# AN IMPROVED MODEL OF PIXEL ADAPTIVE JUST-NOTICEABLE DIFFERENCE ESTIMATION

*Jinjian Wu, Fei Qi, and Guangming Shi*

School of Electronic Engineering, Xidian University
Xi'an, Shaanxi, 710071, P. R. China

## ABSTRACT

A pixel-wise adaptive model for estimating the just-noticeable difference (JND) in spatial domain is proposed in this paper. As the human visual system (HVS) can be considered as a multichannel system, we assume that there exist two channels in the HVS, which deliver luminance adaption factor and texture masking factor, respectively. Both channels affect the JND threshold in a cooperative manner. The texture regions are with abundant redundancy and can tolerate much noise. The disorder degree and spatial masking of the texture are considered to estimate the texture masking effect, for deducing such JND threshold that coincides with the HVS. Finally, the luminance adaptation factor and texture masking factor are combined nonlinearly. Various experiments confirm the improved model has a better visual effect than models proposed before.

***Index Terms***— Just Noticeable Difference, Disordered Degree of Texture, Human Visual System, Texture Masking

## 1. INTRODUCTION

It is well known that there exists redundancy in images and videos. The human visual system can not sense the visual alert under a certain threshold which is called the just-noticeable difference threshold. This physiological and psychophysical phenomena is useful in image and video compression, quality evolution, and watermarking [1–3], etc. Researchers have investigated factors that affect the JND threshold. In most of related works, three factors have been mainly considered, namely, luminance adaptation, texture masking and contrast sensitivity function (CSF) [4–6]. A lot of methods have been proposed to estimate the JND threshold in transform domain or spatial domain in the recent past few years.

Spatial domain methods directly estimate each pixel's JND value which provide a visual view of threshold map for image pixels. These models applied for motion estimation,

quality evaluation, etc. In Chou and Li's model [1], the JND threshold of the background luminance is deduced from the Weber's law. And the texture masking effect is estimated by the maximum signal among the four edge detectors, which is considered as the luminance edge height and used to estimate the luminance difference threshold. Finally, the bigger value between the two masks is chosen as the JND threshold. Yang *et al.* [2] improve Chou and Li's model. Edges are distinguished from the texture regions to avoid over estimating of their JND thresholds, since the HVS is sensitive to edges. And the luminance adaptation and texture masking are integrated with a nonlinear additive formulate for more aggressive estimation. In Chiu and Berger's model [7], texture masking is determined by the bigger luminance difference in horizontal and vertical directions. The JND threshold is defined as a weighted sum of luminance adaptation threshold and texture masking threshold.

Transform domain JND estimation takes the spatial CSF into account. These methods are usually used in perceptual image and video compression. In Zhang *et al.*'s model [3,6], a method which incorporates the spatial CSF into pixel-wise JND estimation based on the DCT-subband is proposed. It includes a quasi-parabolic adjustment for luminance adaptation and classifies the blocks into smooth, edge, and texture areas.

Although texture masking has been under investigation for quite a long time, there are few, if exists, effective estimation. In Chou's model and Yang's model, texture masking is just considered as the luminance difference threshold based on the luminance edge height. This model is effective only if the background masking edge is presented briefly [5]. In Chiu's model, texture masking is determined by the bigger luminance difference between the currently pixel and its horizontal and vertical directions. Texture regions often have no obvious edges in the two directions, or there are always no obvious directions in texture regions, though these texture regions can tolerate abundant noise. The texture masks are always underestimated in the two models.

In this paper, a new pixel adaptive JND estimation method is proposed. The disorder of texture is considered when estimating the texture masking factor, for getting much precise JND threshold of the texture masking part. Then the luminance adaptation factor and the texture masking factor are

nonlinearly combined. In comparison with other pixel adaptive JND estimation models, the proposed model has a more aggressive estimation of the image's JND threshold without subjective visual difference. The rest of the paper is organized as follow. In section 2, the new texture masking computational method is proposed, and the new JND estimation model is presented based on nonlinear combination. Experimental results are shown in section 3. In section 4, we draw some conclusions.

## 2. PIXEL ADAPTIVE JND ESTIMATION

Because of the inherent perceptual redundancies of the image/video, noise below the JND threshold cannot introduce any visible difference in the HVS. Factors which affect the JND threshold have been found in the past. In our model, two major factors, luminance adaptation and texture masking, are chosen to estimate the JND threshold. The texture masking is always underestimated in the past. In this paper a new factor, disorder of the texture, is added to estimate the texture masking for computing a much precise JND threshold of the texture region. We assume there exist two channels in the HVS which convey information of luminance adaptation and texture masking separately and simultaneous, and these information is treated together in the brain. A nonlinear combination is proposed to imitate such a procedure in this paper.

