A Perceptual No-Reference Blockiness Metric for JPEG Images

R. Venkatesh Babu*
IRISA / INRIA
Campus Universitaire de Beaulieu
Rennes Cedex, France
venkat@irisa.fr

Ajit S. Bopardikar   Andrew Perkis
Centre for Quantifiable QoS in Communication Systems
Norwegian University of Science and Technology
Trondheim, Norway
{ajit,andrew}@Q2S.ntnu.no

Abstract

In this paper [1] we present a novel no-reference (NR) metric to measure block impairment (blockiness) in JPEG-coded images. The proposed metric integrates several key human visual sensitivity factors such as edge amplitude, edge length, background activity and background luminance to evaluate the effect of block edge impairment on perceived image quality. The subjective test results of our metric is compared with the Wang-Bovik's NR blockiness metric [9]. The results show that the proposed metric correlates well with the mean opinion score (MOS) than Wang-Bovik's blockiness metric. Further this metric can be extended to predict the quality of the MPEG/H.26x compressed videos.

1. Introduction

The main objective of Image/Video quality assessment metrics is to provide an automatic and efficient system to evaluate visual quality. Such measurements should exhibit good correlation with perception by the human visual system (HVS). The most widely used objective image quality metrics, namely mean square error (MSE) and peak signal to noise ratio (PSNR), do not correlate well with the human perception [9] and require the original image for computing the distortion. Therefore there is a need to develop metrics that closely correlate with human perception. Most of the images available nowadays on the Internet and in multimedia database are in lossy-compressed form. For such a situation, measuring image quality becomes further difficult due to non-availability of the original reference image.

There has been considerable amount of research done to develop objective image/video quality metrics that incorporate perceived quality measurement by considering HVS characteristics. However most of the proposed metrics based on HVS characteristics require the original image as reference [2, 10, 6, 1]. Though it is easy for human observers to assess the quality of the image without any reference image, developing a no-reference (NR) quality metric is a difficult task. Hence apriori knowledge about the artifacts are essential to develop NR metrics. Nowadays, NR quality metrics are the subject of considerable attention by the research community with the emergence of Video Quality Experts Group (VQEG) [8], which is in the process of standardizing the NR and reduced-reference (RR) video quality assessment methods.

The major source of distortion in image/video is due to the block DCT-based compression. The most popular and widely used image format in the Internet as well as in digital cameras is JPEG [4]. Because JPEG uses block-based DCT transform for coding to achieve compression, the major artifact JPEG-compressed images suffers is blockiness. In JPEG coding, non-overlapping 8×8 pixel blocks are coded independently using DCT transform. The compression (bit-rate) and Image quality are mainly determined by the degree of quantization of these DCT coefficients. The effects of quantization manifest as blockiness, ringing and blurring artifacts in the JPEG coded image. The subjective data for all these artifacts are highly correlated [3]. Hence measuring the blockiness in-turn indicates the overall image quality.

In this paper we propose a NR blockiness metric considering various human visual factors. The following important HVS criteria are considered in developing the metric: i) edge amplitude ii) edge length iii) background luminance and iv) background activity or texture.

The paper is organized as follows: Section 2 presents the past work related to blockiness metric. Section 3 presents the basic ideas underlying the blockiness metric and the algorithm to compute the metric. Subjective test results and discussion are presented in Section 4. Finally Section 5 con-

*This work was carried out during the tenure of an ERCIM fellowship of the first author

[1]This work was supported by the Centre for Quantifiable Quality of Service in Communication Systems, Centre of Excellence" appointed by The Research Council of Norway. http://www.ntnu.no/Q2S/

cludes the paper.

