World Academy of Science, Engineering and Technology
International Journal of Electrical and Computer Engineering
Vol:3, No:2, 2009

No-Reference Image Quality Assessment using Blur and Noise

Min Goo Choi, Jung Hoon Jung, and Jae Wook Jeon

***Abstract*—**Assessment for image quality traditionally needs its original image as a reference. The conventional method for assessment like Mean Square Error (MSE) or Peak Signal to Noise Ratio (PSNR) is invalid when there is no reference. In this paper, we present a new No-Reference (NR) assessment of image quality using blur and noise. The recent camera applications provide high quality images by help of digital Image Signal Processor (ISP). Since the images taken by the high performance of digital camera have few blocking and ringing artifacts, we only focus on the blur and noise for predicting the objective image quality. The experimental results show that the proposed assessment method gives high correlation with subjective Difference Mean Opinion Score (DMOS). Furthermore, the proposed method provides very low computational load in spatial domain and similar extraction of characteristics to human perceptional assessment.

***Keywords*—**No Reference, Image Quality Assessment, blur, noise.

I. INTRODUCTION

IN recent years, digital camera is equipped in most of the mobile products like cellular phone, PDA and notebook computer. Image quality is the most important criteria to choose mobile products. In some cases, the benchmarks or reviews of products are based on subjective image quality test and thus are dependent on tester and environment. The subjective image quality assessment often misleads the decision for the image quality control parameters of Image Signal Processing (ISP) algorithm.

The simple and widely used objective image quality metrics are Mean Square Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). But both of them are known not to be well correlated with human perceptual visual quality [1] and need the original reference image. However, it is not always possible to get the reference images to assess image quality. Human observers can easily recognize the distortion and degradation of image without referring to the original image. Therefore, there is absolutely necessary to develop objective quality assessment that correlates well with human perception without the reference image (No-Reference).

In this paper, we propose a method for image quality assessment based on ratio and mean factors of edge blurriness and noise. The proposed quality assessment obtains excellent correlation with subjective image quality scores. There is high correlation between image quality factors and subjective quality scores.

The rest of the paper is organized as follows:

Section II discusses the related work showing the reason why objective image quality assessment is important and necessary. In section III, the new feature extraction algorithm is proposed. We present experimental results and correlation with subjective image quality assessment in section IV. Finally, Section V draws conclusions and provides future works.

Min Goo Choi is with the Department of Image R&D, Samsung electronics, San #24 Nongseo-Dong, Giheung-Gu, Yongin City, Kyungki-Do, 446-711, Korea (corresponding author. phone: 82-31-209-2882; fax: 82-31-209-5324; e-mail: imkchoi@ samsung.com).

Jung Hoon Jung is with the Department of Image R&D, Samsung electronics, San #24 Nongseo-Dong, Giheung-Gu, Yongin City, Kyungki-Do, 446-711, Korea (e-mail: junghoon.jung@samsung.com).

Jae Wook Jeon is with the school of Electrical and Computer Engineering, University of Sungkyunkwan, 300 Chun chun-Dong, Jangan-Gu, Suwon City, Kyungki-Do, 440-746, Korea (e-mail: jwjeon@yurim.skku.ac.kr).

II. RELATED WORK

The test plan for subjective video quality assessment is well guided by Video Quality Experts Group (VQEG) including the test procedure and subjective data analysis [3]. One of test metrics is Double Stimulus Continuous Quality Scale method (DSCQS). In the DSCQS, a subject is presented with a pair of sequences two consecutive times. One of two sequences is the source video input and the other is the test video sequence obtained by processing the source input. The subject is asked to evaluate the picture quality of both sequences using a continuous grading scale. The grading scale is composed of two identical 10cm graphical scales which are divided into five equal intervals. Fig. 1 shows grading scale.

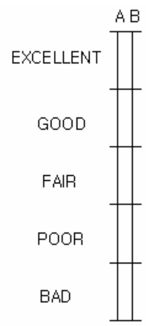

Fig. 1 Grading scale

Source and processed sequences are presented in random order. The DSCQS is considered the most reliable and widely used because it has low sensitivity to contextual effects. Contextual effects mean subjective ratings are influenced by the severity and ordering of impairments within the test session [3].

Although the most reliable method for assessing image quality is subjective test by human observer, the Mean Opinion Score (MOS), which rates results from a subjective test and requires the services of a number of human observers, is inconvenient and expensive. Furthermore, the subjective image quality assessment is generally too slow to be easily used in real-world applications.

