A PERCEPTUALLY SIGNIFICANT BLOCK-EDGE IMPAIRMENT METRIC FOR DIGITAL VIDEO CODING

Shan Suthaharan
Department of Mathematical Sciences
University of North Carolina at Greensboro
Greensboro, NC 27402, USA
ssuthaharan@uncg.edu

ABSTRACT

A new perceptually significant block-edge impairment metric (PS-BIM) is presented in this paper as a quantitative distortion measure to evaluate blocking artifacts in block-based video coding. This distortion measure does not require the original video sequence as a comparative reference and is found to be consistent with subjective evaluation.

1. INTRODUCTION

The blocking artifacts and its propagation through reconstructed video sequences are the most significant of all coding artifacts, especially for bit rates ranging from 64kbps (ITU H.261) to as high as 10-12 Mbps (MPEG-2) [1,2]. It is a well known fact that the quantitative quality measure or distortion measure such as mean squared error (MSE), peak signal-to-noise ratio (PSNR) and mean absolute error (MAE) [1] are not suitable for quantifying the visual strength of the coding artifacts in digital video coding.

However, very minimal research has been done to determine a perceptual distortion measure for compressed video to quantify the visual strength of the blocking artifacts. In [3], a new distortion measure for blocking (edge) artifacts in compressed images based on human visual sensitivity has been proposed. Although the approach shows the significant correlation with the subjective evaluation, it requires both the original and compressed images to form an error image and it is impossible to have the original image at the receiver, especially in the remote site. Similarly, in [4], a spatio-temporal model of the human visual system (HVS) has been proposed for image restoration and quality assessment applications. However, there is a limitation in this approach and that is the requirement for the original image or the image sequence and it is impossible to have the original entire video sequence.

Recently in [5], a generalized block-edge impairment metric (GBIM) has been proposed to evaluate the visual significance of block-edge artifacts in a given image. The GBIM is based on a formulation of constraint sets applied successfully in the post-filtering of compressed images using Projections On to Convex Sets (POCS) algorithm [6]. It is a spatial domain technique and uses luminance masking in extreme bright and dark regions as well as in spatially busy areas. The advantage of this technique compared to [3] is that it does not require the information contained in the original image. However, the technique only provides the statistical strength rather than perceptual strength.

This paper proposes a spatial domain technique called PS-BIM that uses both luminance masking and block noticeable differences (BND) in different perceptual regions known as the Dark Region (or Low Intensity Region), De Vries-Rose Region, Weber Region and Saturation Region [7,8] to measure the visual strength of the blocking artifacts quantitatively.

2. A BLOCK-EDGE IMPAIRMENT METRIC

The perceptually significant block-edge impairment metric (PS-BIM) for measuring the strength of blocking artifacts in a reconstructed video sequence is defined by

$$M_{PS\text{-}BIM} = D_1 / D_2 \tag{1}$$

It measures the blocking artifacts per pixel and the higher the $M_{\text{PS-BIM}}$ is above 1, the greater the severity of the blocking artifacts. The symbols D_1 and D_2 define two types of block noticeable differences along the block boundaries as follows:

$$D_1 = \alpha D_{h1} + \beta D_{v1}$$

$$D_2 = \alpha D_{h2} + \beta D_{v2} \tag{2}$$

Where the terms on the right hand sides are defined by

$$D_{h1} = \sum_{i=8}^{N-8(step8)} \sum_{j=1}^{M} w(I_{ij}).abs(I_{ij} - \sum_{k=-1}^{1}\sum_{l=-1}^{1} I_{(i+k)(j+l)} / 8)$$

$$D_{v1} = \sum_{j=8}^{N-8(step8)} \sum_{i=1}^{M} w(I_{ij}).abs(I_{ij} - \sum_{k=-1}^{1}\sum_{l=-1}^{1} I_{(i+k)(j+l)} / 8)$$

$$D_{h2} = \sum_{i=8}^{N-8(step8)} \sum_{j=1}^{M} w(I_{ij}).abs(I_{ij} - I_{(i+1)j})$$

$$D_{v2} = \sum_{j=8}^{N-8(step8)} \sum_{i=1}^{M} w(I_{ij}).abs(I_{ij} - I_{i(j+1)}) \tag{3}$$