2. Related Work

Algorithms to measure blockiness have used a variety of methods to do so. Wang and Bovik proposed an algorithm based on computing the FFT along the the rows and columns to estimate the strength of the block-edges [12] while Vlachos used cross-correlation of subsampled images to compute a blockiness metric [7]. Wu and Yuen proposed a metric based on computing gradients along block boundaries while tempering the result with a weighing function based on the human visual system (HVS) [10]. The computations yielded a number for each frame that represented the block edge strength for that frame. Similar ideas about the HVS were utilized by Suthaharan [6] and Gao et al., [1]. The general idea behind these metrics was to temper the block-edge gradient with the masking activity measured around it. These approaches utilize the fact that the gradient at a block-edge can be masked by more spatially active areas around it or, in very dark or bright regions [11]. Several of these approaches have proven to be quite effective but can be computationally quite complex for real-time implementation. The the NR quality assessment of JPEG compressed images proposed in [9] uses the features extracted from a set of images to train the model. Hence it is not guaranteed to work well for all images from a different database. The performance of this metric is compared with the proposed metric in Section 4.

3. Proposed Blockiness Metric

The proposed NR metric is designed to take care of various human visual criteria while measuring the blocking artifact. These blocking artifact appears as horizontal and vertical edge distortions at the boundaries of 8×8 blocks. The visual sensitivity to these edges is affected by the following parameters [2]:

- *Edge Amplitude :* Edge amplitude at the 8×8 block boundaries usually indicates the amount of compression the image is subjected to and increases proportionally with the amount of compression. In other words, the edge-amplitude increases with decreasing bit-rates.
- *Edge Length :* Similar to the above, the edge length also increases with increase in compression.
- *Background Luminance :* The visibility of the edge is often affected by the surrounding spatial region. For example the edge lying in a darker region is less visible compared to the edge in the brighter region.
- *Background Activity :* The blocking artifact will be masked by the activity in the background. For example the edge lying in the textured region is less visible to human observer compared the to the edge in the plain background.

The objective of the proposed metric is to integrate the afore mentioned human visual factors to measure the quality of the JPEG-compressed images. The algorithmic steps for computing the metric are shown in Fig. 1. Because we are interested in obtaining the edges along horizontal and vertical direction, the 'prewitt' horizontal and vertical edge operators are used for this purpose. To obtain the activity along the horizontal and vertical edges a high pass filter is used so as to capture the activity on either side of the horizontal and vertical edges. The final binary activity mask is obtained by hard thresholding the activity measure. This mask allows only the less active regions to be considered for measuring the blockiness. The background luminance weights are obtained based on the model proposed by Karunasekera et al., [2]. Here the darker regions (0 to 127) are given less weightage and the brighter regions (128 to 255) are given constant higher weight. The final horizontal and vertical edge maps are obtained by multiplying the edges with the activity mask and luminance weight. The horizontal and vertical edge profiles are computed from these weighted edge maps. These profiles indicate the edge strength along each row and column of the weighted edge map. Since the effect of blockiness is available only at the block boundaries, every eighth location of the horizontal and vertical profiles is considered for measuring blockiness. The amount of deviation at every eighth location from the average value of the neighborhood of the both (horizontal and vertical) profile is calculated as the amount of blockiness. The algorithm is described in detail in the following subsection.

3.1. Algorithm

In this subsection, the implementation details of the proposed blockiness metric are explained. The algorithm is explained for measuring the blockiness along the horizontal direction. A similar approach is used for the vertical direction.

Consider a gray scale image of size $M \times N$ ($rows \times columns$). The intensity of the image at any pixel location (i, j) is given by $I(i, j)$ which lies in the range of 0 to 255.

1. Obtain the horizontal edge map using horizontal prewitt filter.

$$\hat{E_h} = I * P_h \tag{1}$$

$$E_h(i,j) = \begin{cases} \hat{E_h}(i,j) & : \text{ if } \hat{E_h}(i,j) < \tau_e \\ 0 & : \text{ otherwise} \end{cases}$$

