps in the early detection of Diabetic Retinopathy (DR). Information about the structure of blood vessels may assist the expert in grading the severity of disease. The blood vessels have greater contrast, lower intensity values and appear dark in the retinal image. This variation in the intensity levels near the vessel border can be captured by gradient based operators. In this direction a 2-D discrete laplacian operator is used to detect the blood vessels as it captures information relating to edge features. Edge information is known to be an image property to which the human visual system is highly sensitive [7]. Therefore laplacian based error measure is more suitable to decide the quality of the retinal images.

The rest of the paper is organized as follows: Section 2 briefly reviews the most frequently used image quality measures. In Section 3, the performance of various quality measures for evaluating the quality of retinal images is described. The analysis of experimental results and discussion is given in Section 4 and conclusions drawn are briefed in Section 5.

2 Review of Image quality measures

In this Section, various image quality measures which are used for the evaluation of compressed image are discussed.

2.1 Subjective measures

Subjective evaluation is the best method since human experts are ultimately the judges of image quality. In this method, the image quality for medical use is evaluated by visual inspection of the diagnostic features. The MOS are usually obtained by subjectively rating the presented images. The rating would be with respect to a parameter like the degree to which diagnostic features are preserved in the reconstructed image compared to original image. Subjects express their judgement of diagnostic feature qualities according to a given MOS scale as shown in Table 1.

Finally, the scores are averaged across subjects to obtain the final MOS. This measure is correlated with diagnostic information in the image and is preferable to use for the evaluation of objective measure. A drawback of assessment of medical image quality by perceptual measures is that it requires the detailed, time-consuming and expensive efforts of human observers, typically highly trained experts. Another factor is , the change in the intensity value of pixels due to compression/decompression process is not considered in the evaluation process, which plays an important role in automatic feature detection process. The difficulties with subjective measures led to the development of objective quality measures.

MOS	Quality group	Description
5	very good	All the diagnostic features are preserved
4	good	Significant features are preserved
3	fair	Some of the features are distorted
2	poor	Most of the features are distorted
1	bad	All the features are distorted

Table 1: Mean Opinion Score (MOS): quality group and its description

2.2 Objective Quality measures

Various objective image quality measures are used for evaluating the level of image degradation in the compression system. There are basically two classes of objective quality measures. The first class includes simple mathematical measures which evaluates the error in the reconstructed signal. The second class of quality measures consider and incorporate human visual system (HVS) properties for quality evaluation. The perceptual or HVS based measures can provide more consistent estimation of image quality than objective measures. But HVS system is highly nonlinear and complex to understand and implement.

Mathematical measures: The compression and decompression process produces artifacts in the images. The amount and visibility of these distortions strongly depend on the actual image content. In block based compression methods, the blocking artifacts are very common. In wavelet-based compression, the transform is applied to the entire image, therefore no block-related artifacts occur. Instead, blurring and ringing are the most prominent distortions. Various numerical objective measures are used to quantify these distortions.

A number of mathematically defined measures have been used in the literature which includes mean squared error (MSE), peak signal to noise ratio (PSNR), root mean squared error (RMSE), normalized absolute error (NAE), laplacian mean squared error (LSME), normalized cross correlation (NCC) and structural content (SC) [10]. The Table 2 lists the various objective measures with the computation expressions. Traditional simple error measures such as the mean squared error (MSE) or the peak signal-to-noise ratio (PSNR) operate solely on a pixel-by-pixel basis and neglect the important influence of image content and viewing conditions on the actual visibility of artifacts. Hence these measurements are inadequate for the evaluation of the compression artifacts and their predictions often do not agree well with perceived quality. Distortions are often much more disturbing in relatively smooth areas of an image than in texture regions with a lot of activity, an effect not taken into account by pixel-based metrics. Therefore the perceived quality of images with the same PSNR can actually be very different. It is observed that a higher PSNR or equivalently, a lower MSE does not necessarily imply a higher subjective image quality. They typically provide a single number for the entire image and thus cannot reflect spatial variations in image quality.These problems have forced the research group for serious study of vision models and HVS based visual quality measures.

