Unified Spatial Masking for Just-Noticeable Difference Estimation

Jinjian Wu*, Fei Qi†, and Guangming Shi‡
Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China
School of Electronic Engineering, Xidian University, Xi'an, Shaanxi, 710071, P.R. China
Emails: *jinjian.wu@mail.xidian.edu.cn; †fred.qi@ieee.org; ‡gmshi@xidian.edu.cn

***Abstract*—In this paper, we introduce a unified spatial masking function for the estimation of just-noticeable difference (JND). Conventional models estimate several parts independently, and then combine these parts to get the JND. In this work, we treat the spatial masking effect as a nonlinear transformation of the luminance adaptation. To model the transformation, we measure the deviation of image contents from the ideal patterns to establish luminance adaptation rules. Considering both luminance difference and structural regularity, we derive a nonlinear spatial masking function by modulating luminance adaptation with the deviation coefficients. The masking function deduces an accurate estimation of the JND. Experiments demonstrate the validity of the proposed framework.**

***Index Terms*—Just-Noticeable Difference; Spatial Structure; Luminance Adaptation; Luminance Difference**

I. Introduction

Just-noticeable difference (JND) reveals limitations of human visual perception. For applications in which humans are the information sink of visual contents, JND provides great benefits [1], [2]. For instance, in image and video compression, the precision of quantization can be adjusted according to JND to save the code length or to improve the perceptual visual quality [3], [4]. In addition, it also helps to evaluate the image quality [5].

Existing models for JND estimation can be categorized into two classes according the domain knowledge they rely on. In transform domain methods [6], [7], images are divided into small blocks to apply domain transformation. Succeeding operations are performed on each block. As the block division corrupts image contents, such methods do not deal well with spatial structures.

In spatial domain methods [3], [8]–[10], luminance adaptation (LA) and spatial masking are the mainly considered factors. Generally, the LA part is modeled according to Weber's law [8]. As LA is valid only for image patterns with uniform foreground and background, researchers construct various spatial masking functions to make JND estimation feasible for natural images. In Chou and Li's work [8], spatial masking is constructed according to the maximum directional luminance difference (LD). However, since the human visual system are very sensitive to edges, this method over-estimated the JND threshold for edges. Yang *et al.* suggested to weight edge places with small values [3]. This method protects the primary edges, but the secondary edges are corrupted. In addition, the JND of regions with disorderly distributed intensities is still underestimated [9], [10].

The LA models the JND in an ideal condition that uniform foreground is over uniform background. As an image is with abundant textures, we divide the patterns of image contents to two classes. One is with regular textures, the other is with irregular ones. In this work, the JND for both cases are derived by measuring the deviation of the image contents to the ideal case. To achieve this, we build a unified spatial masking function, which is inherently a nonlinear transformation to LA, to estimate the JND. Two components, the LD and the structural regularity (SR), are employed to measure the deviation of regular and irregular textures from the ideal case, respectively.

For LD, the linearity of conventional models is valid only for LD values less than 80 [11]. So we build a nonlinear LD function, which keeps the linearity well for low LD values, while deviates greatly from the linear function at large LD values.

The SR measures the orderliness of image contents. From our empirical perspective, humans are hard to perceive noise in regions with disorderly distributed pixels. On the contrary, uniform regions, or regions with orderly distributed pixels, can hide very little noise. We employ a non-local procedure to get the measurement of the orderliness of image contents, i.e., the SR. This part is also non-linearly proportional to the LA, like the LD part.

The rest part of this paper is organized as follows. Section II introduce the unified spatial masking function based JND estimation framework. Section III provides experimental results demonstrating the validity of the proposed framework. And conclusions and discussions are given in Section IV.

II. Unified Spatial Masking based JND Estimation

The processing flowchart of our unified spatial masking based JND estimation framework is illustrated as Fig. 1. As shown in the flowchart, the LA map is first constructed according to Weber's law. Then the LD and SR maps are created independently. After that, nonlinear transformations are applied to both LD and SR maps. Finally, the LA, LD, and SR components are combined by multiplication to obtain the JND. Based on the estimated JND mask, noise are added to the original image to evaluate the performance of the proposed JND estimation framework.