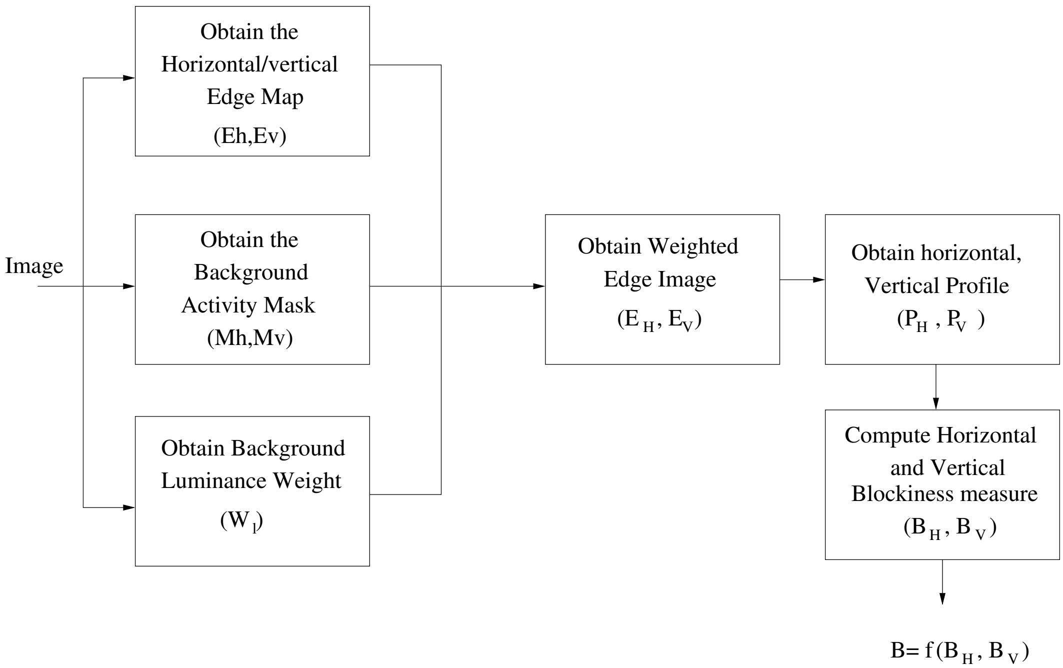

Figure 1. Overview of the proposed blockiness metric

where $P_h = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$ is the prewitt horizontal filter and '$*$' indicates convolution operation. Since the edges caused by blocking artifact has lesser magnitude compared to the image edge, a threshold is used to capture the edges due to blockiness and to avoid the strong image edges. The typical value of the edge threshold used in all our experiments is set at 35. This threshold captures the blocky edges for most of the compression rates.

2. Measure the background activity along horizontal direction

$$\begin{aligned} A_h &= I * F_{ah} \quad \text{where,} \\ F_{ah} &= \frac{1}{8}\begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \end{bmatrix} \end{aligned} \tag{2}$$

The filter F_{ah} captures the background activity along the horizontal edges. The activity values of the entire image is normalized to the range [0-1] by dividing by the maximum activity value of the image.

3. Mask the edges in the active background region by a pre-defined threshold τ_a. The mask M_h is given by

$$M_h(i,j) = \begin{cases} 1 & : \text{ if } \quad A_h(i,j) < \tau_a \\ 0 & : \text{ otherwise} \end{cases} \tag{3}$$

The typical value of the activity threshold used in our experiments is 0.15.

4. The background luminance of every pixel is obtained by using a low-pass filter and the weight of each pixel is assigned as follows:

$$\begin{aligned} W_l(i,j) &= \begin{cases} \sqrt{\frac{I_l(i,j)}{128}} & : \text{ if } \quad 0 \le I(i,j) \le 128 \\ 1 & : \text{ otherwise} \end{cases} \\ \text{where,} & \\ I_l &= I * f_{lp} \end{aligned} \tag{4}$$

where $f_{lp} = \frac{1}{4}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ is the low-pass filter used.

