

BLIND BLOCKINESS MEASURE BASED ON MARGINAL DISTRIBUTION OF WAVELET COEFFICIENT AND SALIENCY

Seungchul Ryu and Kwanghoon Sohn

Department of Electrical and Electronic Engineering, Yonsei University, South Korea
{ryus01, khsohn}@yonsei.ac.kr

ABSTRACT

The objective measurement of blockiness plays an important role in many applications, such as the quality assessment of an image, and the design of image and video coding system. However, most of the existing no-reference blockiness metrics do not consider important influences of grid distortion of an image on the performance of the metric. In this paper, we propose a new blockiness metric, which is robust to grid distortion, based on the marginal distribution of local wavelet coefficients and saliency information. Experiments for several public image databases showed that the proposed metric provides consistent correlations with subjective blockiness scores and outperforms other existing no-reference blockiness metrics.

Index Terms— blockiness metric, blind, wavelet coefficients, saliency

1. INTRODUCTION

Measuring a quality of an image is a considerable task in various applications such as image restoration, processing, and quality monitoring system [1]. In order to determine a quality of an image, subjective evaluation is considered as the most reliable approach since human beings are end users judging the quality in most practical applications. However, subjective evaluations are costly, time-consuming, and thus impractical for real-time implementation and system integration. Therefore, many objective quality metrics have been developed in recent decades. Objective quality metrics are classified into three categories: full-, reduced-, and no-reference metrics [1]. In the full-reference metric, a distorted image is compared to a distortion-free reference image. In the reduced-reference metric, only partial features of the reference image are used. On the other hand, the no-reference metric does not require any information about a reference image, instead a quality score is computed based on inherent characteristics of a given image. Since no reference image is available in many practical implementations, demands for a no-reference approach have significantly increased.

No-reference quality metrics generally seek to capture one or a few distortions due to the lack of reference information [2, 3]. There are several common distortions that may occur during image processing; for example, white noise is presented due to a sensor and transmission over a communication channel, while an image compression introduces blockiness, blurriness, and ringing artifacts. This paper focuses on blockiness due to its importance in image and video compression, enhancement, and quality assessment. Several no-reference blockiness metrics have been proposed in the literature [4-17]. However, most of the existing no-reference blockiness metrics neglect the important influences of grid

distortions of an image on the performance of the metric. In this paper, a new objective metric is proposed for measuring blockiness in images and videos. The proposed metric is based on the marginal distribution of local wavelet coefficients considering the spatial activity masking effect of the human visual system (HVS). In addition, the salient property of the HVS is considered in the pooling step. Unlike other metrics, this metric can effectively predict the perceived blockiness even when the grid of images is distorted.

The remainder of this paper is organized as follows. Section II presents a brief overview of existing no-reference blockiness metrics. Section III describes the proposed blind blockiness metric in detail. In Section IV, the experimental evaluations of the proposed metric for several public image databases are presented. Lastly, Section V concludes this paper.

2. EXISTING BLOCKINESS METRICS

In recent decades, several no-reference blockiness metrics have been proposed. The objective no-reference metrics for measuring blockiness can be classified into the following four categories: boundary pixel based, spatial domain, frequency domain, and HVS based methods. The majority of the metrics are based on boundary pixels of a block where blockiness generally occurs. In [4], the blockiness was computed as the relative ratio of slopes of the boundary and internal pixels in a 1-D pixel vector made of row and column pixels across adjacent blocks. In [5], Perra *et al.* presented a blockiness index, which is a combination of luminance variations between the block boundary pixels and the remaining pixels. In [6], Park *et al.* modeled blockiness using a 2-D linear function, 2-D step function, and Gaussian noise model.

Several blockiness researches have been performed in the spatial and frequency domains [7, 8]. For example, Meesters *et al.* proposed a metric based on the detection of low-amplitude edges computed by Gaussian blurred edges [7]. In [8], edges were modeled as the combination of primary edges, undistorted edges, distorted edges, and blocking artifacts. The blockiness is evaluated in this model using Fourier transformation and least significant bits.

