Structural uncertainty based just noticeable difference estimation

Jinjian Wu
School of Electronic Engineering
Xidian University
Xi'an, Shannxi, China
jinjian.wu@mail.xidian.edu.cn

Weisi Lin
School of Computer Engineering
Nanyang Technological University
Nanyang 639798, Singapore
wslin@ntu.edu.sg

Guangming Shi
School of Electronic Engineering
Xidian University
Xi'an, Shannxi, China
gmshi@xidian.edu.cn

***Abstract*—Just noticeable difference (JND) reveals the minimum visible threshold of the human visual system (HVS), which is useful in visual redundancy reduction. Existing JND models estimate the visible threshold with luminance adaptation and contrast masking. As a result, the smooth and edge regions are effectively estimated, while the disorderly texture regions are always underestimated. The disorderly texture regions possess a large amount of disorderly structures and the HVS cannot fully perceive them. Therefore, in this work, we suggest to consider the disorder degree of structure for JND threshold estimation. According to the correlation among neighboring pixels, the uncertain information is extracted, and the disorder degree of structure is computed, which we called structural uncertainty. Then, taking the effect of background luminance, contrast, and structural uncertainty into account, a novel JND model is deduced. Experimental results demonstrate that the proposed JND can accurately estimate the visible thresholds of different image regions. Moreover, the proposed JND is adopted to remove visual redundancy for JPEG compression, which saves about 14% bit rate while keeping the perceptual quality.**

***Index Terms*—Just Noticeable Difference, Structural Uncertainty, Redundancy Reduction**

I. Introduction

It is well known that the resolution of the human visual system (HVS) is limited, and the HVS cannot perceive the change under a certain visibility threshold [1]. Just-noticeable difference (JND) [2], which accounts for the visible threshold of the HVS, can effectively represent the visual redundancy. Therefore, JND is useful in perception-oriented image/video processing, such as perceptual lossless based image/video coding [3], [4], watermarking [5], image/video quality assessment [6], [7], and so on.

The resolution of the HVS is widely investigated, and many JND estimation models have been introduced in the past decade. Since the HVS presents different sensitivities to different illuminants, background luminance is considered for JND threshold estimation. According to a subjective viewing test, the luminance adaptation is computed with Weber's law [3]. Moreover, the interaction among pixels cause masking effect in the HVS, which also affect the resolution of the human eye. Since the HVS is highly sensitive to the change of luminance, contrast masking is always computed to estimate the masking effect [4]. By considering the two factors, namely, background luminance adaptation and contrast masking effect, many JND estimation models are introduced [3], [8]. These models can only effectively estimate the JND thresholds of the plain and edge regions, while always underestimate the JND threshold of the texture regions [9]. With further analysis, we have found that the texture regions always represent disorderly structures, which contain amount of uncertain information that the HVS cannot fully understand. Therefore, we suggest to take structural uncertainty into account for JND estimation.

Recent research found on cognitive neuroscience, namely, the free-energy principle, indicate that the human brain possesses an internal generative mechanism, which separatively process the orderly and disorderly contents with different procedures [10]. Inspired by the free-energy principle, we suppose that the HVS is adaptive to extract the organization rule of the orderly region, and therefore, the HVS is sensitive to the change of the orderly region; on the contrary, the HVS cannot fully understand the content of the disorderly region with uncertain information, as a result, the HVS is less sensitive to the change on the disorderly region. In other words, the JND threshold of the disorderly region is larger than that of the orderly region. Therefore, we suggest to consider the orderness of an image region for JND threshold computation, and calculate the orderness according to its structural uncertainty. Then, we take luminance adaptation, contrast masking, and structural uncertainty into account, and introduce a novel JND model in this paper. The proposed JND model is applied to remove visual redundancy for JPEG compression, experimental results demonstrate that it saves about 14% bit rate while keeping the perceptual quality.

The organization of this paper is as follows: structural uncertainty of image regions is analyzed, and a novel JND estimation model is created in section II. In section III, the proposed JND model is firstly compared with the latest JND models to demonstrate its effectiveness, and then applies to JPEG compression for visual redundancy reduction. Finally, conclusions are drawn in section IV.