5. The final weighted edge image for the horizontal direction is obtained as:

$$E_H(i,j) = E_h(i,j) \times M_h(i,j) \times W_l(i,j) \tag{5}$$

6. Obtain the vertical profile of E_H (projection of the rows of E_H on a vertical axis).

$$P_v(i) = \sum_{j=1}^{N} E_H(i,j) \tag{6}$$

7. Obtain the blockiness measure along horizontal direction (B_H)

$$B_H = \frac{1}{M}\sum_i |P_{v1}(i) - P_{vm}(i)| \tag{7}$$

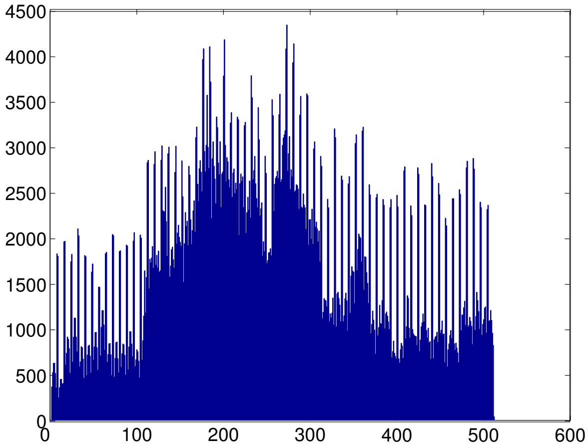

Figure 2. The vertical profile (P_v) of an image.

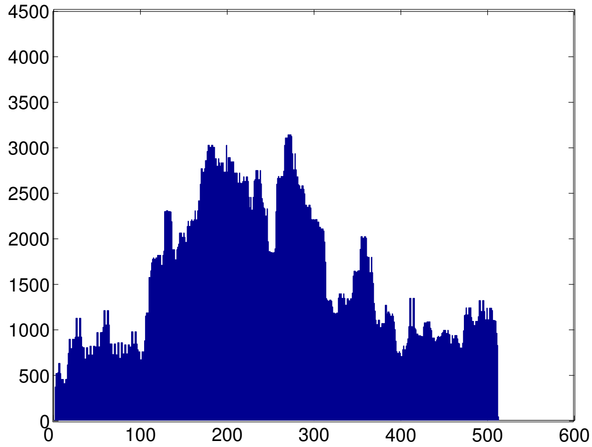

Figure 3. The median filtered vertical profile ($\bar{P}_v$).

$$\text{here,} \quad P_{v1} = P_v(8n) \quad \text{and} \tag{8}$$
$$P_{vm}(n) = \bar{P}_v(8n) \ \text{ where } \ \bar{P}_v = median(P_v)$$

To illustrate this process clearly, the vertical profile P_v and the median filtered vertical profile P_{vm} of an image are shown in Figs. 2 and 3. Now every eighth location of P_v and $\bar{P}_v$ are represented by by P_{v1} and P_{vm}. Fig. 4 shows the difference between P_{v1} and P_{vm}. These differences are added up to give the measure of blockiness along horizontal direction B_H. Similarly the blockiness along vertical direction B_V is also obtained.

8. The final blockiness measure is obtained as a function of both horizontal and vertical blockiness.

$$\widehat{\mathcal{B}} = \sqrt{B_H + B_V} \tag{9}$$

The final blockiness measure B usually lies in the range of 0 to 1. Where 0 indicates the best quality (no blockiness) and 1 indicates the worst quality (extreme blockiness). To compare our metric with the subjective scores, we have inverted it and scaled up to a scale of 0 to 10. The modified final blockiness metric is give by:

$$\mathcal{B} = 10(1 - \widehat{\mathcal{B}}) \tag{10}$$

Hence, the value 10 indicates the best quality and 0 indicates worst blockiness.

4. Subjective Experiments and Discussions

The LIVE Image Quality Assessment Database [5] is used to test the performance of our proposed metric. The details of the database creation are briefly explained in the following lines. To conduct the subjective test, 29 input images were used to create a database. MATLAB's imwrite command was used to write JPEG files at different compression rates to generate a database of 204 compressed images. The total number of images used for the experiment is 233, which includes the original 29 images and 204 compressed images. The study was conducted in two sessions, with the original images included in both. Study 1 contained the first 116 images with 20 subjects and study 2 contained the next 117 images with 13 subjects. The bit rates were chosen such that the resulting distribution of quality scores for the compressed images was roughly uniform over the entire range. Each observer was shown the images randomly. Observers were asked to provide their perception of quality on a continuous linear scale that was divided into five equal regions marked with adjectives Bad, Poor, Fair, Good and Excellent. The scale was then linearly transformed to 1-100 range. The resulting mean opinion score (MOS) was used to test the performance of our blockiness metric.