HVS based quality measures: Image quality assessment can be improved by incorporating some models of HVS into the evaluation process [4]-[7]. In recent years approaches based on HVS-models are slowly replacing the classical numerical or mathematical quality measures. The quality improvement that can be achieved using an HVS-based approach is significant and applies to a large variety of image processing applications. However, the human visual system is extremely complex and many of its properties are not well understood to model it reliably. A recently developed new image quality measure, Structural SIMilarity index (SSIM) does not directly employ HVS model, but is inspired by the functioning of HVS. SSIM is a change in the fundamental assumption from past intensity error based image quality measures and considered to be the best state of the art quality measure [9]. Previous approaches measure perceptual image quality assuming that image intensity is the key component of visual quality. This method often measure intensity error and then penalize these errors according to their visibility. The main idea here is that human visual perception is built to understand a scene based on its structure suggesting that this structural information is the key component of visual quality. The SSIM is applied locally by computing and comparing the luminance, contrast and structural content within a local window which moves across the entire image. The luminance comparison is a function of mean intensity values μ and contrast is compared by standard deviation σ. The structure comparison is a function of correlation coefficient $\sigma(\mathrm{x},\mathrm{y})$ between x and y. This SSIM map gives the local error variation and these local values are averaged to give a single quality index MSSIM [9].

---

mean square error,

$$\mathrm{MSE} = \frac{1}{\mathrm{MN}} \sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} (\mathrm{X}(m,n) - \mathrm{Y}(m,n))^2$$

---

peak signal to noise ratio,

$$\mathrm{PSNR} = 10 \log_{10} \frac{(2^b-1)^2}{\frac{1}{\mathrm{MN}} \sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} (\mathrm{X}(m,n) - \mathrm{Y}(m,n))^2}$$

where b is the number of bits/pixel.

---

root MSE,

$$\mathrm{RMSE} = \sqrt{\frac{1}{\mathrm{MN}} \sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} (\mathrm{X}(m,n) - \mathrm{Y}(m,n))^2}$$

---

normalized absolute error,

$$\mathrm{NAE} = \frac{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} (|\mathrm{X}(m,n)-\mathrm{Y}(m,n)|)}{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} |\mathrm{X}(m,n)|}$$

---

laplacian mean square error,

$$\mathrm{LMSE} = \frac{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} [O(m,n))-O(Y(m,n))]^2}{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} [O(X(m,n))]^2}$$

---

normalized cross correlation,

$$\mathrm{NCC} = \frac{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} \mathrm{X}(m,n)\mathrm{Y}(m,n)}{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} \mathrm{X}(m,n)^2}$$

---

structural content,

$$\mathrm{SC} = \frac{\sum_{m=1}^{\mathrm{M}} \sum_{n=1}^{\mathrm{N}} \mathrm{X}(m,n)^2}{\sum_{m=1}^{\mathrm{N}} \sum_{n=1}^{\mathrm{N}} \mathrm{Y}(m,n)^2}$$

---

structural similarity index,

$$\mathrm{SSIM}(\mathrm{x},\mathrm{y}) = \frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)}$$

Note:- where X and Y are the original and reconstructed image respectively of size M $\times$ N.
$O((\mathrm{X}(m,n)) = \mathrm{X}(m+1,n) + \mathrm{X}(m-1,n) + \mathrm{X}(m,n+1) + \mathrm{X}(m,n-1) - 4\mathrm{X}(m,n)$

Table 2: Image Quality Measures

3 Image Quality assessment

The retinal image should possess some common features which helps to define the quality of the image. The diagnostic features of retinal image are, optic disc (OD), the central vision part macula (MC) and the blood vessel structure. Blood vessel detection is a critical topic in automatic retinal image processing. Blood vessel morphology is an important indicator of many diseases such as diabetes and hypertension. Abnormality of blood vessel includes: change in color, change in width, change in tortuosity and neovascular generation. Morphological changes in retinal vessel structure helps in detecting and grading the Diabetic Retinopathy(DR). The measurement of changes may then be applied to a variety of clinical studies:screening, diagnosis and evaluation of treatment. To identify blood vessel abnormality in large scale screening, its essential to detect the blood vessel map fast and accurately. Then different measurement can then be done to help doctors in making diagnosis. Two strategies have been generally employed for the detection of blood vessels in retinal image. One is edge detection, the other is vessel tracking which needs a priori knowledge of the beginning position in the image. The former method is used in this paper as it is simple to apply and analyze. In this work, the two-dimensional laplacian operator based edge detector expressed by (1), which plays important role in image processing applications is used for detecting the blood vessels [11],[12].

$$\begin{aligned} O((\mathrm{X}(m,n)) = {} & \mathrm{X}(m+1,n) + \mathrm{X}(m-1,n) + \\ & \mathrm{X}(m,n+1) + \mathrm{X}(m,n-1) - 4\mathrm{X}(m,n) \end{aligned} \tag{1}$$