In consequence, objective image quality evaluation approaches are necessarily used in most application. They can be generally categorized into three folds: One is Full Reference (FR) requiring a complete reference image. Second is known as Reduced Reference (RR). RR is useful when the reference image is only partially available. Both of FR and RR need the reference image. However the reference images are not always available in most cases. No Reference (NR) is the third metric to satisfy the cases when reference images are unavailable.

III. PROPOSED WORK

In this paper, we propose NR method which accounts only blur and noise. Although image quality is affected by many features like hue, edge, noise, and contrast, we assume that noise and blur are the most important factors on image quality degradation. The proposed work searches and quantifies the blur and noise as image quality factors.

The framework of the proposed model is shown in Fig. 2.

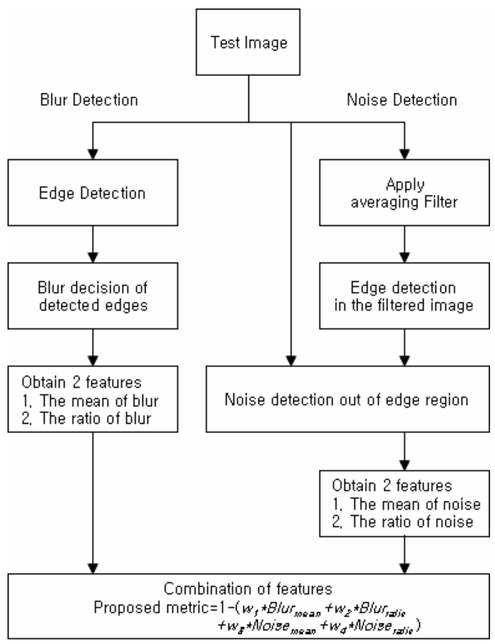

Fig. 2 Framework of the proposed model

Most of the digital cameras have the Image Signal Processor (ISP) to enhance the output of image sensor. One of the important functions of ISP is to remove noise: noise reduction. The strong noise reduction removes noise sufficiently but makes detail and texture blurred. In order to reduce much trial caused by trade-off between noise reduction and detail loss, the criteria for image quality control parameters are required. In case only one feature between blur and noise is considered for quality prediction, the results are to be insufficient for finding the optimized parameters of noise reduction. The proposed image quality metric meets the necessary criteria because our method analyzes both blur and noise simultaneously.

The proposed method calculates blur and noise in a spatial domain. Only the luminance parts of the images are used to estimate blurriness and noise. The blur is measured by simple numeric operations on pixel.

A. Blur Measurement

The blurriness is perceptually determined by human observers regardless of the type of blurring, for example, noise reduction, compression, motion blur, and out of focus. In the paper, we seek to find blur without any assumption about its formation.

Blur estimation is divided into 2 stages: First is edge detection and second is blur decision. The blur in the paper is estimated by difference between the intensity of current pixel and average of neighbor pixels. The difference is then normalized by the average. Fig. 3 shows the blur estimation. If the intensity of center pixel is closer to the average intensity of both side pixels, the center pixel is supposed to be on blurred edge.

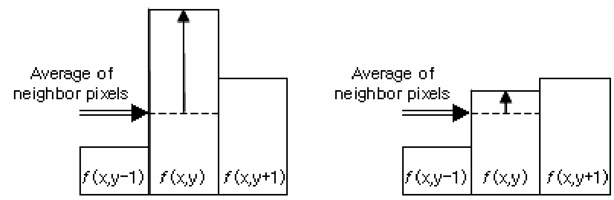

Fig. 3 Blur estimation

We denote the test image with M rows and N columns as $f(x,y)$, for $x \in [1,M]$ and $y \in [1,N]$. The horizontal absolute difference value of a pixel is defined by

$$D_h(x,y) = \left| f(x,y+1) - f(x,y-1) \right|. \tag{1}$$

The mean value of (1) for the image is calculated by

$$D_{h-mean} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} D_h(x,y). \tag{2}$$

In case the value of (1) is larger than (2), the pixel of (1) becomes edge candidate $C_h(x,y)$. If the $C_h(x,y)$ of center pixel is bigger than horizontally adjacent pixels $\{C_h(x,y-1), C_h(x,y+1)\}$, the pixel is determined to be on the edge. The decision of edge pixel $E_h(x,y)$ is summarized as follows

$$C_h(x,y) = \begin{cases} D_h(x,y) & \text{if } D_h(x,y) > D_{h-mean} \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

$$E_h(x,y)=\begin{cases}1 & \text{if } C_h(x,y)>C_h(x,y+1) \text{ and } C_h(x,y)>C_h(x,y-1)\\ 0 & \text{otherwise}\end{cases} \quad (4)$$