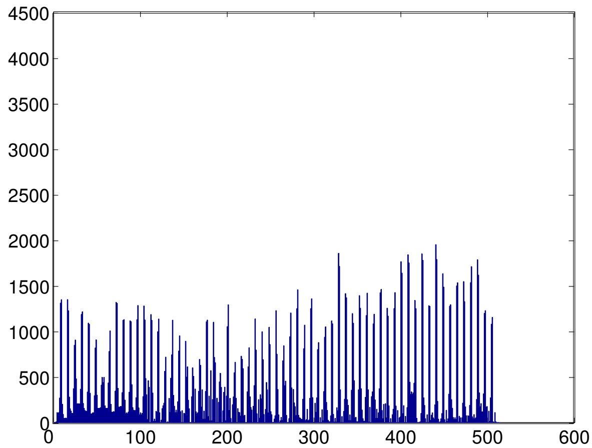

Figure 4. The difference between P_v and $\bar{P}_v$.

The subjective test results for images in the first and sec-

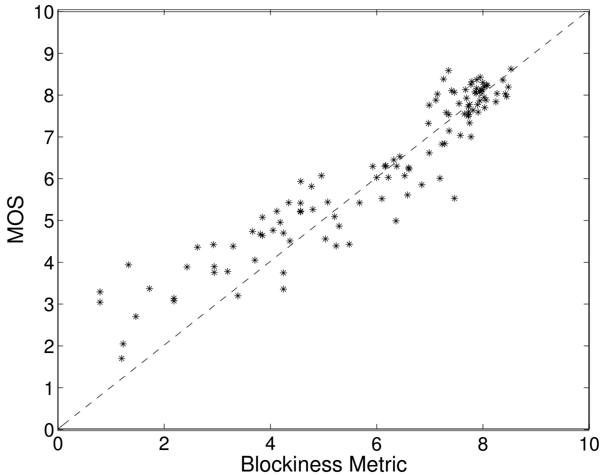

Figure 5. Performance of the proposed metric for Group I images.

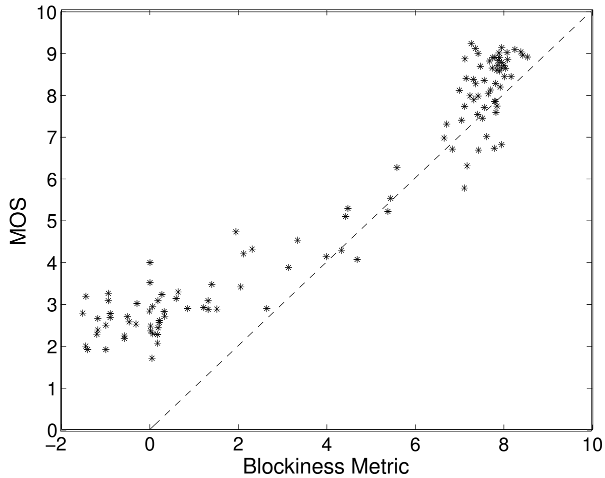

Figure 6. Performance of the proposed metric for Group II images.

ond group are given in Figs. 5 and 6 respectively. The proposed metric correlates well with the subjective test results. To show the efficiency of our proposed metric, we also compare the results with Wang's [9] NR quality metric. In Wang's method the subjective test results are used to train the model for predicting the image quality. The performance of such a model very much depends on the training images used. In our proposed metric there is no such training phase involved. The results of both proposed and Wang's methods for the two groups are shown in Figs. 7 and 8. The results show that the prediction of the proposed metric lies closer to the ideal MOS line (represented by dashed lines). Further the RMS values between MOS and the model prediction of both methods for both groups of images are given in table 1. The results show the proposed method predicts the image quality much better than Wang's method.