HVS based approaches have also been extensively proposed in order to take into account the response of HVS on perceived blockiness. In [9], Wu *et al.* proposed a metric using the weighted mean squared difference along block boundaries, as well as considering luminance masking. In [10], a blocky image was modeled as a non-blocky image invaded by a pure blocky signal, and luminance and texture masking were considered. In [11], the blocking artifacts were modeled as 2-D step functions, and the perceptual blockiness was estimated in the DCT domain exploiting texture masking and luminance masking. In [12], the luminance masking and the concept of the noticeable blockiness in different

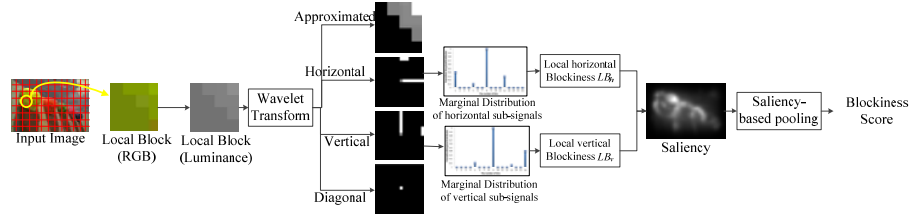


Fig. 1. Block diagram of the proposed blockiness metric.

perceptual regions were employed to measure the visual strength of the blocking artifacts. In [13], Pan *et al.* combined the blockiness and flatness with the local contrast masking and spatial masking.

Those metrics reviewed above are either based on the assumption that grid information, an origin and a size of a block, is known and stationary or do not consider the grid information. However, this assumption is invalid in real-life manipulations such as up/down-scaling and cut and paste operation. Furthermore, in video coding, blocking artifacts can occur at a non-regular grid [14]. In recent researches [15, 16, 17], the grid information is considered into an algorithm. In [15, 16], the size and offset of the grid were detected before computing a blockiness score. However, such approaches that try to determine exact grid information are very sensitive to inaccuracy of a detected grid. In [17], a new approach was proposed based on directions of edges, which does not require the exact location of the block boundary and thus is insensitive to the inaccuracy of the grid-detection compared to the above two methods.

3. PROPOSED BLOCKINESS METRIC

This section describes the proposed no-reference blockiness metric in detail. The proposed metric is based on the marginal distribution of local wavelet coefficients considering characteristics of HVS, i.e., activity masking effect and saliency. Unlike most blockiness metrics, the proposed metric does not require the exact grid information such as origin or size of a block. The approach that employs marginal distribution of local wavelet coefficients is based on our previous work [18]. In this paper, the method is improved considering the concept of activity masking and saliency information.

Fig. 1 illustrates block diagram of the proposed metric. In the method, an image is first divided into $N \times N$ blocks. Note that the size of a block, N , is not restricted to multiple of the size of the grid used for compression, for example, a size of grid is 8 in JPEG. For each block, wavelet coefficients are computed, and then marginal distributions of the wavelet coefficients are constructed for horizontal and vertical sub-signals, respectively. Based on these distributions, local blockiness scores are derived and integrated into an overall blockiness index using saliency-based pooling model. The detailed descriptions are given in following sections. Note that, for the sake of simplicity, we focus on horizontal blockiness, and vertical blockiness can be computed very similarly.

3.1. Marginal distribution of wavelet coefficients

The wavelet transform is an effective way to reveal both spatial and frequency properties of an image. This transform commonly divides the information of an image into approximated and detailed sub-signals. The approximated sub-signal represents a

general trend of pixel values, while the other sub-signals represent detailed features in vertical, horizontal, and diagonal directions.

