

# MODELING BLOCKING VISUAL SENSITIVITY PROFILE

Guangtao Zhai, Wenjun Zhang, Xiaokang Yang, Yi Xu  
Institute of Image Communication and Information Processing  
Shanghai Jiao Tong University, Shanghai 200240, China

## ABSTRACT

Blocking artifact is the most prevailing degradation caused by block-based DCT coding techniques under low bit-rate conditions. To alleviate blockings perceptually, it is desirable to measure the visibility of blocking artifacts. In this paper, we propose an efficient method of estimating the visual sensitivity of blocking artifacts in block-based DCT coding. The differences on block boundaries are measured and transformed into block discontinuity map. We consider the effects of luminance adaptation and texture masking on the blockings and integrate them using nonlinear operator to form an overall masking map. This masking map is then incorporated with the discontinuity map to generate the Blocking Visual Sensitivity Map (BVSM). This map can be used to guide perceptual quality assessment, codec parameter optimization, post-processing, etc. We demonstrate the validity of the BVSM through its application in image quality assessment.

## 1. INTRODUCTION

Many broadly used image/video coding methods are based on block-wise DCT (e.g. JPEG, MPEG). When the bit rate is low, coarse quantization on the DCT coefficients brings some artificial discontinuities across the block boundaries on the decoding end, which is known as blocking effect. Some post-processing schemes are often introduced to supply better visual qualities of the decoded image/video, and that is called “deblocking”. While the deblocking technique has been thoroughly studied through the years [1], we want to take a look at the “blockings” from a psychophysical point of view, and predict the perceptibility of the blocking artifacts by the Human Visual System (HVS).

Abounding research results from psychophysical studies have been assimilated by visual compression techniques and have fundamentally improved the capability of visual communication systems. The influences of spatial masking phenomena (e.g. Contrast Sensitivity Function (CSF), luminance adaptation, and contrast masking) on the HVS have been well simulated by perceptual models (e.g.

Just Noticeable Distortion (JND) profiles [2]). And these perceptual-plausible methods show excellent performance in many applications. However, much attention of engineers in this field goes into estimating the masking effects on a whole-image level and they often compute perceptual models for every pixel in the image [3-5]. Rather than that, in this paper, we emphasize on modeling the masking effects on blocking artifacts and estimate their visual sensitivity.

In this paper, the detailed computation steps of BVSM are given in Section 2; the application of BVSM in no reference image quality assessment is introduced in Section 3; and Section 4 concludes this paper.

## 2. BLOCKING VISUAL SENSITIVITY MAP ESTIMATION

The flow chart of the computation of BVSM is illustrated in Fig. 1, where the output of each step is also presented through an example of “Lena”.

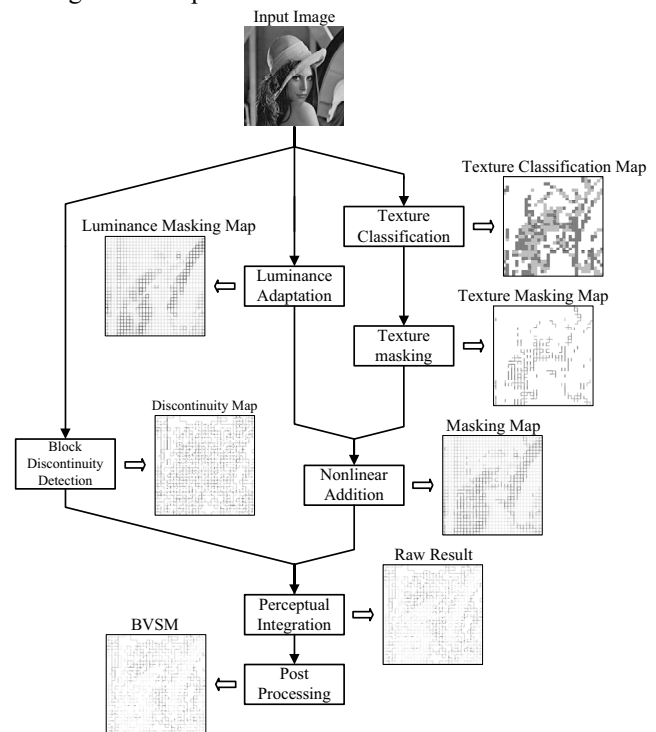


Figure 1. Diagram of estimating the BVSM

This work was supported by National Natural Science Foundation of China under Grant No. 60332030, No.60502034; and Shanghai Rising-Star Program under Grant No. 05QMX1435

