A No-Reference Adaptive Metric for Digital Image Quality Assessment

Jin-Young Lim, Dong-Wook Kang, Ki-Doo Kim and Kyeong-Hoon Jung,

School of Electrical Engineering, Kookmin University.
861-1 Jeongneung-dong, Seongbuk-gu, Seoul, 136-702, Korea
E-mail: {limjy, dwkang, kdk, khjung}@kookmin.ac.kr

ABSTRACT

In this paper, a reference-free perceptual quality metric is proposed for image assessment. It measures the amount of overall blockiness and blurring in the image. And edge-oriented artifacts, such as ringing, mosaic and staircase noise are also considered. In order to give a single quality score, the individual artifact scores are adaptively combined according to the difference between the edge-oriented artifacts and other artifacts. The quality score obtained by the proposed algorithm shows strong correlation with the MOS values by VQEG.

Keywords: Image quality assessment, No-reference, MOS

1. INTRODUCTION

As many multimedia services are based on the digital image and video, how to measure the quality of them is one of the significant issues to determine whether the QoS (Quality of Service) or QoE (Quality of Experience) of services is satisfied or not. However it is not easy to find a practical metric for image or video quality measurement, since the quality assessment is inherently subjective process and to adopt the subjective metric is not efficient in terms of time and cost. Unfortunately, traditional and very popular PSNR or MSE cannot be a reliable metric. It often shows too much discrepancy with the perceived quality by human. Therefore, there have been many efforts to develop a useful objective metric to reflect the human perception.

Objective image quality metrics are generally divided into FR (Full-reference), RR (Reduced-reference) or NR (No-reference) according to the existence of reference image [1]. Among these metrics, NR method is the most difficult and realistic one since it is very plausible to assume there is no information about the original image at the user or consumer side.

It is well known that the perceived quality of image is greatly affected by the blockiness and blurring artifacts in the block-based processed image. Thus many metrics have considered the amount of these artifacts which are mainly due to the loss of high frequency component at block boundary when block-based DCT is applied in JPEG. Wang proposed a metric for the blockiness and blurring measurement by calculating the difference at the block boundary [2]. Meanwhile, Marziliano classified the blurring area through edge mask and calculate the amount of blurring only in that area [3].

In this paper, we have observed the edge-oriented artifacts, such as ringing, mosaic and staircase noise can considerably affect the perceived image quality especially where the blockiness or blurring artifact is not severe. The individual metrics for blockiness, blurring, and edge-oriented artifacts are discussed in Section 2. And an adaptive approach to combine those metrics is devised to generate a single score in Section 3. Section 4 shows the simulation results and Section 5 gives the conclusion.

2. DISTORTION MEASURES

The blockiness, blurring, and edge-oriented artifacts are the main concern in measuring the image quality. Even though there is some correlation between the first two artifacts, it is helpful to separate them because the relationship becomes weak when the image contains text or caption.

It is true that the image quality is almost determined by these two artifacts. But the ringing, staircase, and mosaic noise need to be also considered to increase the reliability of metric. These artifacts are mostly found around the strong edge and called as the edge-oriented artifact [4].

2.1 Metric for blockiness

The blockiness which is the most unpleasant artifact comes from the discontinuity between adjacent blocks in image. And human perception of this artifact is affected by the surrounding neighborhood, that is, it can be easily found in low activity area but is likely to be masked in high activity area. Thus the local activity in image needs to be taken into account in measuring the amount of blockiness.

