Just Noticeable Difference Estimation For Images with Free-Energy Principle

Jinjian Wu, Guangming Shi*, *Senior Member, IEEE*, Weisi Lin, *Senior Member, IEEE*, Anmin Liu, Fei Qi

***Abstract*—In this paper, we introduce a novel just noticeable difference (JND) estimation model based on the unified brain theory, namely the free-energy principle. The existing pixel based JND models mainly consider the orderly factors and always underestimate the JND threshold of the disorderly region. Recent research indicates that the human visual system (HVS) actively predicts the orderly information and avoids the residual disorderly uncertainty for image perception and understanding. So we suggest that there exists disorderly concealment effect which results in high JND threshold of the disorderly region. Beginning with the Bayesian inference, we deduce an autoregressive (AR) model to imitate the active prediction of the HVS. And then we estimate the disorderly concealment effect for the novel JND model. Experimental results confirm that the proposed JND model outperforms the relevant existing ones. Furthermore, we apply the proposed JND model in image compression, and around 15% of bit rate can be reduced without jeopardizing the perceptual quality.**

***Index Terms*—Just Noticeable Difference (JND), Free Energy, Disorder, Autoregressive (AR) Model, Internal Generative Mechanism (IGM)**

I. Introduction

As an important part of the central nervous system, the human visual system (HVS) helps us to know the outside world by processing the visual detail. Rather than literally translating the input scenes, the HVS actively infers the input scenes with an internal generative mechanism (IGM) [1], [2], [3]. Recently, a free-energy principle has been introduced, which tries to formulate the IGM and intends to provide a unified account for human action, perception, and learning [4], [5].

The underlying idea of the free-energy principle is that all adaptive biological agents resist a natural tendency to disorder [5]. In other words, by optimizing the configuration of the IGM, e.g., adaptively adjusting the way it samples the environment and/or the internal encodement of the information [4], [6], the HVS tries to extract as much information as possible to minimize the uncertainty of an input scene. The principle indicates that the HVS cannot fully process all of the sensation information and tries to avoid some surprises (i.e., information with uncertainties, which are usually from disorderly regions) [5], [7]. Such disorderly concealment effect reveals a key limitation of the human perception which should be properly considered in the just noticeable difference (JND) estimation models. The JND threshold, which refers to the minimum visibility threshold of the HVS, is useful in perceptual image/video processing systems [8], [9], [10])

The existing pixel based JND estimation models have not considered the disorderly concealment effect, and they are usually composed of two components, i.e., luminance adaptation and spatial masking [9], [11], [12], [13]. The luminance adaptation component is based on Weber's law and deduced with an experiment under uniform background [14]. And the spatial masking component is investigated with regular luminance edge and is computed as direct proportion to the edge height [15]. Both of them are deduced from experiments with orderly scenes, and therefore, the existing pixel based JND estimation models [9], [11], [12], [13] perform effectively on places with orderly contents (such as smooth, edge, and orderly texture regions), but underestimate the JND thresholds on the disorderly regions (such as disorderly texture regions). Furthermore, the transform domain JND models (e.g., [16], [17]) usually classify image blocks into three types (i.e., plain, edge, and texture), and then three different weights are set for the three types to highlight the texture region. However, the IGM is adaptive to orderly texture contents, and therefore, those orderly texture regions will be overestimated with [16] and [17].

We suggest to consider the disorderly concealment effect in JND threshold estimation. According to the free-energy principle, the orderly stimulus is easy to be precisely predicted, so that the HVS can understand it in detail. However, the disorderly information, which is much complex and with abundant uncertainty, is hard to be precisely predicted; and therefore, the HVS ignores its detail and merely extracts its primary structure. For example, since a straight line located at a uniform background contains little uncertainty, it is easy to sense any changes on the line. But in the case of a disorderly lawn, it is impossible to perceive the detailed structure of each grass and we will naturally ignore this kind of information. Therefore, the disorderly concealment effect is another important factor which determines the JND threshold.

In this paper, we introduce a novel JND estimation model for images based on the free-energy principle. By imitating the inference procedure of the IGM, we deduce an autoregressive (AR) model to predict the orderly contents of an input image, and the prediction residual is regarded as the

Jinjian Wu, Guangming Shi (Corresponding author), and Fei Qi are with Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education of China, School of Electronic Engineering, Xidian University. E-mail: jinjian.wu@mail.xidian.edu.cn; gmshi@xidian.edu.cn; fred.qi@ieee.org.