The first stage of edge detection is the evaluation of derivatives of the image intensity. Gaussian smoothing filters are used to make differentiation more immune to noise.When a two-dimensional laplacian is applied to the reconstructed image, it detects blood vessels present in the image. The Fig. 3 shows the original image, reconstructed, laplacian of the original and laplacian of the reconstructed images. The laplacian MSE (LMSE) is calculated using the corresponding pixel values of the laplacian of original and reconstructed images. Since edge information is an image property to which the HVS is highly sensitive, use of LMSE has been found to be successful in our experiments as it captures information relating to blood vessel

features.
The distortion can be computed locally for every pixel, yielding perceptual distortion maps for better visualization of the spatial distribution of distortions.The pixelwise error between the laplacian of original image and laplacian of reconstructed image gives the laplacian error map as demonstrated in Fig. 2. Such a distortion map can help the expert to make proper decision on diagnostic quality of the image. Hence this can be more useful and more reliable than a global measure in quality assessment applications. The image quality is also evaluated by computing traditional MSE, PSNR and MSSIM. The performance of LMSE is compared with MSE, PSNR and MSSIM quality measures.

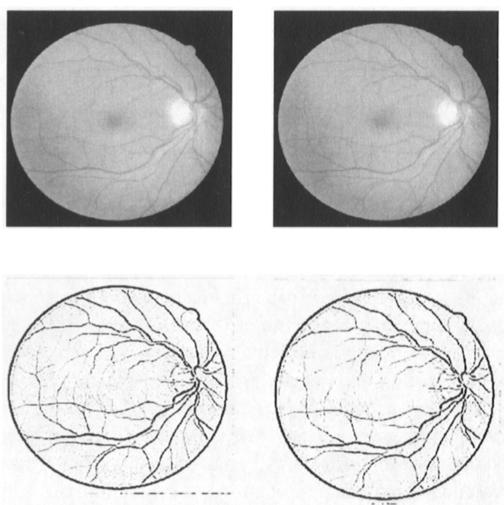


Figure 1: (a) Original image (b) Reconstructed image (c) Laplacian of original image (d)Laplacian of reconstructed image

Images used in the experiment: The retinal images from DRIVE (Digital Retinal Images for Vascular Extraction) database are used for testing and making observations. DRIVE database consists of 40 images (seven are pathological images) captured from a Canon CR5 non-mydriatic 3CCD camera at 45 degree field of view. The set of 40 images has been divided into a training and a test set, both containing 20 images. The images are of size 565 x 585 pixels, 8 bits per color channel. In the present work, the 20 images which includes both test and training images are used. The images are resized to 256 x 256 and stored in JPEG format.

4 Results and Discussion

To produce test images for objective and subjective image quality assessments, many different image compression methods like JPEG2000, Embedded Zerotree Wavelet (EZW) and Set Partitioning in Hierarchical Trees (SPIHT) [13] can be used. SPIHT coding algorithm is a very efficient technique for wavelet image compression. SPIHT is improved and extended version of Embedded Zerotree Wavelet (EZW) coding algorithm and it is one of the best wavelet coder today. Hence only the digital retinal images compressed using SPIHT are considered in this work. The numerical objective quality measures MSE, PSNR, LMSE and MSSIM-a measure inspired by HVS, are computed to quantify the distortion in the reconstructed image after decompression. The spatially varying Laplacian error generates a distortion map of diagnostic vessel features. Similarly the SSIM map is also generated. The LMSE error map generation is simple as it dose not require the computation of mean ,variance and standard deviation as done in SSIM map generation. The laplacian error map and the SSIM map are compared and studied. To have meaningful comparison, the laplacian images are inverted so that brighter region indicates lower distortion and darker region indicates higher distortion. Laplacian error map shows good error visualization of blood vessels compared to SSIM map as shown in Fig. 3. The LMSE for different image quality, is computed as the mean square error between the laplacian of the original and reconstructed images.
While conducting the experiments it is observed that distortion in the smooth region resulted in a lower value of LMSE but when the regions with vessel starts getting distorted, the LMSE value increases greatly. This shows that the LMSE reflects distortion in the edges (blood vessels) more clearly than distortion in the smooth region. To illustrate this, the laplacian of 04-test image is taken and more number of pixels in the nonvessel area are altered manually. This resulted in a low LMSE value. Similarly when few vessel pixels are altered, LMSE value is very high compared to the previous case. These effects are shown in Fig. 4. In the above example, approximately 40,000 nonvessel pixels resulted in an LMSE