Firstly, the average and standard deviation of each block is calculated as follows.

$$\mu_{m,n} = \frac{1}{8\times 8}\sum_{i=0}^{7}\sum_{j=0}^{7} I(m-i, n-j) \qquad (1)$$

$$\sigma^2_{m,n} = \frac{1}{8\times 8}\sum_{i=0}^{7}\sum_{j=0}^{7} [I(m-i, n-j) - \mu_{m,n}]^2 \qquad (2)$$

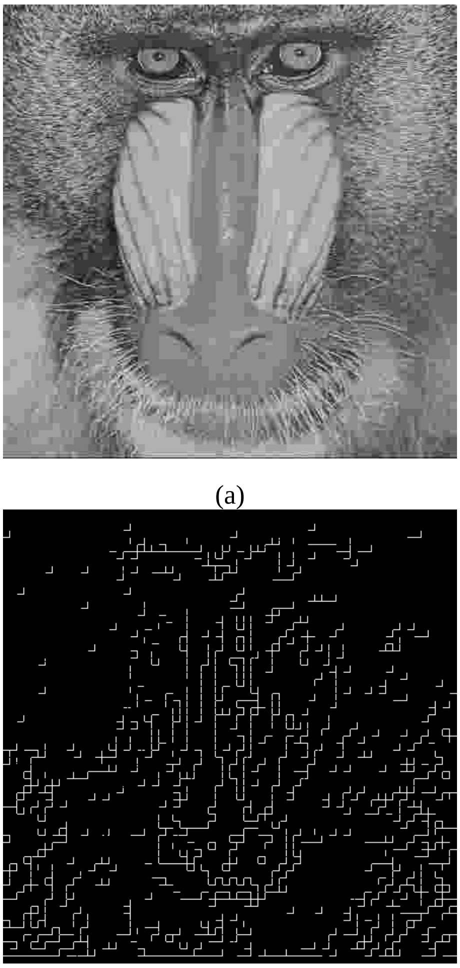
Where we assume the size of block is 8 by 8 and $I(m,n)$ is the intensity value of pixel and (m,n) denotes the position of block boundary. Next, the horizontal and vertical difference values are calculated as in Eq. (3) and (4).

$$\Delta^{h}{}_{m,n} = \frac{1}{8}\sum_{i=0}^{7}[I(m-i,n)-I(m-i,n+1)] \qquad (3)$$

$$\Delta^{v}{}_{m,n} = \frac{1}{8}\sum_{j=0}^{7}[I(m,n-j)-I(m+1,n-j)] \qquad (4)$$

In order to consider the local activity in image, the blockiness artifact is calculated only where the condition of $\Delta^{h}{}_{m,n} \ge \gamma\sigma_{m,n}$ or $\Delta^{v}{}_{m,n} \ge \gamma\sigma_{m,n}$ is satisfied.

Fig. 1 (a) shows the test image and the resulting blockiness mask. As we can see from Fig. 1 (a), the blockiness in the nose area of baboon is clearly noticeable but the one in the cheek is hard to be found. And we can find the mask of Fig. 1 (b) successfully explain this observation.



(a)

(b)

Fig. 1. (a) The test JPEG image, (b) the blockiness mask.

Then the blockiness score in the mask is determined by Sugeno fuzzy integral [5]. The equation of Sugeno fuzzy integral is given as follows.

$$(S)\int_A I_{masked}(m,n)dg = \sup_{\alpha\in[0,1]} \min\{\alpha, g(A \cap F_\alpha(I))\} \qquad (5)$$

$$F_\alpha(I) = \{m,n \mid I_{masked}(m,n) \ge \alpha\} \qquad (6)$$

where $I_{masked}(x,y)$ denotes the masked image with the blockiness mask, A is a set of pixel value in the test image, and g is a fuzzy measure. The result value of the equation (5) is the blockiness score and denoted by S_{blk}.

2.2 Metric for Blurring

Unfortunately, it is impossible to directly measure the amount of blurring artifact without reference image. Therefore a circumvent method is developed by using the fact that the blurring artifact is closely related to the blockiness. If there is no blockiness artifact, it is natural to assume the high frequency components are well preserved. Thus we make use of the blockiness mask to measure the blurring artifact. Fig.2 shows three positions in a single row where the blockiness is detected in the mask. The current position is denoted by P and the previous and the next position are P'_P and P'_N, respectively.