Weisi Lin and Anmin Liu are with the School of Computer Engineering, Nanyang Technological University, Nanyang 639798, Singapore. E-mail: wslin@ntu.edu.sg; liua0002@ntu.edu.sg.

This work is supported by the Major State Basic Research Development Program of China (973 Program) (No.2013CB329402), NSF of China(No. 61033004, 61070138, 61072104, and 61227004), and the Fundamental Research Funds for the Central Universities (No.K50513100005).

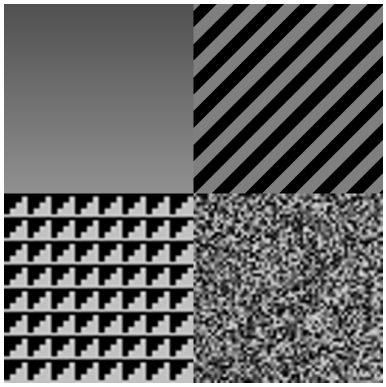

Fig. 1: Analysis on the orderly and disorderly textures. The image is formed by four different textures, namely, ($\mathcal{A}$) a graylevel patch, ($\mathcal{B}$) a zebra patch, ($\mathcal{C}$) a zigzag patch, and ($\mathcal{D}$) a random patch.

disorderly contents of the original image. Then we compute the JND threshold of the orderly term with the existing JND metrics, and propose a new procedure to compute the JND threshold of the disorderly term. Finally, we combine the two parts to deduce the overall JND threshold. The experimental result of the JND guided image noise shaping confirms that the proposed JND model correlates better with the human visual perception, with the evidence that it distributes less noise to sensitive orderly places and more to insensitive disorderly regions. Furthermore, we adopt the proposed JND model to improve the performance of perceptual image coding. It saves about 15% of bit rate while keeping the same perceptual quality.

The rest of this paper is organized as follows. In Section II, based on the Bayesian brain theory, we deduce an AR model to imitate the inference procedure of the brain. And then a novel JND estimation model is proposed and discussed in Section III. Experimental results for the JND based noise shaping and image coding are presented in Section IV. Finally, conclusions are drawn in Section V.

II. Orderly Content Prediction with The AR Model

In this section, we try to build a computational prediction model for the orderly content of the image to imitate the IGM of the brain. Based on the Bayesian brain theory, we analyze the minimum probability of prediction error, and deduce an AR based prediction model.

The HVS can accurately predict and fully understand the orderly content, while roughly perceives the disorderly content and avoids the remaining uncertainty [18], [5]. Fig. 1 shows a synthetic image, which is composed of smooth, zebra, zigzag and random textures. Intuitively, we can easily get the information conveyed by the orderly places, such as the three patches $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ shown in Fig. 1. As a result of these orderly contents have regular patterns, the IGM can accurately infer their contents according to these patterns. However, it is difficult for the IGM to infer the disorderly content in the patch $\mathcal{D}$ of Fig. 1, as it has much uncertain information. In summary, the change of contents in an orderly region is regular and predictable while that in the disorderly region is abrupt and unpredictable. Since the Bayesian inference is a powerful tool for information prediction, in this paper, the Bayesian brain theory [18] is adopted for orderly information prediction.

The rationale of the Bayesian brain theory is that the brain has a model of the world which tries to represent the sensor information probabilistically with minimum error [18]. For example, with an input image $\mathcal{F}$, the Bayesian brain system represents a pixel value x by maximizing the conditional probability $p(x/\mathcal{F})$. Since pixels are highly correlated with their surrounding $\mathcal{X} = \{x_1, x_2, \cdots, x_N\}$ (such as a 11×11 surrounding region which includes 120 neighbor pixels and excludes the central pixel x), $p(x/\mathcal{F})$ can be approximate to $p(x/\mathcal{X})$, and

$$p(x/\mathcal{X}) = p(x)\frac{p(\mathcal{X}/x)}{p(\mathcal{X})}. \tag{1}$$

Taking the logarithm on both sides of (1), we will have

$$\log p(x/\mathcal{X}) = \log p(x) + \log p(\mathcal{X}/x) - \log p(\mathcal{X}). \tag{2}$$