II. JND Estimation

In this section, we firstly analyze the effect of uncertain information on JND, and deduce the equation for structural uncertainty computation. Then, we consider the other two factors, i.e., luminance adaptation and contrast masking, and introduce a novel JND model.

A. Structural Uncertainty Computation

According to the free-energy principle, the HVS cannot fully perceive the contents of the disorderly regions, and the HVS is less sensitive to these regions [10]. Therefore, the structural uncertainty directly affects the resolution of the HVS, and regions with large structural uncertainty have high JND thresholds.

In order to compute the structural uncertainty, the uncertain information is firstly separated from the primary visual information of the original image. Here, the AR prediction model is employed to predict the primary visual information,

$$\mathcal{I}'(x) = \sum_{x_i \in \mathcal{X}} \mathcal{C}_i \mathcal{I}(x_i), \tag{1}$$

where $\mathcal{I}'$ is the reconstructed image, $\mathcal{C}_i$ is the AR coefficient, $\mathcal{X}$ is a surround region, and $\mathcal{I}$ is the original image.

With (1), the uncertain information can be computed as the prediction error, $\Delta\mathcal{I} = \mathcal{I} - \mathcal{I}'$. Then, the structural information of $\Delta\mathcal{I}$ is analyzed with local binary patterns (LBPs) [11]. The pattern of a pixel x_0 is calculated as the comparing value between x_0 and its surrounding x_i,

$$\text{LBP}(x_0) = \sum_{i=1}^{8} s(x_i - x_0)\, 2^i, \tag{2}$$

$$s(x_i - x_0) = \begin{cases} 1, & x_i - x_0 \geq 0 \\ 0, & x_i - x_0 < 0. \end{cases} \tag{3}$$

With the pattern number of each pixel, an image region can be mapped into a 2^8-bin histogram. In order to reduce the bin number, the normal rotate invariant LBPs are chosen, which have 10 different types of patterns [11]. As a result, the structural information of an image region can be represented by a 10-bin histogram. According to the information theory, the structural uncertainty of a pixel can be calculated as the entropy of its local region,

$$\mathcal{U}_s(x) = \sum_{i=1}^{10} -p_i \log p_i, \tag{4}$$

where $\mathcal{U}_s(x)$ is the structural uncertainty of pixel x, and p_i is the probability of the i-th bin. In order to give a clear view, an example is shown in Fig. 1, where (a) is the original image, and (b) is the corresponding structural uncertainty map.

B. JND Model

The JND threshold is determined by background luminance $\mathcal{L}_b$, luminance contrast $\mathcal{C}_l$, and structural uncertainty $\mathcal{U}_s$. Therefore, we try to combine the three factors and deduce a uniform equation for JND estimation.

Firstly, the HVS present different sensitivities to different background luminance, and the luminance adaptation is considered in JND modeling [3],

$$f_1(\mathcal{L}_b) = \begin{cases} 17 \times (1 - \sqrt{\frac{\mathcal{L}_b}{127}}) & \text{If } \mathcal{L}_b \leq 127 \\ \frac{3}{128} \times (\mathcal{L}_b - 127) + 3 & \text{else} \end{cases} \tag{5}$$

(a) Original image

(b) Structural uncertainty

Fig. 1: Structural uncertainty demonstration.