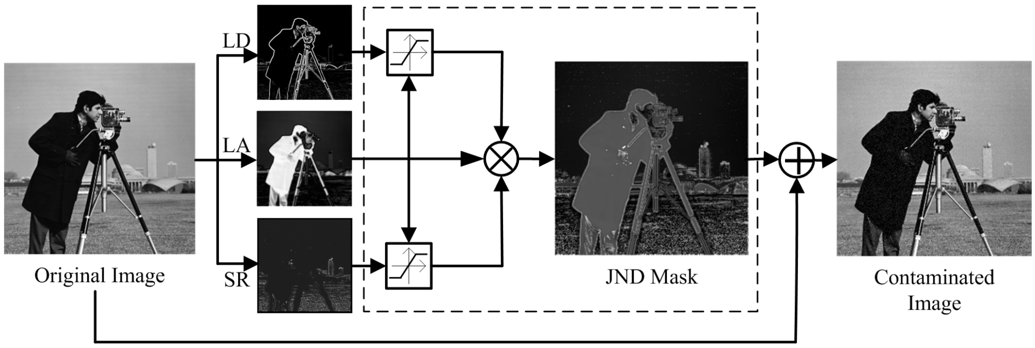


Fig. 1: The processing flowchart of our unified spatial masking based JND estimation framework.

0	0	0	0	0
1	3	8	3	1
0	0	0	0	0
-1	-3	-8	-3	-1
0	0	0	0	0

(a)

0	0	1	0	0
0	8	3	0	0
1	3	0	-3	-1
0	0	-3	-8	0
0	0	-1	0	0

(b)

0	0	1	0	0
0	0	3	8	0
-1	-3	0	3	1
0	-8	-3	0	0
0	0	-1	0	0

(c)

0	1	0	-1	0
0	3	0	-3	0
0	8	0	-8	0
0	3	0	-3	0
0	1	0	-1	0

(d)

Fig. 2: Filters for four directions.

A. Luminance Adaptation

LA is the foundation of JND estimation. As it has been well established, we follow Chou and Li's formulation [8], i.e.,

$$\mathrm{LA}(x) = \begin{cases} a_0 + a_1\sqrt{B(x)/B_0} & \text{If } B(x) < B_0 \\ \gamma[1 + B(x)] & \text{else} \end{cases}, \tag{1}$$

where $a_0 = 20$, $a_1 = -17$, $\gamma = 3/128$, $B_0 = 127$, and $B(x)$ is the background luminance of pixel x.

B. Luminance Difference

In Chou and Li's work [8], LD is defined as

$$\mathrm{LD}'(x) = \max_{k=1,\dots,4} G_k(x), \tag{2}$$

$$G_k = \varphi \nabla_k * I, \tag{3}$$

where φ is the normalizing coefficient with $\varphi = 1/16$, I is the original image, ∇_k are four directional filters, whose kernels are shown in Fig. 2, and symbol $*$ denotes the convolution operation.

Chou and Li used the LD defined in (2) to estimate JND. However, the validity of the linearity of this model is hold only for LD values less than 80, according to the study in [11]. For LD values larger than 80, Chou and Li's model [8] produces over estimated JND. To tackle this problem, we define the following LD coefficient,

$$C_{\mathrm{LD}}(x) = \frac{\mathrm{LD}^2(x)}{(L/3)^2 + \mathrm{LD}^2(x)}, \tag{4}$$

where L is the maximum gray level of the image. The truncated LD value is given by

$$\mathrm{LD}(x) = \max\left\{0, \mathrm{LD}'(x) - \mathrm{LA}(x)\right\}, \tag{5}$$

which makes LD take into effect only when its value is larger than LA.

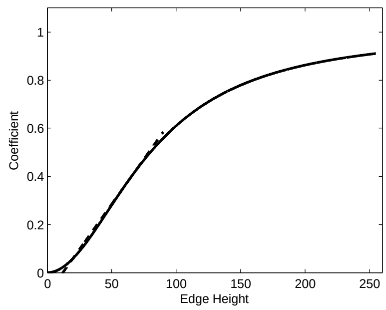


Fig. 3: The LD coefficient. The dashed line is drawn according to the study in [11].

The curve of LD coefficient with respect to LD values is shown in Fig. 3. The linear part is drawn with a dashed line according to the experimental curve given in [11]. It can be seen that, our LD coefficient approximates the dashed line well for small LD values while punishes the over-estimation for LD values larger than 80.