Then we find the expectation of (2). According to the Shannon information theory [19], $E(\log p(x)) = \sum_x p(x)\log p(x) = -H(x)$, and we can express the expectation of (2) as follows:

$$\begin{aligned} E(\log p(x/\mathcal{X})) &= E(\log p(x)) + E(\log p(\mathcal{X}/x)) - E(\log p(\mathcal{X})) \\ &= -H(x) - H(\mathcal{X}/x) + H(\mathcal{X}) \\ &= I(x;\mathcal{X}) - H(x). \end{aligned} \tag{3}$$

where $I(x;\mathcal{X})$ is the mutual information between x and $\mathcal{X}$, and $H(x)$ is the entropy of x. And maximizing $E(\log p(x/\mathcal{X}))$ is a relaxation of maximizing $p(x/\mathcal{X})$ [20].

Intuitively, if the pixel x is located at an orderly place $\mathcal{X}$ with little uncertainty, x can be exactly inferred with $\mathcal{X}$. From the perspective of information theory, if x is strongly correlated to $\mathcal{X}$, the mutual information $I(x;\mathcal{X})$ will be approximate to $H(x)$ [19]. When x possesses uncertainty, it is weakly or even not correlated with $\mathcal{X}$, and therefore, the mutual information $I(x;\mathcal{X})$ is approximate to zero. In summary, the mutual information $I(x;\mathcal{X})$ is an effective measurement of the extent of order of image content, the bigger $I(x;\mathcal{X})$ is the smaller disorderly x will be. Furthermore, $H(x)$ only determines by the original information. According to (1), (2), and (3), we can conclude that maximizing $p(x/\mathcal{X})$ is equivalent to maximize $I(x;\mathcal{X})$. Since the free-energy principle aims at minimizing the disorder, the Bayesian brain theory and the free-energy principle can be united under a uniform cognitive frame of the brain [5].

By defining $\mathcal{X}_{1,k} = (x_1, \cdots, x_k)$, we can decompose the mutual information $I(x;\mathcal{X})$ as follows [20]:

$$I(x;\mathcal{X}) = \sum_k I(x;x_k) + \sum_k [I(x_k;\mathcal{X}_{1,k-1}/x) - I(x_k;\mathcal{X}_{1,k-1})], \tag{4}$$

where x_k is the kth neighboring pixel in $\mathcal{X}$, $I(x;x_k)$ is the mutual information of pixel x with x_k, and the term $I(x_k;\mathcal{X}_{1,k-1}/x) - I(x_k;\mathcal{X}_{1,k-1})$ quantifies the mutual information which relates to the dependencies of these neighboring

(a) (b) (c)

Fig. 2: Image prediction with AR model. (a) The original image, (b) the predicted image, and (c) the residual of the original image and the predicted image (light/dark regions represent low/high residual values, respectively).

pixels [21]. Since the second term in the right side of (4) is much smaller than the first term [21], [20], we suppose that the mutual information $I(x;\mathcal{X})$ is mainly determined by the first term.

The IGM of the brain involves complicated psychological inference, which can effectively use the integral correlations among pixels to predict the sensor information and reduce the disorder. From the perspective of the Bayesian theory, the IGM tries to maximize the mutual information $I(x;\mathcal{X})$ in (4). So, the maximization of the mutual information is a similar procedure with the active inference mechanism of the sensor information.

By imitating the inference mechanism of the IGM, we try to create a computational prediction model, which predicts a pixel based on its surrounding pixels and their mutual information. From (4) we can see that the bigger the value $I(x;x_k)$ is, the more important it plays for mutual information $I(x;\mathcal{X})$ maximization. In other words, if a surrounding pixel x_k has a big value $I(x;x_k)$, it plays a much important role in the prediction of x within the IGM. In this paper, we take the mutual information between the central pixel and each pixel in its surrounding ($I(x;x_k)$) as the autoregressive coefficient, and we build a brain inference based AR model to predict an input image,

$$\mathcal{F}'(x) = \sum_{x_k \in \mathcal{X}} \mathcal{C}_k \mathcal{F}(x_k) + \varepsilon, \tag{5}$$

where $\mathcal{F}'$ is the predicted image, $\mathcal{C}_k = \frac{I(x;x_k)}{\sum_i I(x;x_i)}$ is the normalized coefficient, and ε is white noise. With (5), the orderly information of the input image (Fig. 2 (a)) can be accurately predicted, as shown in Fig. 2 (b).