We observed that the marginal distributions of vertical or horizontal sub-signals are classified into four cases according to existence of blockiness and spatial activity (Fig. 2). In the first case (A), a block has low activity with no blockiness; the marginal distribution presents low-amplitude or no peak-bins with low-amplitude non-peak-bins. In the second case (B), a block has high activity with no blockiness, and the marginal distribution provides high-amplitude peak-bins with high-amplitude non-peak-bins. In the third case (C), a block has low activity and includes a blocking artifact; there are several high-amplitude peak-bins, and the remaining non-peak-bins have low-amplitudes. In the last case (D), a block has high activity and includes a blocking artifact; in this case, both peak- and non-peak-bins have relatively high amplitudes. Among these four cases, the blockiness is only highly perceived by HVS in the third case. Even though there are obviously blocking artifacts in the fourth case, the blocking artifacts are invisible due to the masking effect created by high spatial activity. It is interesting to note that all of the normalized marginal distributions are similar to each other except for the third case. The peak-bins are conspicuously higher than the non-peak-

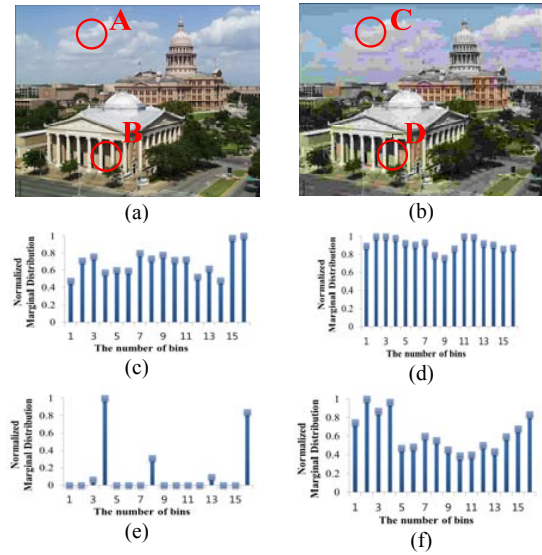


Fig. 2. (a) blockiness-free image, (b) blockiness-corrupted image, (c-f): Normalized marginal distribution of horizontal wavelet coefficient; (c) A: low activity region with no blockiness, (d) B: high activity region with no blockiness, (e) C: low activity region with blockiness, (f) D: high activity region with blockiness.

bins for the third case, while there is no distinct boundary between the peak-bins and the non-peak-bins for the other cases. From this observation, we can estimate the perceptual blockiness exploiting magnitudes of peak-bins and non-peak-bins.

3.2. Local blockiness measure

The horizontal blockiness is estimated based on the marginal distribution of horizontal sub-signal of wavelet transform in a local block. Let us define the i^{th} $N \times N$ block as X_i . The horizontal sub-signal of wavelet transform for X_i is then computed as follows:

$$w_h(u, v) = \sum_{x=0}^N \sum_{y=0}^N \psi_h(x, y) \cdot X_i(2u - x, 2v - y), \quad (1)$$

where w_h represents a horizontal sub-signal, and ψ_h is a horizontal wavelet function composed of two impulse responses of a vertical low-pass filter and a horizontal high-pass filter. The marginal distribution of the horizontal sub-signal is constructed as follows:

$$mw_h(u) = \sum_v |w_h(u, v)|, \quad (2)$$

where $mw_h(u)$ represents the amplitude of the u^{th} bin of the marginal distribution of the horizontal sub-signal. The distribution is then normalized into a range from 0 to 1.

As referred in the previous section, a perceptual blockiness is estimated using a difference between average amplitudes of peak-bins and non-peak-bins. Let us define the peak-bins as $B_p(i)$ and the non-peak-bins as $B_{np}(i)$. We employ a simple thresholding approach for the determination of $B_p(i)$ and $B_{np}(i)$ for simplicity, even though they can be divided by more accurate method. In this paper, 0.7 is empirically used as the threshold value. Accordingly, the horizontal local blockiness LB_h is calculated as follows:

$$LB_h = \begin{cases} 1 & \text{if } mv_h(i) = mv_h(j), \text{ for } i \neq j \\ |mB_p - mB_{np}| & \text{otherwise} \end{cases}, \quad (3)$$

$$mB_p = \frac{1}{N_p} \sum_k B_p(k), \quad mB_{np} = \frac{1}{N_{np}} \sum_k B_{np}(k), \quad (4)$$

where mB_p and mB_{np} are average values of peak-bins and non-peak-bins, respectively. N_p and N_{np} are the numbers of peak-bins and non-peak-bins, respectively. The vertical local blockiness LB_v is calculated with a similar method.