## 2.1 Block discontinuity

Let  $f_{a,b}(x,y)$  and  $g_{a,b}(u,v)$  be the original image block and its DCT coefficients block located in  $(a,b)$ . The decoded image block and DCT coefficients block are denoted as  $\hat{f}_{a,b}(x,y)$  and  $\hat{g}_{a,b}(u,v)$ , respectively. The disparities between consecutive blocks in the vertical and horizontal direction are defined as

$$D_{(a,b)|(a,b+1)}^{\text{ver}}(x) = \hat{f}_{a,b+1}(x,0) - \hat{f}_{a,b}(x,N-1), 0 \leq x \leq M, (1)$$

$$D_{(a+1,b)|(a,b)}^{\text{hor}}(y) = \hat{f}_{a,b}(0,y) - \hat{f}_{a,b}(M-1,y), 0 \leq y \leq N, (2)$$

where  $(M,N)$  is the size of the blocks. We can also compute the disparities in the DCT domain by (3) where we take the vertical disparity as an example.

$$D_{(a,b)|(a,b+1)}^{\text{ver}}(x) = \alpha_u \alpha_v \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} [\hat{g}_{a,b+1}(u,v) - \hat{g}_{a,b}(u,v)(-1)^v] \cdot \cos \frac{(2x+1)u\pi}{2M} \cos \frac{v\pi}{2N}, (3)$$

$$\text{with } \alpha_u = \begin{cases} \sqrt{1/M} & u=0 \\ \sqrt{2/M} & 1 \leq u \leq M-1 \end{cases}, \alpha_v = \begin{cases} \sqrt{1/N} & v=0 \\ \sqrt{2/N} & 1 \leq v \leq N-1 \end{cases}.$$

Since the DC component is usually much larger in magnitude than AC component in the DCT coefficients matrix, we only keep the parts with  $u=v=0$  in (3). Since the typical block size in JPEG and MPEG is 8 by 8, we set  $M=N=8$ . We have

$$\begin{aligned} D_{(a,b)|(a,b+1)}^{\text{ver}}(x) &\approx \frac{1}{8} [\hat{g}_{a,b+1}(0,0) - \hat{g}_{a,b}(0,0)] \\ &= \frac{1}{8} [Q[g_{a,b+1}(0,0)] - Q[g_{a,b}(0,0)]], 1 \leq x \leq 8 \end{aligned}, (4)$$

where  $Q[\bullet]$  denotes the quantization step. From (4) we can find that the disparities between consecutive blocks are constrained by the quantization step size for DC component in the quantization table. Liew [6] showed that the largest step discontinuity caused by quantization for a ramp signal is bounded by twice the DC quantization step size. With consideration of the affects of the AC coefficients we have omitted in (4), Liew [6] suggested that there exist block discontinuities if

$$D_{(a,b)|(a,b+1)}^{\text{ver}}(x) \text{ (or } D_{(a,b)|(a,b+1)}^{\text{hor}}(y)) \in [0.5\Delta, 2.5\Delta], (5)$$

where  $\Delta$  is the DC quantization step size. Now let us combine all the pixels satisfying (5) in  $D_{(a,b)|(a,b+1)}^{\text{ver}}(x)$  and  $D_{(a,b)|(a,b+1)}^{\text{hor}}(y)$  and denote them as  $D^{\text{ver}}(x,y)$  and  $D^{\text{hor}}(x,y)$  for clearer expression. The final block discontinuity map is generated by combining disparities in two directions

$$D(x,y) = D^{\text{ver}}(x,y) \cup D^{\text{hor}}(x,y). (6)$$

## 2.2 Texture masking in blocks

We first categorize the blocks according to the properties for its DCT coefficients and then estimate the masking effects on the aforementioned block discontinuities.

### 2.2.1 Block Classification

The blocks are labeled as ‘smooth’ or ‘texture’ according to the number of its nonzero DCT coefficients. And for the ‘texture’ blocks, we then estimate the direction of the texture by the coordinates of those coefficients.

Assume that there are  $N_{a,b}$  nonzero coefficients, located in  $(u_1, v_1), (u_2, v_2), \dots, (u_{N_{a,b}}, v_{N_{a,b}})$  for a DCT block  $\hat{g}_{a,b}(u,v)$ .