Now, we examine whether the detected edge pixel is blurred or not. The ratio value for blur decision is obtained horizontally by (5) ~ (6)

$$A_h(x,y)=\frac{1}{2}D_h(x,y), \quad (5)$$

$$BR_h(x,y)=\frac{|f(x,y)-A_h(x,y)|}{A_h(x,y)}. \quad (6)$$

In the same way, we can also estimate BR_v in the vertical direction from (1) ~ (6). The larger value between BR_h and BR_v is selected for final decision, which is called inverse blurriness in the paper.

$$B(x,y)=\begin{cases}1 & \text{if } \max(BR_h(x,y),BR_v(x,y))<Th_B\\ 0 & \text{otherwise}\end{cases}. \quad (7)$$

The equation (7) means the center pixel with inverse blurriness (or max(BR_h, BR_v)) under Th_B is considered as blurred. By the experiment, blur can be best detected when Th_B is 0.1. Finally, the mean of blur and ratio to the edge is calculated by

$$Blur_{mean}=\frac{Sum_{blur}}{Blur_{cnt}}, \qquad Blur_{ratio}=\frac{Blur_{cnt}}{Edge_{cnt}} \quad (8)$$

, where Sum_{blur} and $Blur_{cnt}$ are the sum of inverse blurriness and the count of blurred pixels, respectively. $Edge_{cnt}$ is the total number of edge pixels.

B. Noise Measurement

Since the noise along edges perceptually looks less apparent, we measure the noise out of the edge region.

The edge detection can also be affected by noise. Hence, a pre-processing for noise filtering is needed prior to detecting the edge. In the paper, we apply an average filter to the noisy test image to remove the noise. The averaging filtered image $g(x,y)$ is generated by

$$g(x,y)=\frac{1}{3\times 3}\left[\sum_{i=-1}^{1}\sum_{j=-1}^{1} f(x+i,y+i)\right]. \quad (9)$$

We obtain the edge pixels on $f(x,y)$ in the similar way to blur measurement. The noise candidate pixels are estimated as follows

$$D_h(x,y)=|g(x,y+1)-g(x,y-1)|. \quad (10)$$

$$D_{h-mean}=\frac{1}{M\times N}\sum_{x=1}^{M}\sum_{y=1}^{N} D_h(x,y). \quad (11)$$

From (10) and (11) are repeated in vertical direction. Then, we continue to find the final noise as

$$N_{cand}(x,y)=\begin{cases}\max(D_v(x,y),D_h(x,y)) & \text{if } D_h(x,y)\le D_{h-mean} \text{ and } D_v(x,y)\le D_{v-mean}\\ 0 & \text{otherwise}\end{cases} \quad (12)$$

, where $N_{cand}()$ represents the noise candidate which is zero on edge region.

$$N_{ccand-mean}=\frac{1}{M\times N}\sum_{x=1}^{M}\sum_{y=1}^{N} N_{cand}(x,y).$$

$$N(x,y)=\begin{cases}N_{cand}(x,y) & \text{if } N_{cand}(x,y)>N_{cand-mean}\\ 0 & \text{otherwise}\end{cases} \quad (13)$$

The noise decision is made by (13). The mean of noise and ratio to the total number of pixels is generated by

$$Noise_{mean}=\frac{Sum_{noise}}{Noise_{cnt}}, \qquad Noise_{ratio}=\frac{Noise_{cnt}}{M\times N} \quad (14)$$

, where Sum_{noise}, and $Noise_{cn,t}$ are the sum of $N(x,y)$ and the total number of noise pixels, respectively.

C. Combination of Blur and Noise

The extracted features are combined to generate a quality prediction model. The proposed metric is linear equation, which means the computational load is very low. Most of existing NR image quality metrics use exponential or non-linear term [4]-[6].