(where $k \neq 0; l \neq 0$). The values for the parameters α and β may be chosen based on the assumption that the human sensitivity to horizontal and vertical blocking artifacts is similar or not. Where N x M is the size of a video frame and $w(I)$ is the weighting function (see Fig. 1) that is derived using the intensity regions: Dark Region, De Vries-Rose Region, Weber Region and

Saturation Region for the pixel with intensity *I*. This weighting function *w*(*I*) is defined by the following piece-wise function and it is shown in Fig. 1:

$$w(I) = \begin{cases} 1.284, & 31 \geq I \geq 0 \\ -0.433 + 0.5*\log(I), & 81 \geq I > 31 \\ 6.158 - \log(I), & 229 \geq I > 81 \\ 11.592 - 2*\log(I), & 255 \geq I > 229 \end{cases} \quad (4)$$

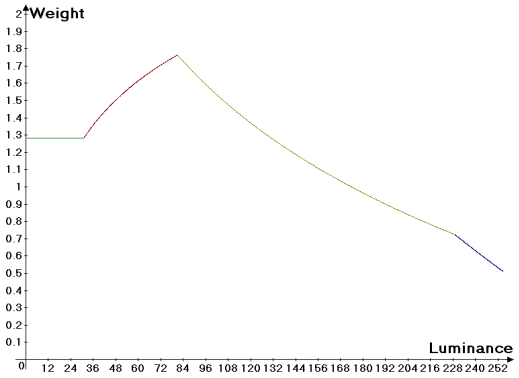


Fig.1 : The weighting function for the PS-BIM

It is interesting to note that the peak of this piece-wise function appears at intensity value 81 and it agrees with the claim in [9] that block distortions are best noticeable when the luminance values lie between 70 and 90, centered approximately at 81for 8-bit gray-scale images. The following section shows the derivation of the weighting function in equation (4).

3. DERIVATION OF WEIGHTING FUNCTION

In [7], the author has defined the just noticeable difference (JND) as the amount of light that needs to be added so that the intensity of a pixel can be discriminated from the background intensity**.** In this paper we define the block noticeable difference (BND) as the amount of light added (or subtracted from) to the intensity of the pixels along the block boundary by the compression schemes, which causes the blocking artifacts along the block boundaries. It is clear from [8], the intensity values (I) can be partitioned into 4 regions namely Dark Region, De Vries-Rose Region, Weber Region and Saturation Region based on the block noticeable difference of I, which is Δ(I), as follows:

$$\Delta(I) = \begin{cases} k, & \text{Dark region} \\ k_1*\sqrt{I}, & \text{DeVries-Rose region} \\ k_2*(1/I), & \text{Weber region} \\ k_3*(1/I^2), & \text{Saturation region} \end{cases} \quad (5)$$

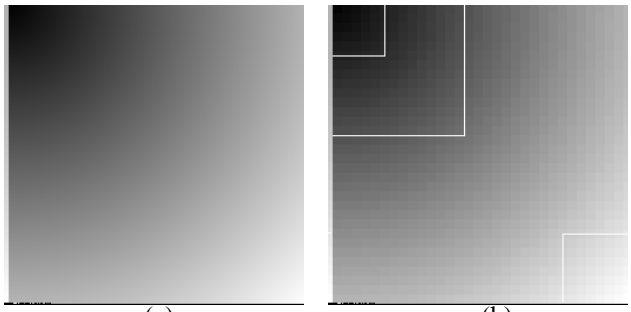
Based on this, our weighting function *w(I)* is chosen as the following logarithmic piece-wise function:

$$w(I) = \begin{cases} \log(k) & \text{Dark region} \\ \log(k_1)+0.5*\log(I) & \text{DeVries-Rose region} \\ \log(k_2) - \log(I) & \text{Weber region} \\ \log(k_3) - 2*\log(I) & \text{Saturation region} \end{cases} \quad (6)$$

In the following two subsections, methods are suggested to determine the 4-region boundaries and suitable values for the constants `k, k1, k2` and `k3`.