### 2.1. Luminance Adaptation

The Weber's law describes the masking effect in the uniform background. In Chou and Li's model [1], the visibility threshold function due to background luminance is acquired based on Weber's law, which is shown as follow

$$T_{lum}(x,y) = \begin{cases} a_0 + a_1\sqrt{\frac{B(x,y)}{B_0}} & \text{If } B(x,y) < B_0 \\ \gamma + \gamma B(x,y) & \text{else} \end{cases}, \quad (1)$$

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

### 2.2. Texture Masking: Spatial and Disorder Degree

It is well known that there exist abundant perceptual redundancies in texture regions, and these regions can hide more JND noise than flat regions. But there exist no exact definition of texture up to the present. In [8], Netravali *et al.* define six factors, coarseness, contrast, directionality, line-likeness, regularity and roughness, as the major characters of the texture. In [5], the spatial masking is used to analysis the texture masking with a simple background. The spatial masking effect is deduced from the luminance edge height. Chou and Li [1] propose the texture masking equation based on spatial masking. Yang *et al.* [2] improve Chou and Li's model, and
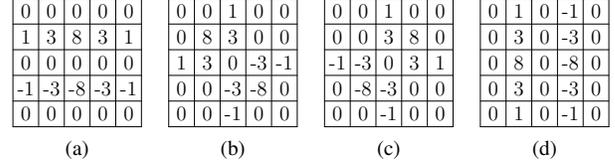


**Fig. 1**: High pass filters for texture detection at four directions
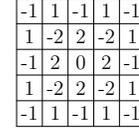


**Fig. 2**: Operator for calculating the disordered degree

propose a texture masking function which weighted by the magnitudes of gradients, shown as following

$$T_{dif}(x,y) = S(x,y)W(x,y), \quad (2)$$

where $S(x,y)$ is the spatial masking effect

$$S(x,y) = [\alpha_1 B(x,y) + \alpha_2][\beta_1 G(x,y) + \beta_2] + c, \quad (3)$$

$$G(x,y) = \max_{k=1,2,3,4} Grad_k(x,y), \quad (4)$$

$$Grad_k = \varphi f * G_k, \quad (5)$$

where $\alpha_1 = 0.01$, $\alpha_2 = 11.5$, $\beta_1 = 0.01$, $\beta_2 = -1$, the constant $c = -12$, $\varphi = 1/16$; $G$ is the maximum gradient, $f$ is the original image, and $G_k$ are four directional filters detecting textures, as shown in fig. 1.

The weight $W(x,y)$ in equation (2) relates to the magnitude of the edge at the pixel $(x,y)$. And $W$ is computed by edge detection followed with a Gaussian filter: $W = E * h$, where $E$ is the edge map of the image, detected by canny operator with threshold of 0.5, $h$ is a $k \times k$ Gaussian low pass filter with standard deviation $\sigma$, the size is set as $k = 5$ and $\sigma = 0.8$ is chosen in experiments.

The spatial masking effect is only effective when the texture is with a simple background. Texture masking is usually underestimated because the texture's background in natural images are always complex. For example, the grass regions have low luminance edge magnitude but can tolerate much noise. From experiments we know that the HVS is less sensitive to the irregular and less coarseness texture regions. An operator is used to estimate the degree of disorder of a given textural area, shown as fig. 2,

$$D = \lambda f * K, \quad (6)$$

where $K$ is the operator of the disorder degree, $f$ is the input image, the parameter $\lambda = 1/16$.

As the major factors affecting the texture masking, spatial masking and disorder degree are combined linearly. The perceptual threshold for texture masking factor is deduced as

$$T_{tex}(x,y) = T_{dif}(x,y) + \eta D(x,y). \quad (7)$$

We set the parameter $\eta = 2$ in our experiments.

## 2.3. JND Threshold of Each Pixel

The two factors, luminance adaptation and texture masking, affect the visibility threshold in the HVS in a cooperative manner. In the flat regions, the luminance adaptation is dominating whereas the texture masking effect is primary in the texture regions. The JND threshold of each pixel is a nonlinear combination of the two factors, its relationship is shown by the following equation,

$$T(x, y) = \theta_1(x, y)T_{lum}(x, y) + \theta_2(x, y)T_{tex}(x, y), \quad (8)$$

where $\theta_1(x, y) = \frac{T_{lum}(x,y)}{T_{lum}(x,y)+T_{tex}(x,y)}$ and $\theta_2(x, y) = 1 - \theta_1(x, y)$.