Let us define the sum of the coordinates in *horizontal* and *vertical* direction as

$$U_{a,b} = \sum_{i=1}^n u_i, V_{a,b} = \sum_{i=1}^n v_i. (7)$$

We label each block as either ‘smooth’ or ‘texture’ according to the number of its nonzero DCT coefficients as follows:

$$BC_{(a,b)} = \begin{cases} \text{'texture'}, & \text{if } N_{a,b} > R_1 \text{ or } U_{a,b} + V_{a,b} > R_2 \\ \text{'smooth'}, & \text{otherwise} \end{cases} (8)$$

For the ‘texture’ one, we estimate the dominant direction of the texture by the coordinates of those coefficients:

$$DC_{(a,b)} = \begin{cases} \text{'horizontal'}, & \text{if } U_{a,b} / V_{a,b} > R_3 \\ \text{'vertical'}, & \text{if } U_{a,b} / V_{a,b} < 1 / R_3 \\ \text{'oblique'}, & \text{otherwise} \end{cases}, (9)$$

where  $R_1, R_2$  and  $R_3$  are some experimentally generated thresholds and we find  $R_1 = 6, R_2 = 10, R_3 = 1.3$  satisfy most images.

### 2.2.2 Masking Effect Estimation

We first define the abovementioned ‘horizontal’ and ‘vertical’ blocks as ‘directional’ blocks. The texture masking effect between different blocks can then be estimated as

$$TM(a_0, b_0) | (a_1, b_1) = \begin{cases} T_1 & \text{if } (a_0, b_0), (a_1, b_1) \\ & \text{are 'un-directional' blocks} \\ T_2 & \text{if } (a_0, b_0), (a_1, b_1) \text{ are one 'un-directional'} \\ & \text{and one 'directional' blocks} \\ T_3 & \text{if } (a_0, b_0), (a_1, b_1) \text{ are 'directional'} \\ & \text{blocks of the same direction} \\ 0 & \text{otherwise} \end{cases} (10)$$

$T_1, T_2$  and  $T_3$  are some constants, and we set  $T_1 = 5, T_2 = 8, T_3 = 10$  in this paper.

## 2.3 Luminance masking in blocks

We first conduct an experiment to measure the visibility threshold of the blocking under different luminance conditions and then use it to estimate the luminance masking effect on the blocking.

### 2.3.1 Luminance Adaptation of Blocking

The measured just visible difference between blocks under certain luminance conditions is shown in Fig. 2. As can be found that the blocking visibility threshold in middle luminance range is low, which means that the HVS is more sensitive. Similar result has been reported for pixels in [3]. We use a piecewise parabola to approximate the measured thresholds

$$LUM(l) = \begin{cases} \lambda_1 \left(1 - \frac{l}{128}\right)^{\gamma_1} + 2 & \text{if } 0 \leq l < 128 \\ \lambda_2 \left(\frac{l}{128} - 1\right)^{\gamma_2} + 2 & \text{if } 128 \leq l < 255 \end{cases}, \quad (11)$$

where  $l$  is the luminance difference, and here we set  $\lambda_1 = 16, \lambda_2 = 11, \gamma_1 = 3, \gamma_2 = 2$ .

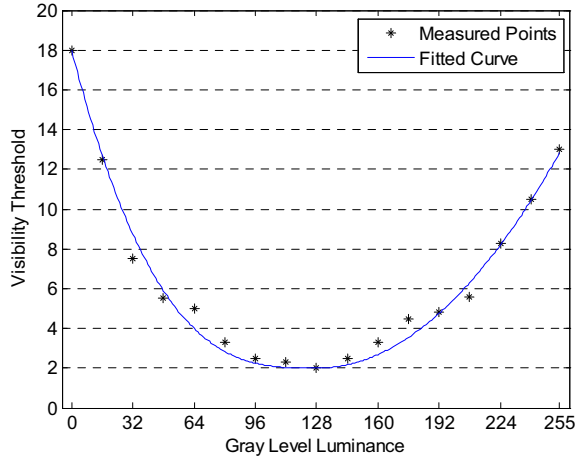


Figure 2. Blocking visibility threshold measurement and fitting

### 2.3.2 Masking Effect Estimation

With the fitted visual sensitivity curve, we now can define the luminance masking as

$$LM(a_0, b_0)|(a_1, b_1) = LUM\left(\frac{1}{8} \cdot \min\{\hat{g}_{a_0, b_0}(0,0), \hat{g}_{a_1, b_1}(0,0)\}\right), \quad (12)$$

where  $\hat{g}_{a_0, b_0}(0,0)$  and  $\hat{g}_{a_1, b_1}(0,0)$  are the DC coefficients of the neighboring blocks, which divided by 8 are equivalent to the average luminance of those blocks.