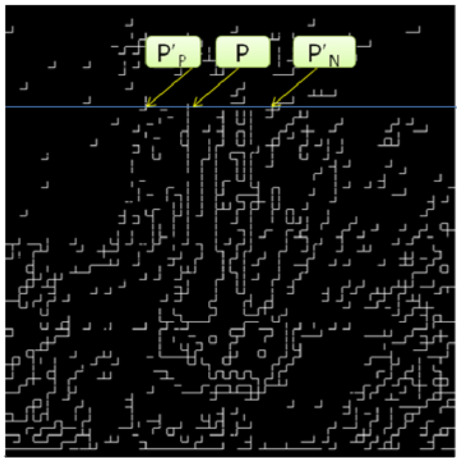


Fig. 2. Three positions to calculating the blurring artifact.

As the blurring increases, the variation of pixel intensity decreases. Also the blurring depends on the spatial distance. To get the blurring score, we calculate the difference of the maximum and minimum value along the row and divide it by the distance from P'_P to P'_N. The positions of P'_P, P, and P'_N are (m,n_{i-1}), (m,n_i), and (m,n_{i+1}), respectively. The final blurring score S_{blur} is obtained by the reciprocal of the result as Eq. (7) and (8), where $K(j)$ is the number of blockiness position in m-th row.

$$S_{blur} = \left[\sum_{j=0}^{M-1}\sum_{i=1}^{K(j)-1}\frac{Diff(i)}{n_{i+1}-n_{i-1}}\right]^{-1} \qquad (7)$$

$$Diff(i) = \max(I(m,n_{i-1}) : I(m,n_{i+1})) - \min(I(m,n_{i-1}) : I(m,n_{i+1})) \qquad (8)$$

2.3 Metric for Edge-oriented Artifacts

When there are strong edges in image, the blockiness and blurring are not sufficient to measure the quality. We can easily found the ringing, staircase, or mosaic artifacts around edge in relatively low texture area. These artifacts are considered to be related with the quantization noise and have different characteristic from the blockiness or blurring.

In order to measure the amount of these artifacts, we need to detect the principal edge and its direction. Three masks for edge detection are given in Fig. 3. These masks are applied to the DCT coefficients of 8x8 image blocks after masking operation and the sum of coefficients is obtained. If the difference between the largest and second largest sum exceeds the threshold, we consider there exists an edge with corresponding direction.

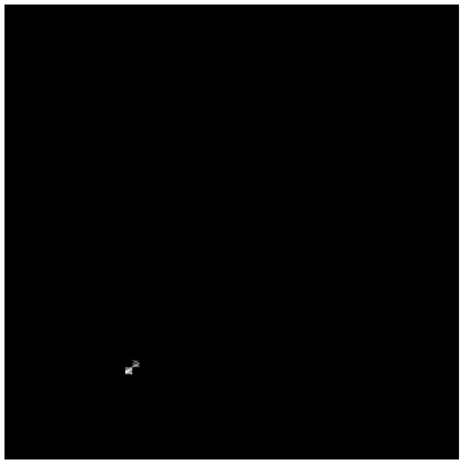
$$\begin{bmatrix} 0&1&1&1&1&1&1&1\\ 0&1&1&1&1&1&1&1\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&0&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0\\ 1&1&0&0&0&0&0&0 \end{bmatrix} \begin{bmatrix} 0&0&0&0&0&0&0&0\\ 0&0&1&0&0&0&0&0\\ 0&1&1&1&0&0&0&0\\ 0&0&1&1&1&0&0&0\\ 0&0&0&1&1&1&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix}$$

(a) (b) (c)

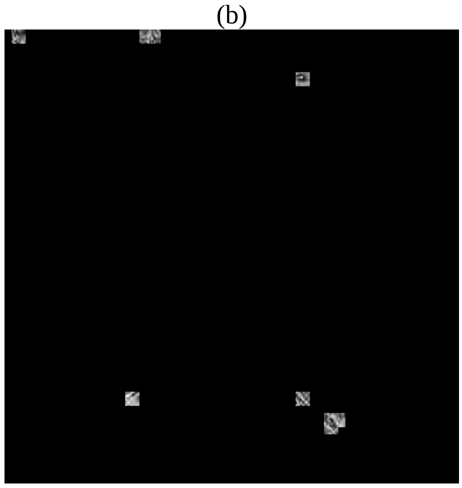
Fig. 3. Edge detecting masks in DCT domain,
(a) vertical mask, (b) horizontal mask, (c) diagonal mask.