From eq. (8) we know that, in flat regions, $T_{tex}(x, y)$ is nearly equal to 0 and $\theta_1$ is approximate to 1, so the JND threshold is almost composed by the luminance adaption threshold. Whereas in texture regions, the texture masking affect is prominent. This conduces the weighted parameter $\theta_2$ has a bigger value and decrease the weighted value of $\theta_1$, the JND threshold is mainly the texture masking effect and the luminance adaption effect is restrained. The two factors affect the JND threshold in a nonlinear manner cooperatively, controlled by the two weighted parameter $\theta_1$ and $\theta_2$.

## 3. EXPERIMENTAL RESULTS AND DISCUSSIONS

We inject the JND-guided noise into an image $f(x, y)$ and make a comparison with other two pixel adaptive JND estimation models, which are Chou and Li's model and Yang *et al.*'s model. The JND-guided noise shaping equation is

$$f_{jnd}(x, y) = f(x, y) + srT(x, y), \quad (9)$$

where $s$ takes $+1$ or $-1$ randomly, and the parameter $r$ regulates the energy of JND noise.

As a subjective image quality metric, mean structural similarity (SSIM) [9] is effective to predict perceived image quality. We adjust the parameter $r$ to make sure each JND estimation model has the same SSIM value and compare the energy of injected noise. The baboon image is chosen in our experiment and we set the $SSIM = 0.9738$. The size of the baboon image is $512 \times 512$. The noise-contaminated images and its noise mask of each model are shown in fig. 3.

From the JND noise mask of Chou's and Yang's models, shown as fig. 3(b) and (d), it can be seen that less JND noise is injected into the images. In the proposed model, shown as fig. 3(f), much more JND noise is injected into the texture regions of the image. For the texture region contains abundant redundancy and can tolerate much JND noise. However the JND noise in the texture regions is underestimated both in Chou's and Yang's models. In the proposed model, the



(a) $SSIM = 0.9738$      (b) $MSE = 33.35$

(c) $SSIM = 0.9738$      (d) $MSE = 38.09$

(e) $SSIM = 0.9738$      (f) $MSE = 54.59$
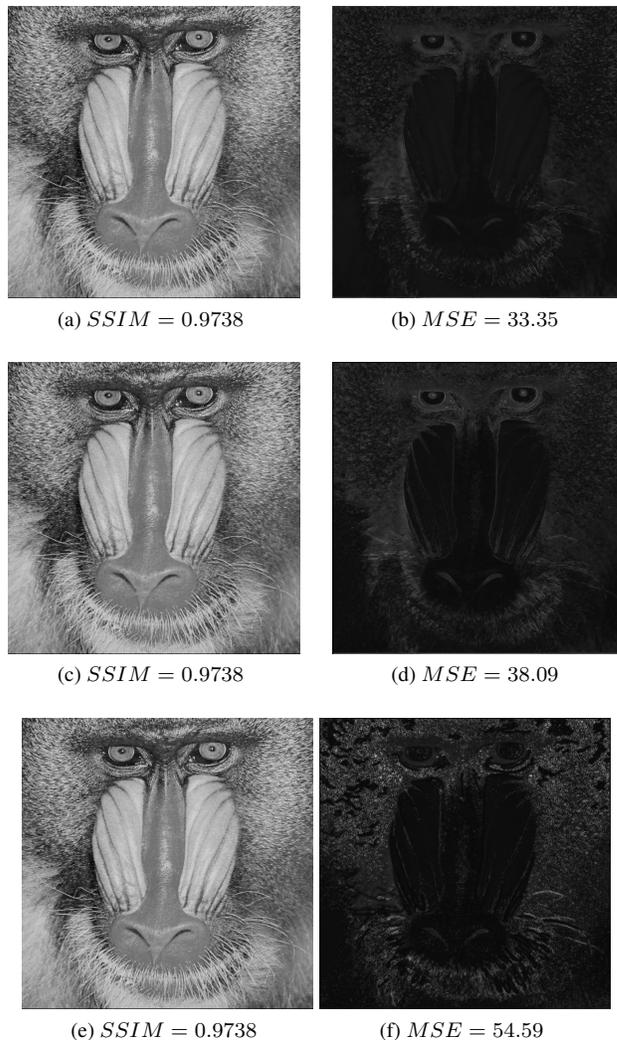
**Fig. 3**: The noise-contaminated images (left) and noise masks (right) of baboon. (a)(b) Chou's model, (c)(d) Yang's model, (e)(f) the proposed model.

disordered degree of the texture is considered when estimating the texture masking. Comparing the three models under the same SSIM value, the proposed model makes a more aggressive estimation and injects much more JND noise (with $MSE = 54.59$) without causing visible distortion.