3.3. Saliency-based pooling strategy

The HVS attends to salient regions in an image because the field of view of HVS is very restricted. Given the fact that the artifacts present in attended regions are better perceived by the HVS than artifacts present in non-attended areas, we utilize the saliency information to draw an overall blockiness. The employed method for generating the saliency map is graph-based visual saliency (GBVS) [19], owing to its simplicity and good performance. A bottom-up visual saliency model, GBVS is based on the random walks on the graphs constructed using edge strengths between two nodes. Based on the saliency map, the overall horizontal blockiness B_h and vertical blockiness B_v are computed as follows:

$$B_h = \frac{\sum_i LB_h(i) \cdot S(i)}{\sum_i S(i)}, \quad B_v = \frac{\sum_i LB_v(i) \cdot S(i)}{\sum_i S(i)}, \quad (5)$$

$$S(i) = \sum_x \sum_y S_{xy}(x, y), \quad (6)$$

where $S_{xy}(x, y)$ is the pixel-wise saliency information, and $S(i)$ is an average value of the saliency in the i^{th} local block. M and N are the

number of local blocks and the size of a local block, respectively. Lastly, an overall blockiness of an image is estimated as follows:

$$B = \alpha \cdot B_h + (1 - \alpha) \cdot B_v. \quad (7)$$

Here we assume that the horizontal and vertical blockiness have the same importance, i.e., $\alpha = 0.5$ is used in (7). MATLAB implementation of the proposed metric can be available at <http://diml.yonsei.ac.kr/~sryu/bmws/>.

4. PERFORMANCE EVALUATIONS

To evaluate the performance of the metric, JPEG-compressed images from LIVE [20] and IVC [21] databases, as well as their modified versions, are used. LIVE database includes 233 JPEG-compressed images, including the 29 original color images. The subjective experiments for the LIVE database are conducted using a continuous linear scale. The difference mean opinion score (DMOS) for each image is then calculated from the raw scores.

The IVC database consists of 10 original images and 235 distorted images. The distortion types are JPEG, JPEG 2000, Locally Adaptive Resolution (LAR) coding, and Gaussian blurring. The subjective tests are conducted using a double stimulus impairment scale method (DSIS), in which the reference and distorted images are sequentially displayed. In our evaluation, JPEG-compressed images (50 images) are used.

In addition to the original LIVE and IVC databases, cropped images and resized images are employed to evaluate the robustness to grid distortions. The cropped images are obtained from LIVE and IVC databases with removing 1, 3, and 5 pixels from boundary. The resized images are obtained from being magnified (with a scale of $\times 1.2$) and contracted (with a scale of $\times 0.8$), respectively.

In our experiments, LIVE JPEG-compressed images (233 images), their cropped images ($233 \times 3 = 699$ images), their resized images ($233 \times 2 = 466$ images), IVC JPEG-compressed images (50 images), their cropped images ($50 \times 3 = 150$ images), and their resized images ($50 \times 2 = 100$ images) are used, totaling 1698 JPEG-compressed images. To evaluate the performance of the proposed metric, we follow the suggestions of the VQEG report [22]. A four parameter logistic function, as recommended in [22], is used for non-linear regression before calculating the performance measure, i.e., Pearson correlation coefficient (PCC). Note that, for a well-defined metric, a value of PCC should be high.