C. Structure Regularity

According to our empirical perspective, disordered regions can conceal large quantity of noise. And most conventional JND estimation model under-estimates the JND level for disordered regions. In this work, we employ the following non-local procedure to measure the structure regularity,

$$\mathrm{SR}'(x) = \sum_{y \in \mathcal{R}(x)} S(x,y)\big(I(x) - I(y)\big), \tag{6}$$

where x and y are two distinct pixels, $S(x,y)$ is the similarity between x and y, I is the original image, and $\mathcal{R}(x)$ is a region surrounding pixel x. The similarity $S(x,y)$ between two pixels x and y is calculated by

$$S(x,y) = \frac{1}{\alpha(x)} \exp\left(-\frac{d(x,y)}{2h_x^2}\right), \tag{7}$$

where $\alpha(x) = \sum_y \exp\left(-\frac{d(x,y)}{2h_x^2}\right)$ is a normalizing constant, parameter h_x controls the decay rate of the similarity, and $d(x,y)$ denotes the distance between two local regions $\Omega(x)$ and $\Omega(y)$ centering at x and y, respectively. Here, we take

$d(x, y)$ as the sum of squared differences between corresponding image patches $\Omega(x)$ and $\Omega(y)$, i.e.,

$$d(x, y) = \left\|\Omega(x) - \Omega(y)\right\|_2^2. \tag{8}$$

As an effective metric of the roughness of a region, the variance $\sigma(x)$ of the local block $\Omega(x)$ is adopted to adjust the decay variable h_x,

$$h_x = \begin{cases} \sigma_0 & \text{If } \sigma(x) \leq \sigma_0 \\ \sigma_0\sqrt{\sigma_0/\sigma(x)} & \text{else} \end{cases}, \tag{9}$$

with $\sigma_0 = 10$ in this paper.

Similar to the LD coefficient, we define the following nonlinear structural regularity coefficient,

$$C_{\text{SR}}(x) = \frac{\text{SR}(x)}{5\sigma(x) + \text{SR}(x)}. \tag{10}$$

The truncated SR is obtained as

$$\text{SR}(x) = \max\left\{0, \text{SR}'(x) - \text{LA}(x)\right\}. \tag{11}$$

D. JND Estimation

Since we model the JND for non-uniform image contents as deviations from uniform patterns, the final JND is estimated by applying the spatial masking of both LD and SR as

$$\text{JND}(x) = \left[1 + k_{\text{LD}}C_{\text{LD}}(x)\right]\left[1 + k_{\text{SR}}C_{\text{SR}}(x)\right]\text{LA}(x), \tag{12}$$

where parameters k_{LD} and k_{SR} are set to 2 and 8, respectively.

III. Experimental Results

In this section, we show the performance of the proposed JND estimation framework. We implement Chou and Li's method [8], Yang *et al.*'s method [3], and Zhang *et al.*'s method [6]. And we compare these methods with the proposed one on several test images. Two typical images, the tank and cameraman, are illustrated in Fig. 4 and Fig. 5, respectively. Other test images are with similar performance.

In the tank test case, which is taken from the USC-SIPI image database [12], as shown in Fig. 4, our method adds less noise to the tank while more to the messy grassland than the other three methods. As the tank is a very regularly structured target, the visual quality degradation is heavy in the noisy images generated according to the other three models. This is because edges on the tank are corrupted by noise. In addition, the tank is the perceptually salient object in this image, so corrupting the tank gives uncomfortable visual impression.

For the cameraman image, as shown in Fig. 5, our method gives the most comfortable visual experience. Because our method adds most noise to the grassland in the image where more noise can be concealed. Chou and Li's model [8] and Yang *et al.*'s method both over-estimate the JND level on the cloth of the man. Adding more noise on these parts produce very uncomfortable visual experience. Zhang *et al.*'s method adds more noise on edges, which corrupt the sharpness of edges around the salient objects, the man and the tripod. On the contrary, our method adds noise to edges without corrupting the smoothness in both tangent and normal directions of edges.

From these experiments we can summarize the contributions to the final JND of the three components. LA is mainly for the uniform regions in images, LD are dominant at edges. Both LA and LD are over-estimated in conventional methods. SR is for regions with disordered (or irregular) textures. This factor is generally under-estimated in conventional methods. The proposed unified spatial masking gives a well balance of these factors and produces accurate JND estimation.

IV. Conclusions

In this paper, we propose a unified framework for pure spatial masking based JND estimation. Our contributions are in two aspects. On one hand, we provide a method to fix the deviation from the linear Weber's law. On the other hand, we introduce a procedure to measure the regularity of image structures to improve the performance of JND estimation on regions with disordered textures. Experimental results show the validity and effectiveness of the proposed framework.