### 2.4 Nonlinear addition of masking effects

To overcome the possible overlap between texture masking and luminance masking, we use a nonlinear addition model [7] to combine these two masking effects

$$M(a_0, b_0)|(a_1, b_1) = TM(a_0, b_0)|(a_1, b_1) + LM(a_0, b_0)|(a_1, b_1) - \xi \cdot \min\{TM(a_0, b_0)|(a_1, b_1), LM(a_0, b_0)|(a_1, b_1)\} \quad (13)$$

$\xi (\in (0,1))$  is a gain reduction factor, and we set  $\xi = 0.3$  in this paper.

### 2.5 Perceptual Integration and post-processing

Hereinafter, like the depiction for block discontinuity map, we denote  $M(a_0, b_0)|(a_1, b_1)$  as  $M(x, y)$  for convenient description. The blocking visual sensitivity of the image can now be defined as

$$BVSM(x, y) = \begin{cases} D(x, y) \cdot [M(x, y)]^{-1} & \text{if } D(x, y) \cdot M(x, y) \neq 0 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

Note that the values in masking map, discontinuity map and BVSM should actually being at some ‘half-pixel’ positions. The resultant blocking visual sensitivity map usually has some isolated points and small regions, we can impose a morphology-based post-processing step to clear them and generate the final BVSM.

### 3. APPLICATION IN IMAGE QUALITY ASSESSMENT

The BVSM can be easily pooled to generate an impairment score of the BDCT coded image, which is named as mean blocking visual sensitivity (MBVS)

$$MBVS = \frac{1}{M \cdot N} \sum_{x=1}^M \sum_{y=1}^N [BVSM(x, y)]^{\zeta} \quad (15)$$

Where  $M$  and  $N$  are image dimensions. And we find  $\zeta = 0.3 \sim 0.5$  often gives the best pooling results.

We now compare the performance of MBVS with those of other two well-known no-reference image quality metrics for BDCT coded images, namely Wu’s mean generalized blocking impairment metric (MGBIM) [8] and Wang’s JPEG-Quality-Score (JQS) [9]. The experiment was conducted on the JPEG dataset in the LIVE image database release 2 [10], which consists of 233 images with their Difference Mean Opinion Score (DMOS). We evaluated the quantitative performance of the proposed metrics using methods introduced by Video Quality Experts Group (VQEG) [11]. Those metrics are correlation after nonlinear regression, and Spearman rank-order correlation between objective and subjective scores, which evaluate the prediction accuracy and the prediction monotonicity respectively.

Fig. 3 shows the scatter plots of the prediction of different models versus DMOS with 95% confidence intervals. (The final MGBIMs are multiplied by 100 for better fitting results.) It is found that the proposed MBVS creates fewer outliers as compared to MGBIM and JQS. The quantitative performances listed in table 1 also justify that MBVS outperforms MGBIM and JQS in both prediction accuracy and monotonicity.

### 4. CONCLUSION

An effective method of evaluating blocking visual sensitivity has been proposed. We integrate the luminance adaptation and the texture masking effects of the HVS with

a nonlinear addition model and incorporate it into block discontinuity detections to generate the final BVSM. The performance of the proposed BVSM has been demonstrated by its application in no-reference image quality assessment. It can be also used to codec parameter optimization and post-processing tasks in visual communication systems, which is of course our future work.

TABLE I. PERFORMANCE OF IMAGE QUALITY ASSESSMENT MODELS  
CC: CORRELATION COEFFICIENTS

Models	correlation after nonlinear regression			Spearman rank-order correlation		
	Low bound	CC	High bound	Low bound	CC	High bound
MGBIM	0.3303	0.5073	0.6499	0.2329	0.4245	0.5844
JQS	0.8964	0.9341	0.9583	0.8358	0.8942	0.9327
MBVS	0.9219	0.9568	0.9795	0.8471	0.9004	0.9358