III. The JND Estimation Model

In this section, we take three factors, which are luminance adaptation, spatial masking, and disorderly concealment effect, into account to create a novel JND model. The architecture of the proposed model is shown in Fig. 3. An input image is firstly preprocessed by an internal generative mechanism (IGM), which is decomposed into orderly and disorderly parts. Then the non-linear additivity model for masking (NAMM) [9] procedure and the disorderly concealment effect (DCE) procedure are adopted to estimate the JND thresholds of the two parts, respectively. Finally, the thresholds for the two parts are combined to achieve the overall JND value.

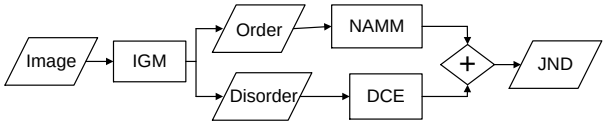

Fig. 3: The architecture of the proposed JND model.

Since the content of the predicted image is orderly, the existing formulations for luminance adaptation (LA) and spatial masking (SM) are chosen to estimate the JND threshold of the orderly component ($\mathcal{F}'$). And therefore, the NAMM [9] model is adopted to compute the JND threshold of the predicted part,

$$\mathrm{LA}(x) = \begin{cases} 17 \times (1 - \sqrt{\frac{B(x)}{127}}) & \text{If } B(x) \leq 127 \\ \frac{3}{128} \times (B(x) - 127) + 3 & \text{else} \end{cases} \tag{6}$$

$$\mathrm{SM}(x) = [0.01B(x) + 11.5][0.01G(x) - 1] - 12 \tag{7}$$

$$\begin{aligned} \mathrm{JND}_p(x) =& \mathrm{LA}(x) + \mathrm{SM}(x) \\ & - C^{gr} \times \min\{\mathrm{LA}(x), \mathrm{SM}(x)\}, \end{aligned} \tag{8}$$

where $B(x)$ is the background luminance of pixel x, i.e., the mean luminance of an image region; $G(x)$ is the maximum edge height of its 5×5 neighborhood [9]; JND_p is the JND threshold of the predicted part $\mathcal{F}'$; C^{gr} is the gain reduction parameter due to the overlapping between LA and SM, and is set as $C^{gr} = 0.3$ (the same as in [9]).

According to the theory of the IGM [18], [5], the brain works as an active inference system, which accurately predicts orderly content and avoids disorderly information. We regard the residue between the original image and its prediction (computed with (5)) as the disorderly image $\mathcal{D}$, which is computed as follow (as shown in Fig. 2 (c)),

$$\mathcal{D} = |\mathcal{F} - \mathcal{F}'|. \tag{9}$$

Under the IGM, the HVS ignores the detail of the disorderly region and only extracts a survey. So we can inject more noise into the disorderly region. Since the disorderly image represents the uncertainty of the original image, a pixel in

the disorderly image with large value means its corresponding original information is highly uncertain, and more noise can be concealed into it. Therefore, the JND threshold due to the disorderly concealment effect (JND_d) can be computed as

$$\mathrm{JND}_d(x) = \alpha \cdot \mathcal{D}(x), \tag{10}$$

where α is a disorderly adjustable parameter. In the current work we regard it as a fixed value for simplicity, and according to subjective viewing results on four different images we set $\alpha = 1.125$.

We have computed the JND thresholds of the orderly part and disorderly part with (8) and (10), respectively. But overlapping exists between the two parts which needs to be removed when combining them. And therefore, the nonlinear combination procedure of the NAMM model [9] is adopt again and the overall JND threshold is acquired

$$\begin{aligned}\mathrm{JND}(x) =& \mathrm{JND}_p(x) + \mathrm{JND}_d(x) \\ &- C^{gr} \times \min\{\mathrm{JND}_p(x), \mathrm{JND}_d(x)\}.\end{aligned} \tag{11}$$

IV. Experimental Results

In this section, we firstly inject JND noise into images and make a comparison between the proposed model and three existing JND models [9], [13], [16]. Then we apply the JND model to image compression to save bit rate under the same perceptual quality.