We adjust the parameter $r$ to inject equivalent energy of JND noise into the image with each JND estimation model. The camera-man image is chosen in our experiment, whose size is $256 \times 256$. The JND noise contained images with $MSE = 89.34$ are shown in fig. 4. From fig. 4(a) we can see that much noise is injected into the edge and flat regions in Chou's model. And the distortion in these regions is apparent, shown as fig. 4(b). In Yang's model, though the edge region has been separated from the texture region, too much JND noise is injected and there exist obvious distortion in this region, as fig. 4(c) and (d) show.
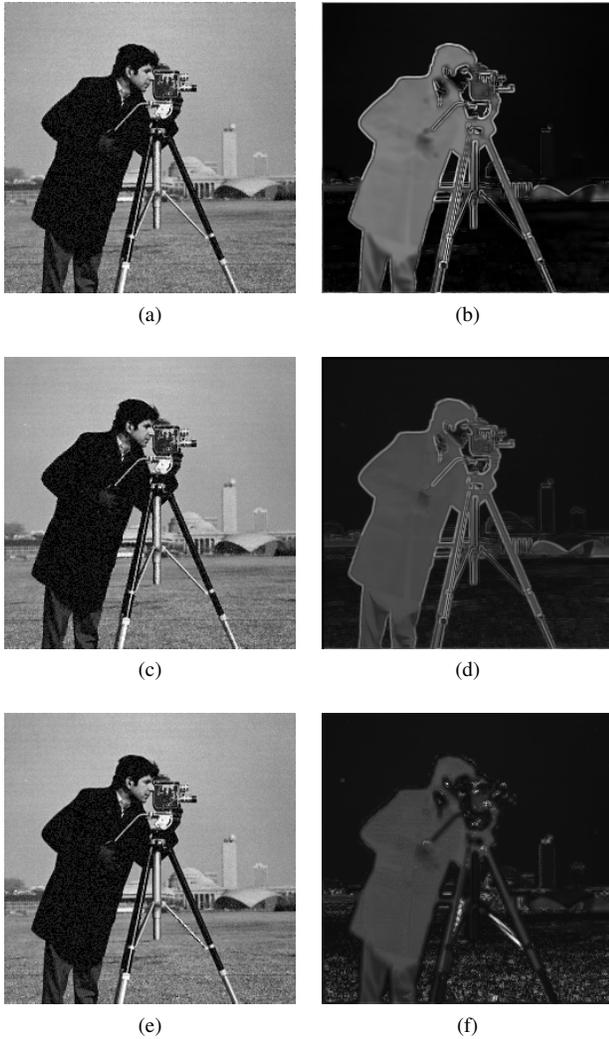
**Fig. 4**: The noise-contaminated images (left) and noise masks (right) of cameraman (with $MSE = 89.34$). (a)(b) Chou's model, (c)(d) Yang's model, (e)(f) the proposed model.

In the proposed model, shown as Fig. 4(f), much noise is injected into the texture regions, such as the grass region, in which much noise can be tolerated. With the nonlinear combination, little noise is injected into the edge and the flat region. Comparing with Chou's and Yang's models, as shown in fig. 4(a), (c) and (e), the proposed model has less distorted edge and flat region, and has a better perceptual quality.

## 4. CONCLUSIONS

It is well known that flat region can conceal less JND noise whereas texture region can tolerate much more. Existing pixel adaptive JND estimation models always injected much JND noise into flat regions and underestimated the JND threshold of texture regions. In the proposed model, we consider the disorder degree of the texture when estimating the tex-

ture masking. The luminance adaption and texture masks are combined with a nonlinear formula. Experiments show that the proposed model has an aggressive estimation in texture region without causing visible distortion.

However, many aspects remain for improvement. In the proposed model, the CSF effect, which is considered as a very important factor affects the JND estimation, has not been considered. Characters of texture are so complex that there are no clear definition. Further investigation on texture masking effect can be done.

## 5. REFERENCES

[1] Chun-Hsien Chou and Yun-Chin 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.

[2] 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.

[3] Xiaohui Zhang, Weisi Lin, and Ping Xue, "Just-noticeable difference estimation with pixels in images," *Journal of Visual Communication and Image Representation*, vol. 19, no. 1, pp. 30–41, 2008.

[4] Arun N. Netravali and Birendra Prasada, "Adaptive quantization of picture signals using spatial masking," *Proc. IEEE*, vol. 65, no. 4, pp. 536–548, 1978.

[5] Arun N. Netravali and Barry G. Haskell, *Digital Pictures: Representation and Compression*, Plenum, New York, 1988.

[6] X. H. Zhang, W. S. Lin, and P. Xue, "Just-noticeable distortion estimation for image pixels," in *IEEE 7th Workshop on Multimedia Signal Processing*, Shanghai, China, 2005.

[7] Yi-Jen Chiu and Toby Berger, "A software-only videocodec using pixelwise conditional differential replenishment and perceptual enhancement," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 3, pp. 438–450, 1999.

[8] Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.*, vol. 8, no. 6, pp. 460–473, 1977.

[9] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, and Eero P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.