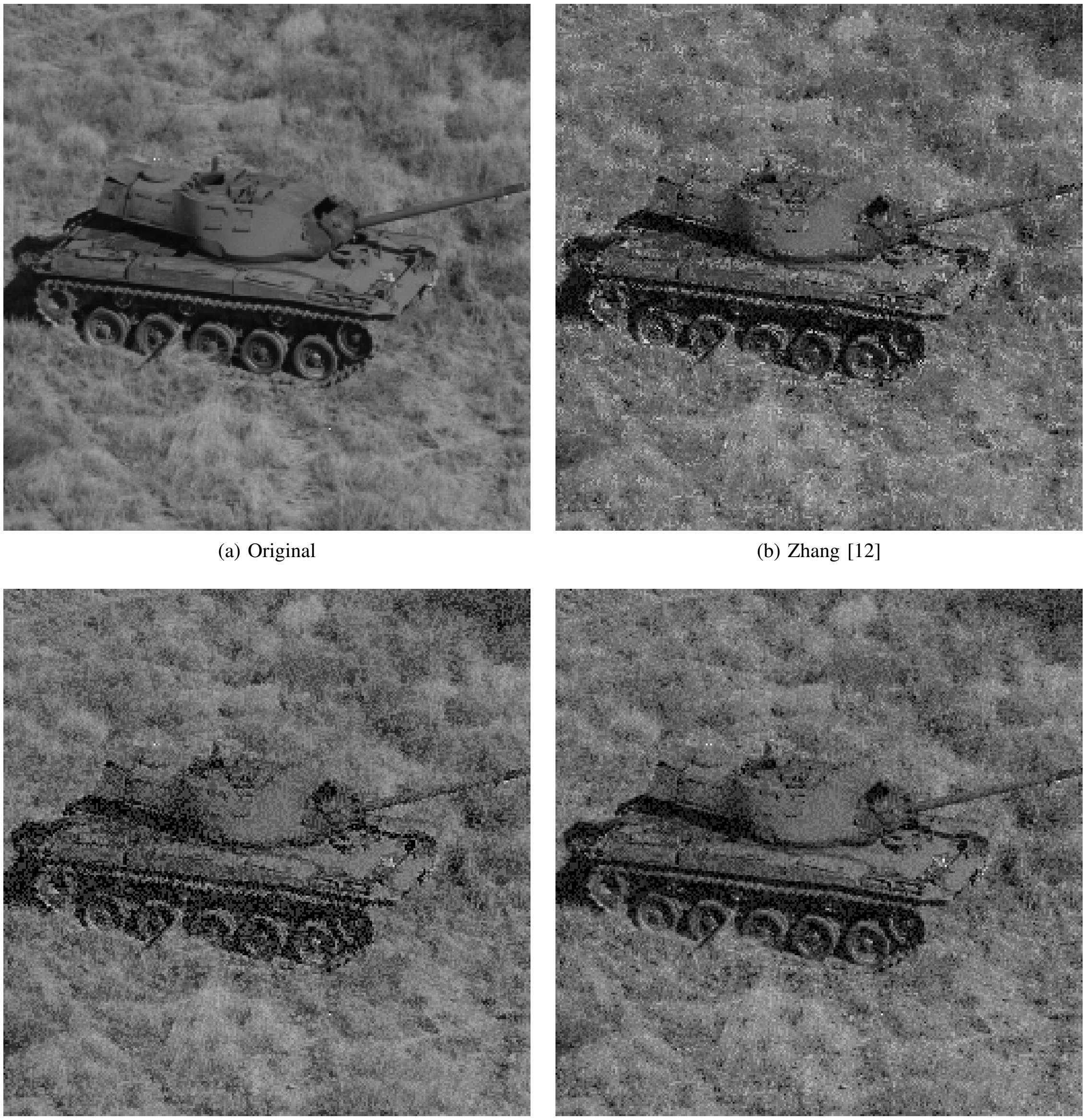
Next, the functions of luminance contrast and structural uncertainty on JND threshold are analyzed. From experiment we have found that the two factors are correlated: an image region with large luminance contrast and large structural uncertainty presents high JND threshold; on the other conditions, image regions present low JND thresholds. Therefore, we combine the two factors for JND estimation,

$$f_2(\mathcal{C}_l, \mathcal{U}_c) = 0.115\, \mathcal{C}_l\, \mathcal{I}(\mathcal{U}_l, \mathcal{C}_c) \tag{6}$$

$$\mathcal{I}(\mathcal{U}_l, \mathcal{C}_c) = 1 + (1 + 2\exp(-0.6\, \mathcal{C}_l))\, \mathcal{U}_c \tag{7}$$