Table 1. Performance evaluation of the proposed metric and existing metrics for the LIVE databases

Metric	Pearson correlation coefficient			
	LIVE	Cropped LIVE	Resized LIVE	ALL
Proposed	0.928	0.906	0.914	0.915
M1 [4]	0.684	0.530	0.352	0.378
M2 [5]	0.902	0.883	0.880	0.876
M3 [6]	0.603	0.272	0.238	0.280
M4 [9]	0.740	0.856	0.723	0.500
M5 [10]	0.850	0.872	0.563	0.737
M6 [11]	0.929	0.593	0.904	0.627
M7 [12]	0.894	0.590	0.321	0.399
M8 [13]	0.904	0.766	0.883	0.724
M9 [15]	0.876	0.885	0.809	0.844
M10 [16]	0.724	0.734	0.743	0.736
M11 [17]	0.906	0.887	0.854	0.883
M12 [18]	0.911	0.880	0.870	0.889

Table 2. Performance evaluation of the proposed metric and existing metrics for the IVC databases

Metric	Pearson correlation coefficient			
	IVC	Cropped IVC	Resized IVC	ALL
Proposed	0.825	0.858	0.813	0.807
M1 [4]	0.504	0.291	0.204	0.210
M2 [5]	0.787	0.731	0.711	0.717
M3 [6]	0.539	0.432	0.351	0.240
M4 [9]	0.817	0.380	0.481	0.383
M5 [10]	0.797	0.823	0.172	0.559
M6 [11]	0.791	0.406	0.786	0.466
M7 [12]	0.562	0.188	0.038	0.156
M8 [13]	0.746	0.293	0.759	0.285
M9 [15]	0.806	0.810	0.781	0.759
M10 [16]	0.601	0.595	0.367	0.520
M11 [17]	0.761	0.759	0.763	0.760
M12 [18]	0.787	0.801	0.723	0.762

Tables 1 and 2 summarize the results of the proposed metric compared with twelve existing no-reference blockiness metrics for LIVE and IVC databases, respectively. As shown in Tables 1 and 2, the proposed metric shows good performance for all of the databases. In fact, the proposed metric exhibits the highest correlations for all cases, except for the case of non-modified LIVE images. On the other hand, most existing metrics exhibit low correlations, especially when all the images, including non-modified, cropped, resized images, are mixed (the right-most columns in Tables 1 and 2, respectively). The metrics M6 [11], M7 [12], and M8 [13] showing high correlations for the non-modified LIVE images fail to predict perceived blockiness when they are applied to other databases. In fact, the obtained results show that the metric M7 [12] exhibits very low correlation for IVC databases. Only four metrics (M2 [5], M9 [15], M11 [17], and M12 [18]) among eleven ones show consistent and high correlations for all of the databases. Nevertheless, the proposed metric outperforms these three metrics for all the databases.

5. CONCLUSION AND FUTURE WORKS

In this paper, a novel metric for measuring perceptual blockiness is proposed based on the marginal distribution of local wavelet sub-signals. The local blockiness derived from the relative magnitudes between peak-bins and non-peak-bins are integrated into an overall objective blockiness score through saliency-based pooling method. To evaluate the performance of the proposed metric, we conducted experimental verifications on several public image databases. The results show that the proposed metric can provide consistent and outstanding performance in diverse images with grid-distortions. Furthermore, the proposed metric outperforms eleven existing no-reference blockiness metrics.

Future research will focus on enhancing the metric to be robust to more general grid-distortions. For example, the proposed metric cannot be applied when the grid is rotated in different directions. Other directions of future research include extending blockiness metric into videos considering blockiness fluctuation.