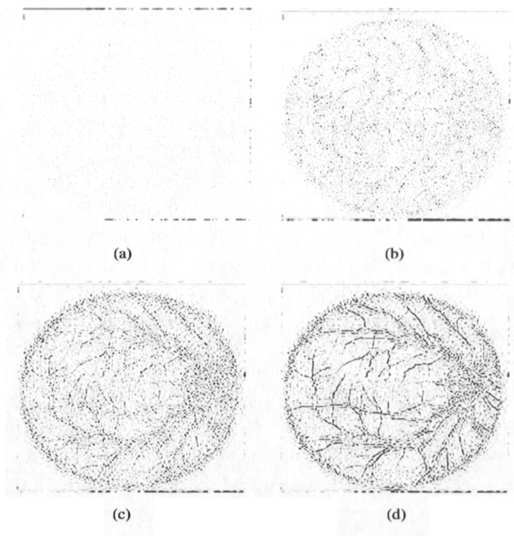


Figure 2: Laplacian error map for (a) CR=4 (b) CR=8 (c) CR=12 (d) CR=16

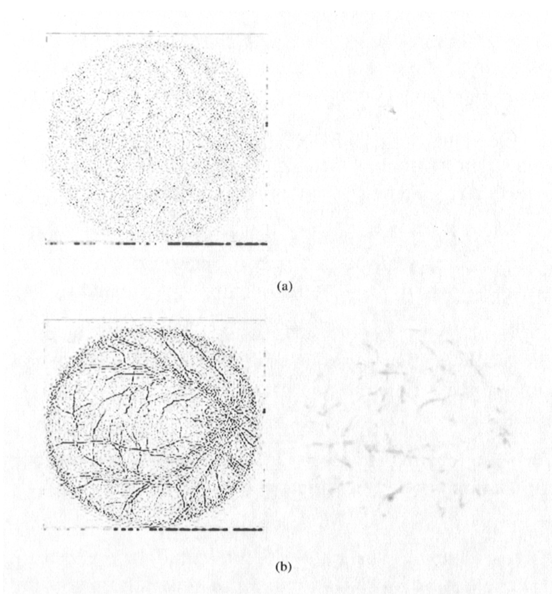


Figure 3: Laplacian error map and SSIM map for (a) CR=12 (b) CR=16

of 3 and around 1500 altered vessel pixels gives an LMSE of 20. Hence LMSE can be used as a good quality measure for retinal images.

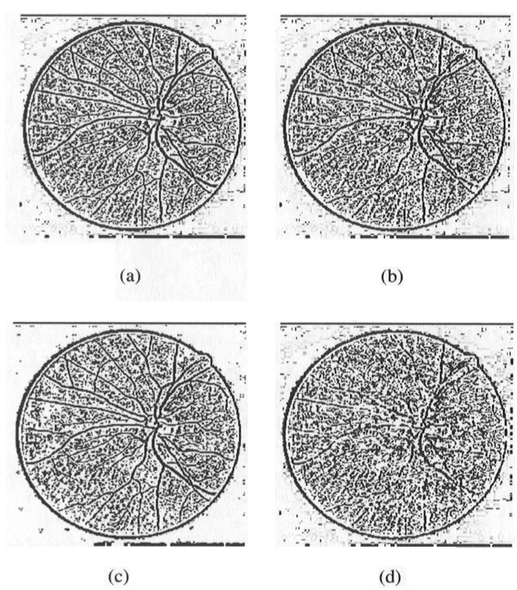
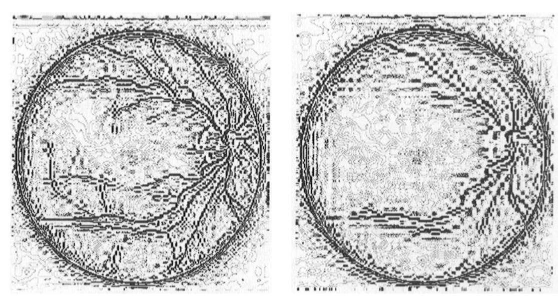


Figure 4: (a) Laplacian of original image (b) Nonvessel pixels are distorted (c) Some vessel pixels are distorted (d) more vessel pixels are distorted

In some applications like telemedicine , it would be helpful to know, the given test image is useful for diagnosis or not before applying the quality assessment procedure. This is possible, if we get an overall qualitative view about the degree of the distortion in the reconstructed image. Laplacian produces closed edge contours, if the image meets certain smoothness constraints. At low compression rate, the smoothed regions in the image is not strong enough to show the closed loop like formations. It is observed that, a properly scaled laplacian image exhibits small closed loop like formations in the blurred (over smoothed) regions due to wavelet compression at moderate and high compression ratio. This indicates laplacian localizes the distorted smooth regions by showing loop like formations as shown in Fig. 5. Hence laplacian of the reconstructed image gives an overall idea of whether the reconstructed image is distorted or not without requiring for a reference image.