Table 1. RMS between MOS and Prediction

Metric	Group I	Group II
Proposed	0.78	1.95
Wang's	1.45	3.33

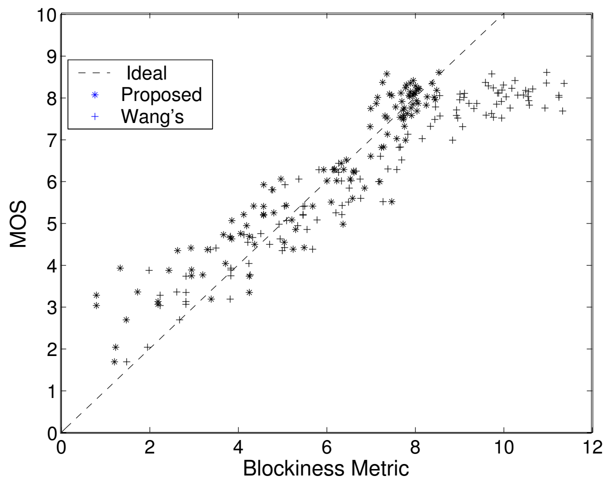

Figure 7. Comparison of proposed and Wang's metrics for Group I images.

5. Conclusions

In this paper we have presented a new NR blockiness metric based on human visual characteristics. Unlike other NR metrics, this metric is not trained for a particular set of images. The performance of the proposed metric is better than the previously reported Wang-Bovik's NR blockiness metric. This metric can be easily extended to measure the quality of MPEG/H.26x compressed videos which use the similar block DCT-based compression. A Matlab implementation of the proposed metric is available at *http://www.q2s.ntnu.no/˜venkat/Quality.html*.

6. Acknowledgment

The authors would like to thank Prof. Bovik and his lab members for providing the JPEG image quality assessment database to test our metric.

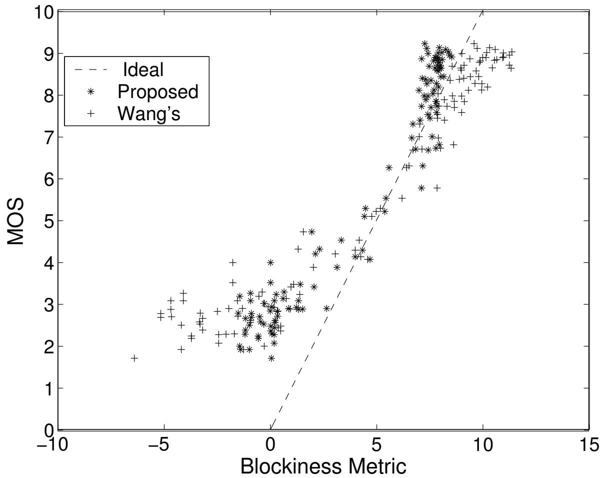

Figure 8. Comparison of proposed and Wang's metrics for Group II images.

References

[1] W. Gao, C. Mermer, and Y. Kim. A de-blocking algorithm and a blockiness metric for highly compressed images. *IEEE Transactions on Circuits and Systems for Video Technology*, 12(12):1150–1159, December 2002.

[2] S. A. Karunasekera and N. Kingsbury. A distortion measure for blocking artifacts in images based on human visual sensitivity. *IEEE Transactions on Image Processing*, 4(6), June 1995.

[3] L. Meesters and J.-B. Martens. A single-ended blockiness measure for JPEG-coded images. *Signal Processing*, 82(3):369–387, March 2002.

[4] JPEG official site. http://www.jpeg.org/

[5] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik. Live image quality assessment database. http://live.ece.utexas.edu/research/quality

[6] S. Suthaharan. A perceptually significant block-edge impairment metric for digital video coding. In *Proc. ICASSP'2003*, volume 3, pages 681–684, Hong Kong, 2003.

[7] T. Vlachos. Detection of blocking artifacts in compressed video. In *Electronics Letters*, volume 36, pages 1106–1108, 2000.

[8] Video Quality Experts Group. (VQEG). website: http://www.vqeg.org

[9] Z. Wang, H. R. Sheikh, and A. C. Bovik. No-reference perceptual quality assessment of JPEG compressed images. In *Proc. ICIP'02*, volume 1, pages 477–480, September 2002.

[10] H. R. Wu and M. Yuen. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 70(3):247–278, November 1998.

[11] M. Yuen and H. R. Wu. A survey of hybrid MC/DPCM/DCT video coding distortions. *Signal Processing*, 4(11):317–320, November 1997.

[12] Z. Wang, A. C. Bovik and B. L. Evans . Blind measurement of blocking artifacts in images. In *Proc. ICIP'00*, volume 3, pages 981–984, September 2000.