3.1 Calculation of Region Boundaries

We suggest a simple approach to determine the region boundaries. Consider the template image (Fig. 2(a)) that has continuous intensity changes from dark (0) to bright (255) from top-left hand corner to bottom-right hand corner on the 2D-image plane. This is a suitable image to identify the strength of blocking artifacts perceptually at different intensity levels within the 4-regions. We compress this image using an intra frame coding; an example using MPEG2 I-frame compression is shown in Fig. 2(b).



(a) (b)

Fig. 2: (a) The template image. (b) MPEG2 I frame compressed version of (a)

It is clear from Fig. 2(b) that the blocking artifacts cannot be seen between intensity values 0 and 31 (approximately) and it follows the constant behavior of HVS property in dark region. Between 31 and 81 (approximately) the visibility of blocking artifacts increases and it can be assumed that it follows the DeVries-Rose law. Between 81 and 229 (approximately), the visibility of the blocking artifact decreases and it can be assumed that it follows inverse Weber's law. Then the blocking artifact decreases deeply in the saturation area with the inverse quadratic law (above 229).

In summary, we say that the intensity regions [0, 31], (31, 81], (81, 229] and (229, 255] form Dark region, De Vries-Rose region, Weber region and Saturation region respectively, and thus we can rewrite eq.(6) in terms of intensity as follows:

$$w(I) = \begin{cases} \log(k) & 31 \geq I \geq 0 \\ \log(k_1)+0.5*\log(I) & 81 \geq I > 31 \\ \log(k_2) - \log(I) & 229 \geq I > 81 \\ \log(k_3) - 2*\log(I) & 255 \geq I > 229 \end{cases} \quad (7)$$

Note that this is a piece-wise function and it is continuous within the 4-regions and intersects at region boundaries. Therefore, the constants `k, k1, k2` and `k3` can be calculated using the equality conditions at the boundaries as discussed in the following section.

3.2 Calculations of k, k_1, k_2 and k_3

In this section, a method is presented to determine suitable values for the constants k, k_1, k_2 and k_3, in equation (7) using the block noticeable difference derived from the template image in Fig. 2(b). It is clear from equation (7) that the dark region and De Vries-Rose region intersect at I=31, thus

$$\log(k) = \log(k_1) + 0.5*\log(31)$$

It gives $k = 5.568*k_1$.

(8)

Similarly, the De Vries-Rose region and Weber region intersect at I=81 and the Weber region and Saturation region intersect at I=229. Therefore,

$$\log(k_1) + 0.5*\log(81) = \log(k_2) - \log(81)$$

$$\log(k_2) - \log(229) = \log(k_3) - 2*\log(229)$$

By solving these equations, we obtain

$$k_2 = 729*k_1;\ k_3 = 729*229*k_1,$$

(9)

It is clear that these constants depend on one another and finding a value for one of these constants lead to finding the values for the others. In order to find these constants, we carry out the following procedure; using the template image and block noticeable difference, to determine the value for the constant k_1 which falls in the De Vries-Rose region. Note that the blocking artifacts are significantly noticeable in the pixel area where the intensity values are around 81, and so we first select a block, say B, in this pixel region such that it falls inside the De Vries-Rose region.