(a) (b)

Figure 5: formation of closed contours for (a) CR=16 (b) CR=32

While conducting the subjective evaluation of images by medical experts, no constraints were placed on viewing time, viewing distance or illumination (lighting) conditions. The experts were allowed to simulate the conditions they would use in their everyday work. It is focused more on the experts assessment of diagnostically acceptable quality of compressed images to ensure sufficient diagnostic accuracy. The MOS is obtained from two medical experts and eight students. For each of the image, the MOS is obtained by taking the average of the individual MOS of all the viewers. The closeness of MSE, PSNR, LMSE and SSIM measures with the medical experts opinion is computed using Rank Order Correlation Coefficient (ROCC) and tabulated as shown in Table 3.

5 conclusion

Preservation of diagnostic information in the reconstructed retinal image is the basic requirement in telemedicine. Twenty retinal images from DRIVE database are compressed by using SPIHT, a Wavelet based compression method. The quality of retinal images is evaluated by various objective quality measures. The quality opinion scores of the same images are

Quality measure	*ROCC*
MSE	0.9679
PSNR	- 0.928
LMSE	0.9776
SSIM	- 0.933

Table 3: Performance Comparison of Image Quality Measures; ROCC:Rank Order Correlation Coefficient.

obtained by a group of subjects. After comparing the results, it has been found that LMSE is better in giving the information about the blood vessel structure present in the reconstructed image than other measures. LMSE has good correlation of 0.9776 with the subjective score compared to other quality measures. Detection of blood vessels is an important stage in automatic retinal image processing. Abnormality of blood vessel indicates many eye related diseases. To identify blood vessel abnormality in large scale screening, it is essential to detect the blood vessel map fast and accurately. Different measurements can then be done to help doctors in making diagnosis. Hence the present work of laplacian based vessel detection can be considered as an initial step towards automatic image quality assessment.

References

References

[1] P. Cosman , R M Gray and R A Olshen, "Evaluating the Quality of Compressed Medical Images: SNR, Subjective Rating and Diagnostic accuracy" Proc. IEEE, Vol.82 No.6,pp 919-932,(1994).

[2] R. H. Eikelboom, K. Yogesan, Chris J. Barry, Ian J. Constable, Mei-Ling Tay-kearney, L. Jitskaia and Philip H. House, " Methods and Limits of Digital Image COmpression of Retinal Images for Telemedicine", Investigative Ophthalmology and Visual Science, Vol.41, No.7, pp 1916-1924(2000).

[3] ITU-R Recommendation BT.500-10, "Methodology for the subjective assessment of the quality of the television pictures". 2000.

[4] S. Daly, "The Visible Difference Predictor: An algorithm for the assessment of image fidelity". Digital Image and Human Vision, MIT Press ,1993.

[5] M. Miyahara, K. Kotani and V. R. Algazi,"Objective Picture Quality Scale (PQS) for Image Coding", IEEE Trans. Commun., Vol.26, No.9, pp 1215-1226(1998).

[6] Michael P. Eckert and Andrew P. Bradley, " Perceptual quality metrics applied to still image compression", Signal Processing, Elsevier, Vol.70, No. 3, pp 177-200(1998).

[7] Andrew P. Bradley, " Wavelet Visible Difference Predictor", IEEE Trans. on Image Processing, Vol.8, No. 5, pp 717-730(1999).

[8] Z.Wang and A.C. Bovik, "A Universal Image Quality Index". IEEE Signal Processing Letters, Vol.9, No.3, pp 81-84(2002).

[9] Z.Wang, A.C. Bovik, H. Sheikh and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity". IEEE Trans. Image Processing, Vol.13, No.4, pp 600-612(2004).

[10] Sonja Grgic, Mislav Grgic and Marta Mrak, "Reliability of Objective Picture Quality measures", Journal of Elecrical Engg., Vol. 55, No. 1-2, pp 3-10(2004).

[11] Mitra Basu, "Gaussian-Based Edge Detection Methods- A survey", IEEE Trans. Systems, Man and Cybernetics- PartC:Application and Reviews, Vol. 32, No. 3,pp 252- 260(2002).

[12] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Pearson Education.

[13] A. Said and W.A. Pearlman, "A New Fast and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees", IEEE Trans. on Circuits and Systems for Video Technology, vol. 6, No.3, pp. 243-250(1996).