6. REFERENCES

- [1] Z. Wang and A.C. Bovik, *Modern Image Quality Assessment*, Morgan & Claypool, USA, 2006.
- [2] P. Marziliano, F. Dufaux, S. Winkler, and T. Ebrahimi, "Perceptual blur and ringing metrics: Applications to JPEG2000," *Signal Process.-Image Commun.*, vol. 19, no. 2, pp. 163-172, Feb. 2004.
- [3] S. Ryu and K. Sohn, "No-reference sharpness metric based on inherent sharpness," *IET Electron. Lett.*, vol. 47, no. 21, pp. 1178-1180, Oct. 2011.
- [4] J. Yang, H. Choi, and T. Kim, "Noise estimation for blocking artifacts reduction in DCT coded images," *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 10, no. 7, pp. 1116-1120, Oct. 2000.
- [5] C. Perra, F. Massidda, and D. Giusto, "Image blockiness evaluation based on Sobel operator," *Proc. IEEE ICIP*, Genoa, Italy, Sep. 2005.
- [6] C. Park, J. Kim, and S. Ko, "Fast blind measurement of blocking artifacts in both pixel and DCT domains," *J. Math. Imaging Vis.*, vol. 28, no. 3, pp. 279-284, Jul. 2007.
- [7] L. Meesters and J. Martens, "A single-ended blockiness measure for JPEG-coded images," *Signal Process.*, vol. 82, no. 3, pp. 369-387, Mar. 2002.
- [8] S. Suthaharan, "No-reference visually significant blocking artifact metric for natural scene images," *Signal Process.*, vol. 89, no. 8, pp. 1647-1652, Aug. 2009.
- [9] H. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," *IEEE Signal Process. Lett.*, vol. 4, no. 11, pp. 317-320, Nov. 1997.
- [10] Z. Wang, A.C. Bovik, and B. Evan, "Blind measurement of blocking artifacts in images," *Proc. ICIP*, Vancouver, Canada, Sep. 2000.
- [11] S. Liu and A.C. Bovik, "Efficient DCT-domain blind measurement and reduction of blocking artifacts," *IEEE Trans. Circuits and Syst. Video Technol.*, vol. 12, no. 12, pp. 1139-1149, Dec. 2002.
- [12] S. Suthaharan, "A perceptually significant block-edge impairment metric for digital video coding," *Proc. IEEE ICASSP*, Hong Kong, Apr. 2003.
- [13] F. Pan, X. Lin, S. Rahardja, W. Lin, E. Ong, S. Yao, Z. Lu, and X. Yang, "A locally adaptive algorithm for measuring blocking artifacts in images and videos," *Signal Process.-Image Commun.*, vol. 19, no. 6, pp. 499-506, Jul. 2004.
- [14] A. Leontaris, P.C. Cosman, and A.R. Reibman, "Quality evaluation of motion-compensated edge artifacts in compressed video," *IEEE Trans. Image Process.*, vol. 16, no. 4, Apr. 2007.
- [15] I. Kirenko, and R. Muijs, "Coding artifact reduction using non-reference block grid visibility measure," *Proc. IEEE ICME*, Toronto, Ont. Canada, Jul. 2006.
- [16] H. Liu and I. Heynderickx, "A perceptually relevant no-reference blockiness metric based on local image characteristics," *EURASIP J. Adv. Signal Process.*, vol. 2009, no. 263540, pp. 1-14, Jan. 2009.

- [17] F. Pan, X. Lin, S. Rahardja, E. Ong, and W. Lin, "Using edge direction information for measuring blocking artifacts of images," *Multidimens. Syst. Signal Process.*, vol. 18, no. 4, pp. 297-308, Dec. 2007.
- [18] S. Ryu and K. Sohn, "Wavelet-based no-reference blocking artifact measure," *Proc. IVCNZ*, Auckland, New Zealand, Nov. 2011.
- [19] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," *Adv. Neural Inform. Process. Syst.*, vol. 19, pp. 525-552, 2007.
- [20] H.R. Sheikh, Z. Wang, L. Cormack, and A.C. Bovik, "LIVE image quality assessment database release 2," 2006, <http://live.ece.utexas.edu/research/quality>.
- [21] P.L. Callet, and F. Autrusseau, "Subjective quality assessment IRC-CyN/IVC databases," 2005, <http://www2.irccyn.ec-nantes.fr/ivcdb/>.
- [22] VQEG, "Final report from the Video Quality Experts Group on the validation of objective models of video quality assessment Phase II," Aug. 2003.