Suppose m_1 is the intensity value of a pixel along the block boundary of B, but inside the block and m_2 is the intensity value of the closest neighboring pixel lies outside the block. We can simply assume that the difference m_2–m_1 makes the block noticeable along the edge and it is the block noticeable difference value of the intensity m_1 and thus

$$\log(m_2 - m_1) = \log(k_1) + (1/2)*\log(m_1).$$

(10)

Using the template image in Fig. 2(b), we find the exact intensity values m_1=77.56 and m_2=83.27. Using these intensity values in the above equation, we obtain

$$\log(5.71) = \log(k1) + (1/2)*\log(77.56)$$

It gives $k_1 = 0.6484$. Therefore, using equations (8) and (9), we have:

$$k=3.61;\ k_2=472.68;\ k_3=108244.54$$

(11)

Therefore, equation (7) becomes:

$$w(I) = \begin{cases} 1.284, & 31 \geq I \geq 0 \\ -0.433 + 0.5*\log(I), & 81 \geq I > 31 \\ 6.158 - \log(I), & 229 \geq I > 81 \\ 11.592 - 2*\log(I), & 255 \geq I > 229 \end{cases}$$

(12)

4. EXPERIMENTAL RESULTS

In our experiments we used several MPEG coded video sequences (only the reference frames are considered) and assumed that the human visual sensitivity to horizontal and vertical blocking artifacts are similar [3]. Therefore we selected $\alpha = \beta = 0.5$. The subjective evaluation was conducted using Sun Microsystems Ultra 60. A reconstructed I-frame of the flower garden video sequence (MPEG coded at 0.358bpp) is shown with its PSNR and $M_{PS\text{-}BIM}$ values in Fig. 3(a). It is then filtered using POCS [6], WLS [10] and IWF [11] techniques and the filtered video frame of Fig. 3(a) using WLS is shown in Fig. 3(b) with its PSNR and M_{PSBIM} values. These values indicate that the video quality is improved and blocking artifacts are reduced.

Further experiments show that the WLS reduces the blocking artifacts better than the POCS filtering and also shows that our proposed metric $M_{PS\text{-}BIM}$ is very effective measure for measuring blocking artifacts and consistency with subjective evaluation.

Table-I provides a compilation of PSNR and $M_{PS\text{-}BIM}$ values for several other MPEG coded video sequences. For example, it shows that WLS improves the visual quality (or reducing blocking artifacts) better than POCS by 0.0042 units/pixel for flower garden video frame. There are 12250 pixels used and thus it gives us the blocking artifacts reduction of 51.45 units/frame.

(a) (b)

Fig.3: (a) MPEG coded flower garden video frame at 0.358 bpp, its PSNR=23.14 and M_{PSBIM}=0.8050 units/pixel; (b) WLS filtered video frame and its PSNR=24.11 and M_{PSBIM}=0.7768 units/pixel.

We also investigated the effectiveness of our measure at different bit rates. The graphs in Fig.4 and Fig.5 show the relationship between D_1 and D_2 and PS-BIM values of flower garden video with respect to different bit rates. Our finding is that the visual quality is unacceptable (i.e. severe blocking artifacts present) at bit rates below the bit rate at the intersection between the two curves of D_1 and D_2 and the PS-BIM value is greater than 1 for these unacceptable video quality (see Fig. 5). Our research also indicates that the quality of video at these bit rates cannot be improved to a required level of visual quality by using the block filtering techniques such as POCS, WLS and IWF.

Table I: Comparison of PSNR and PS-BIM

MPEG Encoded Images	PSNR	PS-BIM	Improvements of WLS over POCS	
			PSNR	PS-BIM
Flower Garden 0.358bpp	23.14	0.8050 units/pixel	0.967	0.0042 units/pixel
Football Image 0.233bpp	26.48	0.9351 units/pixel	0.977	0.0046 units/pixel
Calendar Image 0.339bpp	22.42	0.8523 units/pixel	0.878	0.0037 units/pixel

Fig.5 also shows that the PS-BIM increases sharply from 1.0086 units/pixel as bit rate decreases below 0.1893bpp. It can also be seen that the PS-BIM decreases slowly from 1.0086 units/pixel as bit rate increases above 0.1893 and it indicates that the blocking artifacts presented in these bit rate can be removed by the algorithms like POCS, WLS and IWF.