As mentioned above, the JND threshold is determined by three factors, i.e., background luminance, luminance contrast, and structural uncertainty. The luminance adaptation and structural masking effect have been computed with (5) and (6), respectively. In order to acquire the JND threshold, we need to combine the two portions. However, there exists overlap between the two portions. Here, we adopt the nonlinear combination method, i.e., NAMM [4], to calculate the JND threshold,

$$\text{JND} = f_1(\mathcal{L}_b) + f_2(\mathcal{C}_l, \mathcal{U}_c) - 0.3 \times \min\{f_1(\mathcal{L}_b), f_2(\mathcal{C}_l, \mathcal{U}_c)\}. \tag{8}$$



(a) Original (b) Zhang [12]

(c) Liu [9] (d) Proposed

Fig. 2: JND model comparison. The tank image (a) contaminated by white noise (PSNR=$24.7dB$) with the guide of Zhang et al.'s model (b), Liu et al.'s model (c), and the proposed JND model (d).

III. Experimental Result

In this section, we firstly compare the proposed JND model with two latest JND models, i.e., Zhang et al.'s model [12] and Liu et al's model [9], to demonstrate the effectiveness. Then, the proposed JND model is applied to JPEG compression for visual redundancy reduction.

A better JND model can effectively present the visible thresholds of different image regions. In order to verify the accuracy of JND models, we inject white noise (with same energy) into test images under the guide of different JND models,

$$\hat{\mathcal{I}}(x) = \mathcal{I}(x) + \alpha \cdot \mathrm{rand}(x) \cdot \mathrm{JND}(x), \tag{9}$$

where $\hat{\mathcal{I}}$ is the contaminated image, α keeps the energy of white noise from different JND models as the same, and $\mathrm{rand}(x)$ randomly takes $+1$ or -1.

A representative image, the tank image (which includes smooth, edge, and texture regions), is chosen for JND performance comparison. White noise is injected into this image with the guide of different JND models (i.e., Zhang, Liu, and the proposed models). The energy of white noise injected into the image is the same, with PSNR=$24.7dB$. As shown in Fig. 2, Zhang et al.'s model mainly guides the noise into the edge regions, which results in obvious distortion; Liu et al.'s model distributes too much noise into the smooth region (i.e., the tank object), which is also quite sensitive to the HVS; the proposed model injects much noise into the grassland where the HVS is not sensitive to, and less noise into the tank object.

TABLE I: JPEG compression VS. JND guide JPEG compression.

Image	JPEG (bpp)	Proposed (bpp)	Bit Rate Saving	Subjective Quality	
				Mean	Std
Lena	0.644	0.545	15.4%	0.631	1.174
Babara	0.936	0.851	9.0%	0.300	1.138
Tank	0.773	0.556	28.1%	-0.188	1.117
Indian	0.960	0.817	14.9%	0.465	1.293
Huts	0.894	0.754	15.7%	-0.089	1.085
Mandrill	1.395	1.298	7.0%	0.125	1.091
Boat	0.894	0.756	15.4%	0.010	0.918
Airplane	0.947	0.836	11.7%	-0.125	1.264
House	0.687	0.598	13.0%	-0.007	1.185
Couple	0.826	0.715	13.4%	0.250	1.182
Average	0.897	0.773	13.8%	0.137	1.187

By comparing the three contaminated images, we can see that though the energy of the noise is same, the proposed model returns a better quality than the other two models.

Since the JND model can effectively represent the visual redundancy of the HVS, we apply the proposed JND model to JPEG compression. With the guidance of the JND threshold, we can modify the original signal according to its visual redundancy,

$$\tilde{I}(x) = \begin{cases} I(x) + \mathrm{JND}(x), & \text{if } I(x) - \overline{I}_B(x) < -\mathrm{JND}(x), \\ \overline{I}_B(x), & \text{if } |I(x) - \overline{I}_B(x)| \leq \mathrm{JND}(x), \\ I(x) - \mathrm{JND}(x), & \text{if } I(x) - \overline{I}_B(x) > \mathrm{JND}(x), \end{cases} \tag{10}$$

where $\overline{I}_B(x)$ is the mean value of a compression block B. Detailed information about this equation can be found in [13].