The staircase noise is expected to be found along the diagonal edges. Therefore the section where the edge is detected and its direction is diagonal is classified to the area of staircase artifact. Meanwhile, the mosaic phenomena come from the cohesion of visible distortions. So if the cohesion size is larger than 16x16, these sections are classified to the area of mosaic artifact. And the remaining edge section is the area of ringing artifact. Fig. 4 (a), (b) and (c) show the areas of these edge-oriented artifacts.

(a)



(b)



(c)

Fig. 4. (a) ringing artifact area, (b) staircase artifact area, (c) mosaic artifact area

To obtain the individual artifact score, the Sugeno fuzzy integral in Section 2.1 is used again. The ringing, staircase, and mosaic scores are denoted by S_R, S_S and S_M, respectively.

3. ADAPTIVE COMBINATION OF METRICS

The ultimate goal of quality metric is to generate a single score to measure the image quality as a whole. The five individual scores in Section 2 need to be combined. We observed that the blockiness or blurring artifact is dominant in most case and the influence of the edge-oriented artifact is relatively small. If blockiness and blurring artifacts are too large in image, it is needless to consider the ringing, staircase, and mosaic artifacts. However, these artifacts become significant factor around strong edge.

Therefore, it is reasonable to use an adaptive method. The edge-oriented artifact is considered only where the blockiness or blurring is not severe. Fig.5 shows the overall block diagram of the proposed adaptive approach.

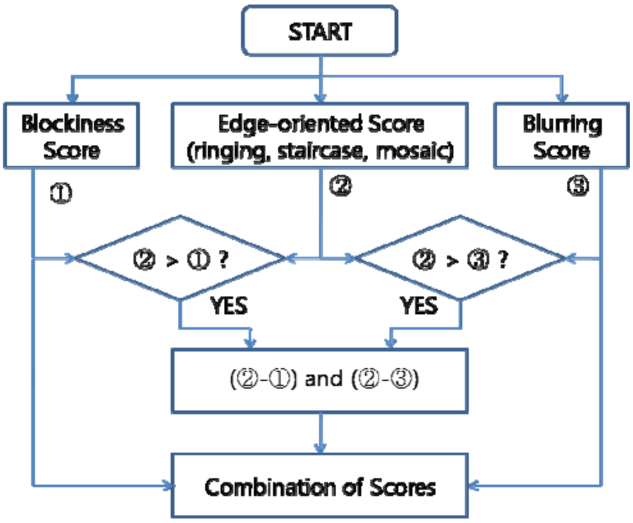


Fig. 5. Adaptive combination of individual scores.

When we have the individual scores, the differential ringing, staircase, and mosaic scores are calculated as Eq. (9) and (10). And the final quality score, QS, is calculated as Eq. (11), where α and β are the parameters which can be easily determined by using nonlinear regression.

$$\Delta S_x = (S_x - S_{blk})u(S_x - S_{blk}) + (S_x - S_{blur})u(S_x - S_{blur}) \quad (9)$$

$$u(S_x - S_y) = \begin{cases} 1 & S_x \geq S_y \\ 0 & otherwise \end{cases} \quad (10)$$

$$QS = \alpha S_{blk} + \beta S_{blur} - (\Delta S_R + \Delta S_S + \Delta S_M) \quad (11)$$

4. EXPERIMENTAL RESULTS

As we can see in the Fig.1 (b) and Fig.4 (a)~(c), various artifacts are successfully detected and well matched to the human perception. And the MOS (Mean Opinion Score) results of the proposed method are given in Table.1 which shows the MOS results of VQEG (Visual Quality Expert Group) [6] and Wang's method [2]. The five test JPEG images are also obtained from VQEG's web site [6].