Our proposed metric is given by

$$\text{Proposed Metric}=1-(w_1 Blur_{mean}+w_2 Blur_{ratio} + w_3 Noise_{mean}+w_4 Noise_{ratio}) \quad (15)$$

, where $w1$, $w2$, $w3$, and $w4$ are the weights estimated from intensive subjective test data. We obtain the weights by using linear regression analysis. As a result of linear regression analysis, $w1$, $w2$, $w3$, and $w4$ are 1.55, 0.86, 0.24, and 0.66 respectively with R-Square 83.2%. The R-Squared value stands for the % variation of the data explained by the fitted line. The closer the points are to the line, the better the fit is. We optimize the obtained parameters by exhausted experiments as follows

$$w_1=1,\ w_2=0.95,\ w_3=0.3, \text{ and } w_4=0.75 \quad (16)$$

The weights in (16) show that the spread of noise is more sensitive to human perception. We also can see the more important fact that the blurriness is more sensitive to human perceptual quality assessment than noise. The result provides important direction for adjusting the image quality control parameters of noise reduction of ISP.

IV. Experimental Results

We choose JPEG2000 and white noise images from LIVE image database for experiments. The LIVE image database is composed of 29 color reference images and their distorted images using the following distortion type: JPEG2000, JPEG, White noise, Gaussian blur, and bit errors. Every distorted image includes corresponding Difference Mean Opinion Score (DMOS). The 29 reference images are shown in Fig. 4. The JPEG2000 images are generated with various bit rates and contain edge distortion like blur and ringing. The white noise images are made of White Gaussian noise.

In order to obtain and verify the weights of (16), the test images composed of JPEG2000 and white noise are divided into 2 parts. The first part for training consists of randomly selected 30 images including 15 JPEG2000 images and 15 white noise images. The second part for testing contains the rest of the JPEG2000 and white noise images: 154 JPEG2000 and 53 white noise images.

Fig. 4 The reference images of LIVE database

We use Pearson linear coefficient for the performance accuracy of the proposed metric and compare our proposed metric with PSNR and NR method by Z. Wang, H.R. Sheikh, and A.C. Bovik. They provide the Matlab source code for No-reference perceptual quality assessment of JPEG compressed images [7], which analyzes blocking artifacts and blur of JPEG compressed image. Wang's NR metric shows low correlation with DMOS because of focusing on the predefined artifacts of JPEG like blockiness. In experiment using training images, our proposed metric is better correlated with the subject ratings than PSNR by 9.7%. The comparison result is shown in Table I. Fig. 5 and Fig. 6 show the scatter plot of DMOS versus PSNR and DMOS versus proposed metric respectively.

TABLE I
Pearson Linear Correlation Coefficient

Metric	Training images
PSNR	-0.833
Wang	-0.437
Proposed	-0.914

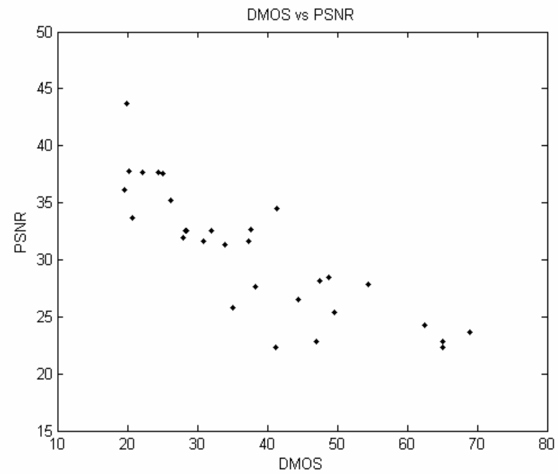

Fig. 5 PSNR compared with DMOS using training images

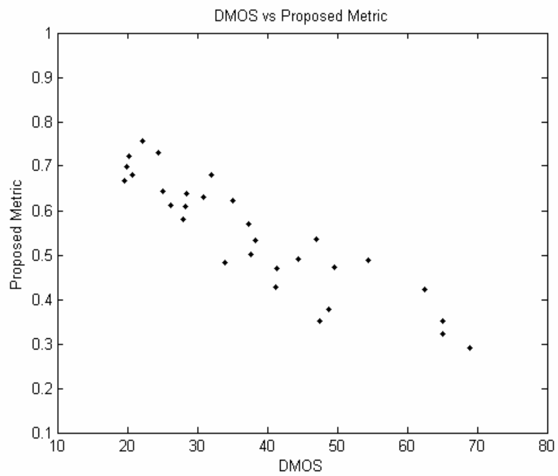

Fig. 6 Quality predictions by proposed metric using training images

We verify the validity of obtained weights (16) using the testing images. Again, the proposed metric provides the results that correlate well with DMOS. The overall performance of the proposed metric is outstanding as compared with PSNR. Since Pina Marzilizno et al. use the same JPEG2000 images from LIVE for their perceptual blur and ringing metrics and provide result as linear correlation, relative comparison with the proposed metric is possible [8]. In test using only JPEG2000 images of LIVE, our proposed metric is more correlated than Marzilizon's: 0.9 and 0.86 respectively. The summary of both experimental results are shown in Table II. Fig. 7 and Fig. 8 illustrate the correlation of PSNR and the proposed metric respectively.