In the future, we plan to investigate the intrinsic biological mechanism associated with the two nonlinear equations.

Acknowledgments

This work is supported National Natural Science Foundation of China under grant NO. 60805012, 61070138, and 61033004, and by the Fundamental Research Funds for Central Universities under grant NO. JY10000902006.

References

[1] I. Hontsch and L. Karam, "Adaptive image coding with perceptual distortion control," *IEEE Trans. Image Process.*, vol. 11, no. 3, pp. 213–222, 2002.

[2] D. Xu, R. Wang, and J. Wang, "Video watermarking based on spatio-temporal JND profile," in *Digital Watermarking*, H. Kim, S. Katzenbeisser, and A. T. S. Ho, Eds. Springer Berlin Heidelberg, 2009, vol. 5450, pp. 327–341.

[3] X. K. Yang, W. S. Ling, Z. K. Lu, E. P. Ong, and S. S. Yao, "Just noticeable distortion model and its applications in video coding," *Signal Processing: Image Communication*, vol. 20, no. 7, pp. 662–680, 2005.

[4] Z. Chen and C. Guillemot, "Perceptually-Friendly H.264/AVC video coding based on foveated Just-Noticeable-Distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 6, pp. 806–819, 2010.

[5] W. Lin, L. Dong, and P. Xue, "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 900–909, 2005.

[6] X. Zhang, W. Lin, and P. Xue, "Just-noticeable difference estimation with pixels in images," *Journal of Visual Communication and Image Representation*, vol. 19, no. 1, pp. 30–41, Jan. 2008.

[7] Z. Wei and K. Ngan, "Spatio-Temporal just noticeable distortion profile for grey scale Image/Video in DCT domain," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 3, pp. 337–346, 2009.

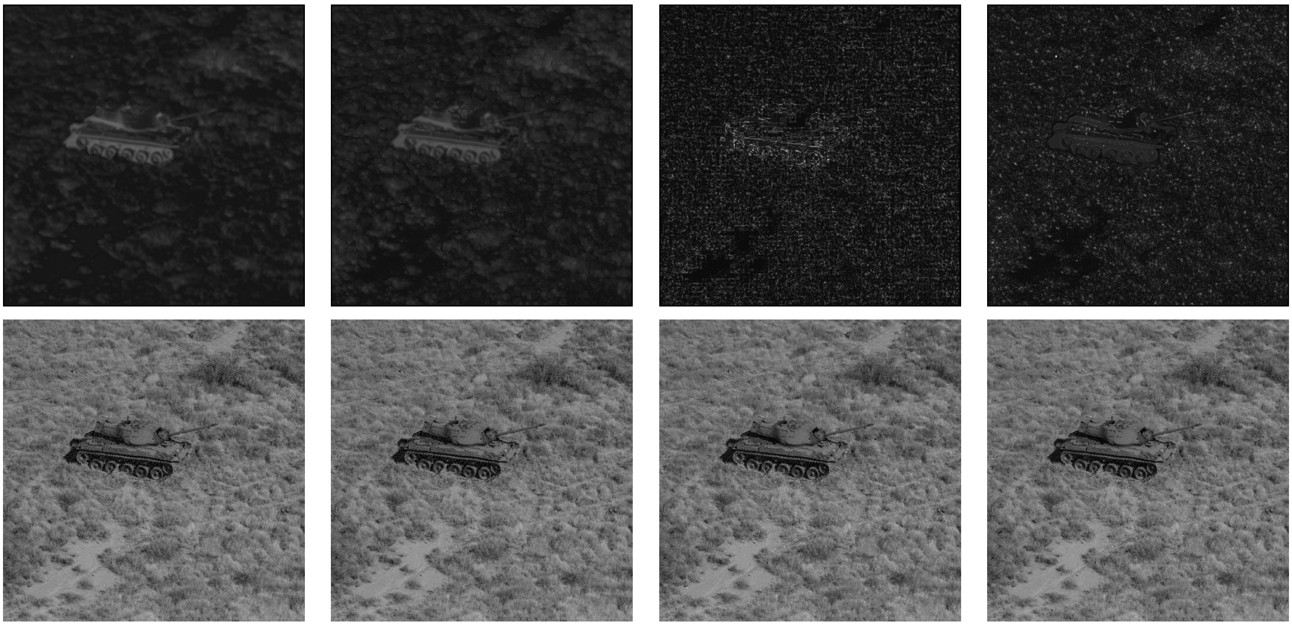
[8] C.-H. Chou and Y.-C. Li, "A perceptually tuned subband image coder based on the measure of just-noticeable distortion profile," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 6, pp. 467–476, 1995.