### REFERENCES

- [1] M. Y. Shen and C.-C. J. Kuo, "Review of postprocessing techniques for compression artifact removal," *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2-14, 1998.
- [2] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proceedings of the IEEE*, vol. 81, no. 10, pp. 1385-1422, 1993.
- [3] C. H. Chou and Y. C. Li, "Perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 467-476, 1995.
- [4] X. Yang, W. Lin, Z. Lu, E. Ong, and S. Yao, "Motion-compensated residue preprocessing in video coding based on just-noticeable-distortion profile," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 742-751, 2005.
- [5] W. Lin, L. Dong, and P. Xue, "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 7, pp. 900-909, 2005.
- [6] A. W. C. Liew and H. Yan, "Blocking Artifacts Suppression in Block-Coded Images Using Overcomplete Wavelet Representation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 4, pp. 450-461, 2004.
- [7] X. K. Yang, S. Yao, Z. Lu, W. S. Lin, and E. P. Ong, "Just-noticeable-distortion profile with nonlinear additivity model for perceptual masking in color images," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 3 ed Hong Kong, Hong Kong: Institute of Electrical and Electronics Engineers Inc., 2003, pp. 609-612.
- [8] H. R. Wu and M. Yuen, "Generalized block-edge impairment metric for video coding," *IEEE Signal Processing Letters*, vol. 4, no. 11, pp. 317-320, 1997.
- [9] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No reference perceptual quality assessment of JPEG compressed images," in *IEEE International Conference on Image Processing*, pp. 477-480, 2002.
- [10] 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/subjective.htm>.

[11] VQEG, "Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment," <http://www.vqeg.org>, Mar.2000.

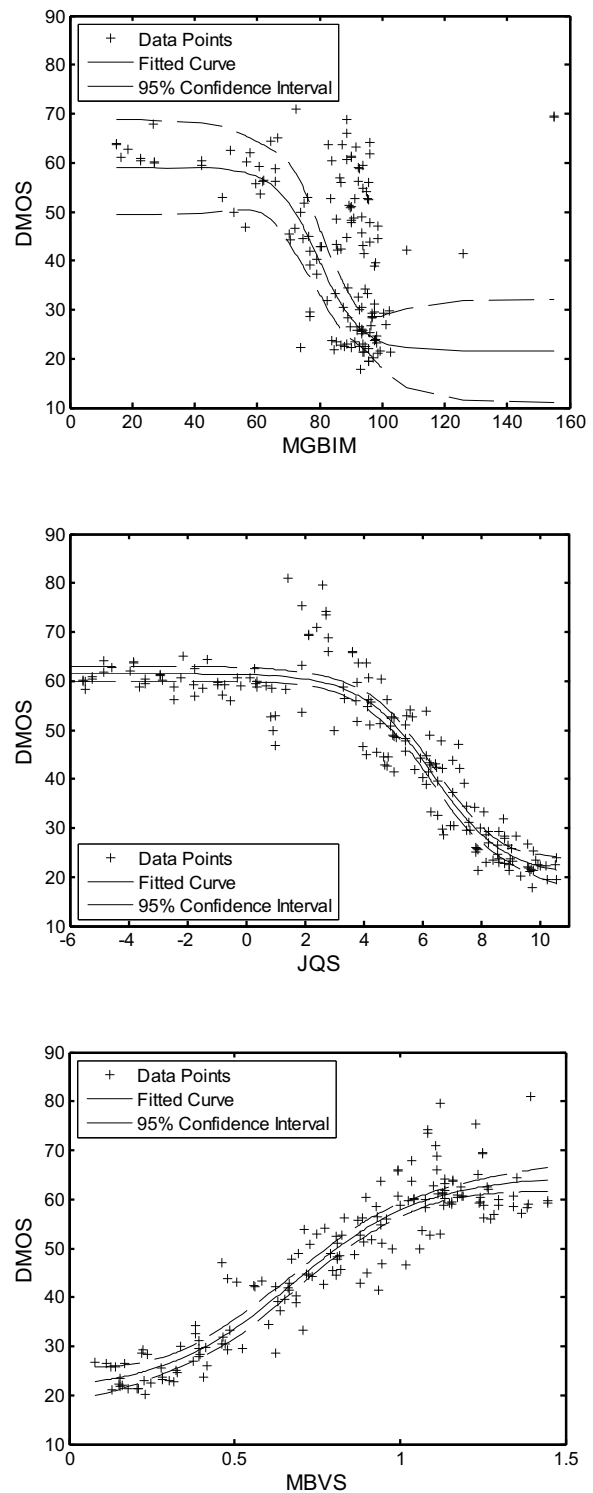


Figure 3. Scatter plotting results of the predictions of no-reference image quality metrics vs. DMOS.