TABLE II
PEARSON LINEAR CORRELATION COEFFICIENT

Metric	Training images	Testing images
PSNR	-0.833	-0.824
Proposed	-0.914	-0.910

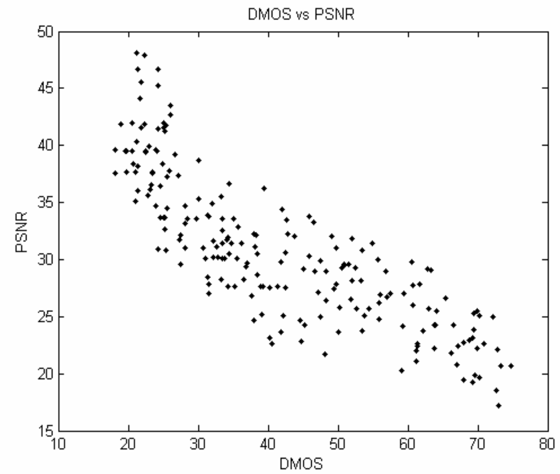

Fig. 7 PSNR compared with DMOS using testing images

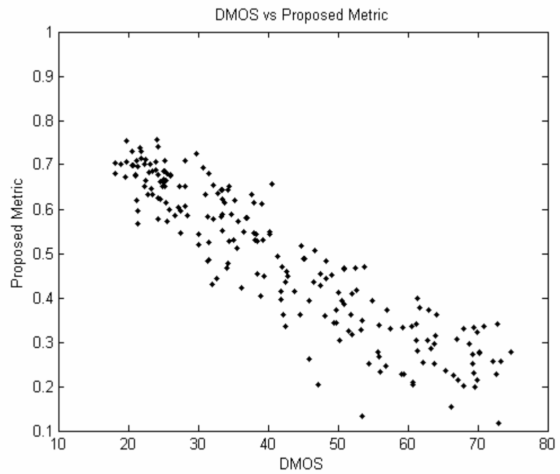

Fig. 8 Quality predictions by proposed metric using testing images

V. CONCLUSION

The growing popularity of camera function in Information Technology (IT) product has made it absolute necessary to develop NR image quality assessment. We demonstrated a novel NR perceptional image quality assessment. We find critical criteria to determine objective quality metric. The exhausted experimental results show the new method highly correlates with subjective ratings. From the relation among the weighting parameters (16), we find important guidance that the sharpness is more significant than noise. The blur is more sensitive than noise to human perceptual assessment. Our future work includes investigating and improving the dependencies between the subjective opinion score and pooling method.

REFERENCES

[1] Z. Wang, A.C. Bovik, and L. Lu, "Why is image quality assessment so difficult?", Proc. IEEE Inter. Conference Acoustics, Speech, and Signal Processing(ICASSP-2002),Vol.4, pp. 3313-3316, Orlando, FL, 13-17 May 2002.
[2] Video Quality Experts Group (VQEG), http://www.vqeg.org.
[3] M. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies", Proceedings of the SPIE, Vol. 5150, pp. 573-582, 2003.
[4] Z.M. Parvez Sauzzad, Y. Kawayoke, and Y. Horita, "No reference image quality assessment for JPEG2000 based on spatial features", Signal Process: Image Communication 23 (2008) pp.257-268.
[5] R. Venkatesh Babu, S. Suresh, and Andrew Perkis, "No-reference JPEG-image quality assessment using GAP-RBF", Signal Processing 87 (2007) pp.1493-1503
[6] Z. Wang, H.R. Sheikh, and A.C. Bovik, "No-reference perceptual quality assessment of JPEG compressed images", Proceedings of the ICIP,02, vol. 1, pp. 477-480, 2002
[7] H.R. Sheikh, Z. Wang, L. Cormack and A.C. Bovik, "LIVE Image Quality Assessment Database Release 2", http://live.ece.utexas.edu/research/quality.
[8] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: application to JPEG2000", Signal Process: Image Communication 19 (2004) pp.163-172.