[9] A. Liu, W. Lin, F. Zhang, and M. Paul, "Enhanced just noticeable difference (JND) estimation with image decomposition," in *17th IEEE Int'l Conf. Image Processing, ICIP*, 2010, pp. 317–320.

[10] A. Liu, W. Lin, M. Paul, C. Deng, and F. Zhang, "Just noticeable difference for images with decomposition model for separating edge and textured regions," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1648 –1652, 2010.

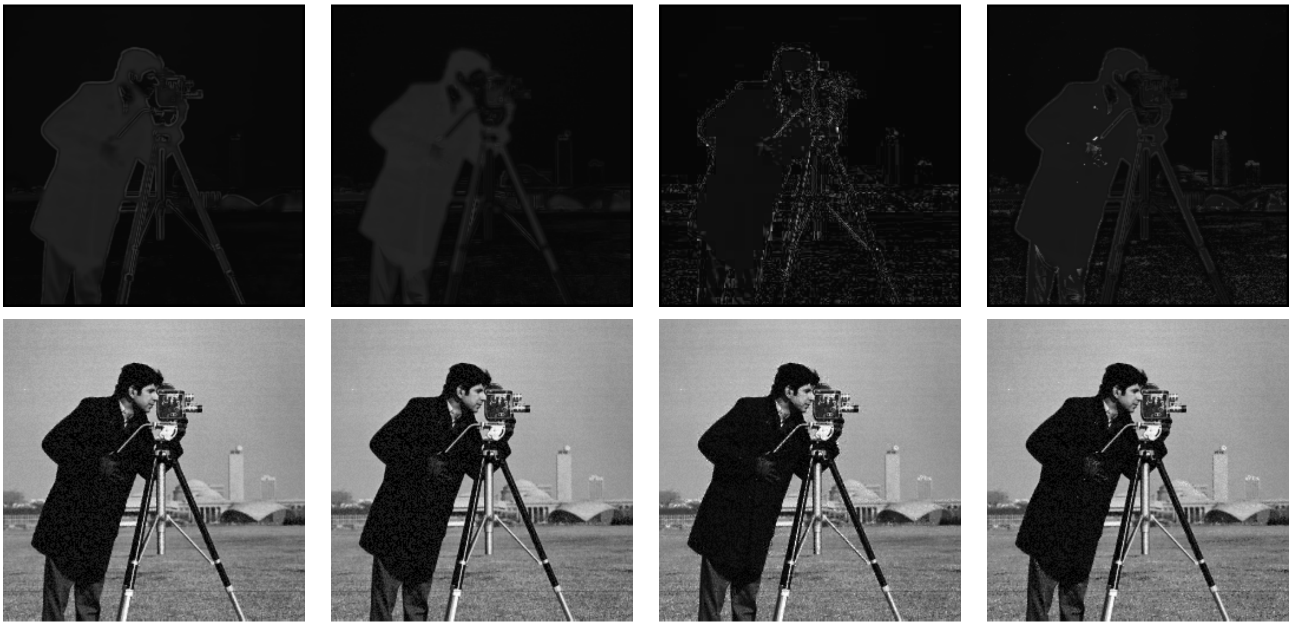
[11] A. Netravali and B. Prasada, "Adaptive quantization of picture signals using spatial masking," *Proc. IEEE*, vol. 65, no. 4, pp. 536–548, 1977.

[12] "The USC-SIPI image database," http://sipi.usc.edu/database/database.php, [Online; accessed June 1, 2011].



Chou Yang Zhang Ours

Fig. 4: Performance comparison of JND estimation methods on the test image tank. The top row shows JND masks, and the bottom row the JND guided noisy images (under the same energy level of noise with MSE = 71.29). Our method adds less noise to the visually salient tank while more to the messy grassland than the other three methods.



Chou Yang Zhang Ours

Fig. 5: Performance comparison of JND estimation methods on the test image cameraman. The top row shows JND masks and the bottom row the JND guided noisy images (under the same energy of noise level with MSE = 63.35). Our method gives the most comfortable visual experience, because our method adds most noise to the grassland in the image where more noise can be concealed.