Ten oft-used images are chosen for compression (under the same QP) [14], and the bit rates for these compressed images are listed in Table I. Moreover, in order to compare the qualities of these JPEG compressed images and their corresponding JND guided JPEG compressed images, a subjective viewing test is designed, which follows the ITU-R BT.500-11 standard. In this test, 18 people from our laboratory are invited. The observers are asked to evaluate the qualities of these images: if the JPEG compressed image is better than its corresponding JND guided JPEG compressed image, then a '-1' value is given; on the contrary, a '+1' value is given; and a 0' value is given if they are with a same quality.

The valuation results are listed in Table I. As can be seen, the proposed JND threshold can effectively remove the visual redundancy for JPEG compression. An average of 13.8% bit rate can be saved for the 10 test images. Meanwhile, the qualities of these JND processed images are almost remained unchange, as the values under the '*Mean*' column shown in Table I. Therefore, the proposed JND model can accurately estimate the visible threshold, and it can be effectively applied into image compression for visual redundancy reduction.

IV. Conclusion

In this paper, we have proposed a novel JND estimation model. Existing JND models mainly consider the background luminance and luminance contrast for JND threshold computation. As a result, they can only effectively estimate the JND thresholds of the smooth and edge regions, while always underestimate the JND threshold of the texture region. Inspired by the recent research found on cognitive neuroscience, we suggest that structural uncertainty is another factor which affect the visibility of the HVS. By considering three factors, i.e., background luminance, luminance contrast, and structural uncertainty, a novel JND model is deduced. Experimental results demonstrate that the proposed JND model outperforms the existing JND models. Moreover, the proposed JND model can effectively predict the visual redundancy, and it saves about 14% bit rate when applying to JPEG compression.

Acknowledgment

This work was supported by NSF of China (No. 61227004, 61372131, and 11204014), Ph.D. Programs Foundation of Ministry of Education of China (No. 20130203130001), the 111 Project (No. B07048), and the Fundamental Research Funds for The Central Universities (No. JB140227).

References

[1] S. L. Macknik and M. S. Livingstone, "Neuronal correlates of visibility and invisibility in the primate visual system," *Nature Neuroscience*, vol. 1, no. 2, pp. 144–149, Jun. 1998.

[2] W. Lin, *Computational Models for Just-noticeable Difference*. CRC Press, 2006, chapter 9 in Digital Video Image Quality and Perceptual Coding, eds. H. R. Wu and K. R. Rao.

[3] C. H. Chou and Y. C. Li, "A 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. K. Yang, W. S. Lin, 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.

[5] Y. Niu, J. Liu, S. Krishnan, and Q. Zhang, "Combined just noticeable difference model guided image watermarking," in *IEEE Int'l Conf. Multimedia and Expo (ICME)*, 2010, pp. 1679–1684.

[6] J. Wu, W. Lin, G. Shi, and A. Liu, "Perceptual quality metric with internal generative mechanism," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 43–54, Jan. 2013.

[7] W. Lin and C. J. Kuo, "Perceptual visual quality metrics: A survey," *J. Visual Communication and Image Representation*, vol. 22, no. 4, pp. 297–312, 2011.

[8] J. Wu, G. Shi, W. Lin, A. Liu, and F. Qi, "Just noticeable difference estimation for images with free-energy principle," *IEEE Transactions on Multimedia*, vol. 15, no. 7, pp. 1705–1710, 2013.

[9] 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 Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1648–1652, Nov. 2010.

[10] K. Friston, "The free-energy principle: a unified brain theory?" *Nat Rev Neurosci*, vol. 11, no. 2, pp. 127–138, Feb. 2010.

[11] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

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

[13] X. Yang, W. Lin, Z. Lu, E. Ong, and S. Yao, "Motion-compensated residue pre-processing in video coding based on just-noticeable-distortion profile," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 742–750, June 2005.

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