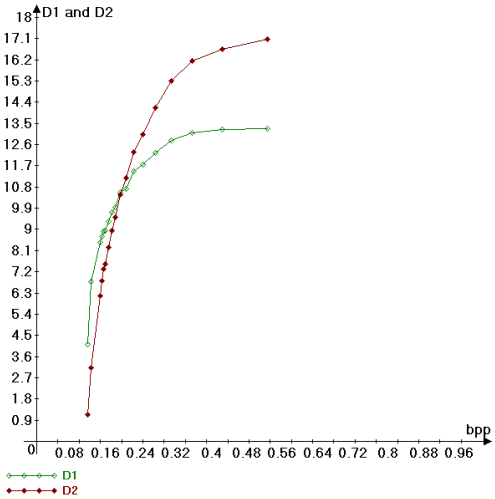


Fig. 4: Plot of PS-BIM1 and PS-BIM2 at different bit rates

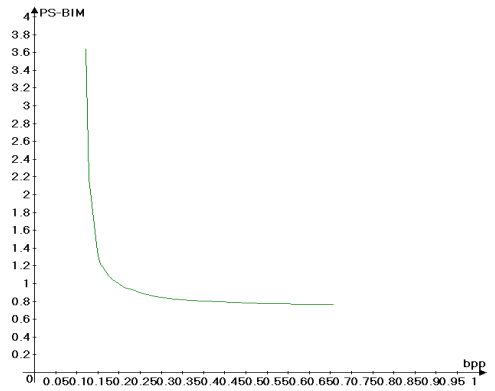


Fig. 5: Plot of PS-BIM against different bit rates

5. CONCLUSION

A new perceptually significant block-edge impairment metric that is suitable for compressed video has been proposed in this paper. It uses contrast masking and block noticeable differences to enhance the accuracy of the measure with the subjective evaluation. The advantage of this metric is that it uses only the information contained in the compressed video and provides consistent results with the subjective evaluations. Object motion and related perceptual characteristics will be included in the future development of PS-BIM.

6. REFERENCES

[1] R. J. Clarke, *Digital Compression of Still Images and Video*, New York: Academic, 1995.

[2] K. R. Rao and J. J. Hwang, *Techniques and Standards for Image, Video, and Audio Coding*, Prentice Hall, 1996.

[3] S. A. Karunasekera and N. G. Kingsbury, "A distortion measure for blocking artifacts in images based on human visual sensitivity," IEEE Trans. Image Processing, vol. 4, pp. 713-724, June 1995.

[4] Christian J. van den Branden Lambrecht, "A working spatio-temporal model of the human visual system for image restoration and quality assessment applications," in Proc. ICASSP-96, May 7-10, Atlanta, GA.

[5] H. R. Wu and M. Yuen, "A generalized block-edge impairment metric for video coding," IEEE Signal Processing Letters, Vo. 4, No. 11, pp. 317-320, November 1997.

[6] Y. Yang, N. P. Galatsanos and A. K. Katsaggelos, "Projection-based spatially-adaptive reconstruction of block-transform compressed images," in Proc. SPIE Conf. Visual Communications and Image Processing, Sept. 1994, vol. 2308, pp. 1477-1488.

[7] G. Buchsbaum, "An analytical derivation of visual nonlinearity," IEEE Trans. on Biomedical Engineering, vol. 27, pp. 237-242, 1980.

[8] M. K. Kundu and S. K. Pal, "Thresholding for edge detection using human psychovisual phenomena," Pattern Recognition Letters 4, pp. 433-441, 1986.

[9] B. Girod, "The information theoretical significance of spatial and temporal masking in video signals," in Proc. SPIE Conf. Human Vision, Visual Processing and Digital Display, 1989, vol. 1077, pp. 178-187.

[10] S.S.O. Choy, Y.-H.Chan and W.-C.Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients," IEEE Signal Processing Lett., vol. 4, no. 1, Jan. 1997.

[11] S. Suthaharan, S.W. Kim, H. R. Wu and K.R. Rao, "Transform domain techniques for multimedia image and video coding," Multimedia Image and Video processing, Ed. By L. Guan, S.-Y Kung and J. Larsen, CRC Press, ISBN: 0-8493-3492-6, pp. 363-381, 2000.