Table 1. The comparison of MOS results

Test image		VQEG[6]	proposed	Wang[2]
barba	r1	4.462	3.699	3.761
	r2	3.731	2.987	3.392
	r3	2.423	2.059	2.703
	r4	1.385	1.491	2.139
	r5	1.077	1.250	1.758
clown	r1	4.462	3.812	3.772
	r2	3.577	3.543	3.263
	r3	2.500	2.930	2.300
	r4	1.769	2.539	1.851
	r5	1.077	1.634	0.890
fruit	r1	4.577	4.816	3.743
	r2	4.231	3.946	3.406
	r3	3.346	2.401	2.521
	r4	2.192	1.892	1.777
	r5	1.500	1.670	1.220
isabe	r1	4.577	5.420	3.683
	r2	4.231	4.466	3.276
	r3	3.154	3.421	2.682
	r4	2.231	2.397	1.843
	r5	1.462	1.995	1.169
mandr	r1	4.577	4.332	3.546
	r2	4.500	3.945	3.304
	r3	3.346	2.852	2.374
	r4	2.923	2.373	2.028
	r5	1.615	1.602	1.131

Generally speaking, the proposed metric have generated satisfactory quality score to show strong relationship with the MOS of VQEG and its performance is superior to that of Wang's metric [2]. Table 2 shows the MAD of MOS values with respect to the VQEG's results. The proposed metric could improve about 30% over the Wang's metric on the average.

The performance is, of course, image-dependant but the proposed method has some noticeable gain when the high-activity areas are irregularly distributed like the test image of 'mnrd', because the ringing artifact is incorporated. In case of 'clown' image, however, Wang's method outperformed to the proposed method, because there are too closely located check or stripe patterns.

Table 2. MAD(Mean Absolute Difference) of MOS

Test image	Proposal	Wang[2]
barba	0.430	0.551
clown	0.488	0.294
fruit	0.388	0.636
Iabe	0.409	0.600
mndr	0.371	0.916
Mean	0.417	0.599

5. CONCLUSIONS

In this paper, we proposed a no-reference algorithm to measure the quality of digital image. The individual metrics for several visual artifacts, such as blockiness, blurring, ringing, staircase, and mosaic, were discussed and each artifact scores were adaptively combined to give a single quality score by comparing the edge-oriented artifacts with other artifacts. The MOS of the proposed metric showed a strong relationship with the MOS of VQEG. And it produced better results over Wang's metric, in most case.

As the size of test image pool increases, it is expected that the resulting MOS is more reliable. And the adaptive scheme needs to be carefully devised to get more trustworthy results since the performance of quality metric basically depends on the local characteristics of image.

6. REFERENCES

[1] H. R. Wu, K. R. Rao, Digital Video Image Quality and Perceptual Coding, CRC Press 2006.

[2] Zhou Wang, H. R. Sheikh, Alan C. Bovik, "No-reference perceptual quality assessment of JPEG

compressed images," IEEE International Conference on Image Processing, vol.1, pp 477-480, Sep. 2002.
[3] Pina Marziliano, Frederic Dufaux, Stefan Winkler, Touradj Ebrahimi, "A no-reference perceptual blur metric," http://infoscience.epfl.ch.
[4] M. Yen, H.R. Wu. "A survey hybrid MC/DPCM/DCT video coding distortions", Signal Processing, vol. 70 pp.248-278, 1998.
[5] Junli Li, Zheru Chi, Gang Chen, "Image retrieval based on Sugeno fuzzy integral," International Conference on Image and Graphics, vol. 18, pp.160-163, Dec. 2004.
[6] http://www.vqeg.org/