A. JND Models Comparison

An effective JND model should be able to guide shaping more noise into the insensitive places and less into the sensitive regions. Under the same injected noise energy, a more accurate JND model would result in better perceptual quality. In this subsection, we try to compare the proposed model with three latest relevant models, which are Yang et al.'s model [9], Liu et al.'s model [13], and Zhang et al.'s model [16]. To a testing image, the noise is injected with the guidance of its corresponding JND, which is shaped as follow,

$$\hat{\mathcal{F}}(x) = \mathcal{F}(x) + \beta \mathrm{rand}(x)\mathrm{JND}(x), \tag{12}$$

where $\hat{\mathcal{F}}$ is the JND noise contaminated image, β regulates the energy of JND noise, which makes the same energy for different JND models, and $\mathrm{rand}(x)$ randomly takes $+1$ or -1.

To make a clear comparison among these JND models, a representative synthetic image, as Fig. 4 shows, is chosen [22]. It is composed of smoothness, edge, orderly texture, and disorderly texture, i.e., the background of the rings, the black and white rings, the zebra crossing patches, and the disorderly blocks. The JND threshold of the image is computed with all of these JND models, and the JND noise is injected into the original image according to (12) with a same noise energy (MSE $= 160$).

From the JND mask maps (as shown in the first row of Fig. 4) we can see that, in Yang et al.'s model [9], most of noise is distributed into regions with dark background and secondary edge (such as the zebra crossing regions). Liu et al.'s [13] try to distinguish texture from edge. As a result, the grid regions, the zebra crossing regions, and the disorderly patches are considered as textural regions, and much more noise is injected into these regions. Zhang et al.s model [16] only injects noise into places with luminance changes, especially the edge/texture regions with high edge heights (i.e., the black ring and zebra crossing regions).

From the perspective of the computational formulations of the three existing JND models, they mainly adopt some orderly factors for JND estimation, and none has taken the disorderly concealment effect into account. As a result, too much noise is injected into the orderly texture regions. And therefore, as the contaminated images shown in the second row of Fig. 4, the zebra crossing blocks (locate at the left bottom) in the three images (Fig. 4 (e)-(g)) are obviously distorted.

Fig. 4 (d) and (h) show the JND mask map and the contaminated image of the proposed model, respectively. By considering the disorderly concealment effect, most of the noise is injected into the disorderly regions, especially the disorderly texture blocks locate at the top right corner. And little noise is injected into the orderly zebra crossing. Therefore, the contaminated image of the proposed model has a much better perceptual quality than those of the other three models.

In order to make a comprehensive comparison between the proposed model and the three existing models, a subjective viewing test is conducted on 10 frequently-used images (with size of 512×512). The setting (e.g., the viewing distance and environment) of the viewing test follows the ITU-R BT.500-11 standard [23]. In each test, two contaminated images of the same scene, where one image contaminates by the proposed JND model and the other one contaminates by another JND model (one of the three comparison models), are juxtaposed on a 17-in monitor (randomly at left or right) for comparison. Sixteen subjects are asked to evaluate which one is better and how much better it is. Seven of them are experts in image processing and the others are naive, and their eyesight is either normal or has been corrected. The evaluation rule and its quantitative score are shown in Table I. Meanwhile, if the left image is better (worse) than the right image, then a positive (negative) score is given. Furthermore, we set the viewing distance as four times of the image height.

TABLE I: Scores for Subjective Viewing Test

Score	0	1	2	3
Description	Same quality	Slightly better	Better	Much better

The comparison result of the subjective viewing test is shown in Table II, the standard deviation (std) values of the subjective scores are around 1, which means the subjective evaluation results from the sixteen people are steady and believable. In addition, as shown in the left of Table II, a positive (negative) mean score indicates that the proposed model has a better (worse) perceptual quality than the corresponding compared model. From the table we can see that the proposed model outperforms (the mean scores are positive) the other three models on most of the images, especially on the Tank and Port images. That is because the disorderly regions, such as the lawn in the Tank and the sea surface in the Port, are less sensitive to the HVS and can conceal much noise. And the proposed model considers the disorderly

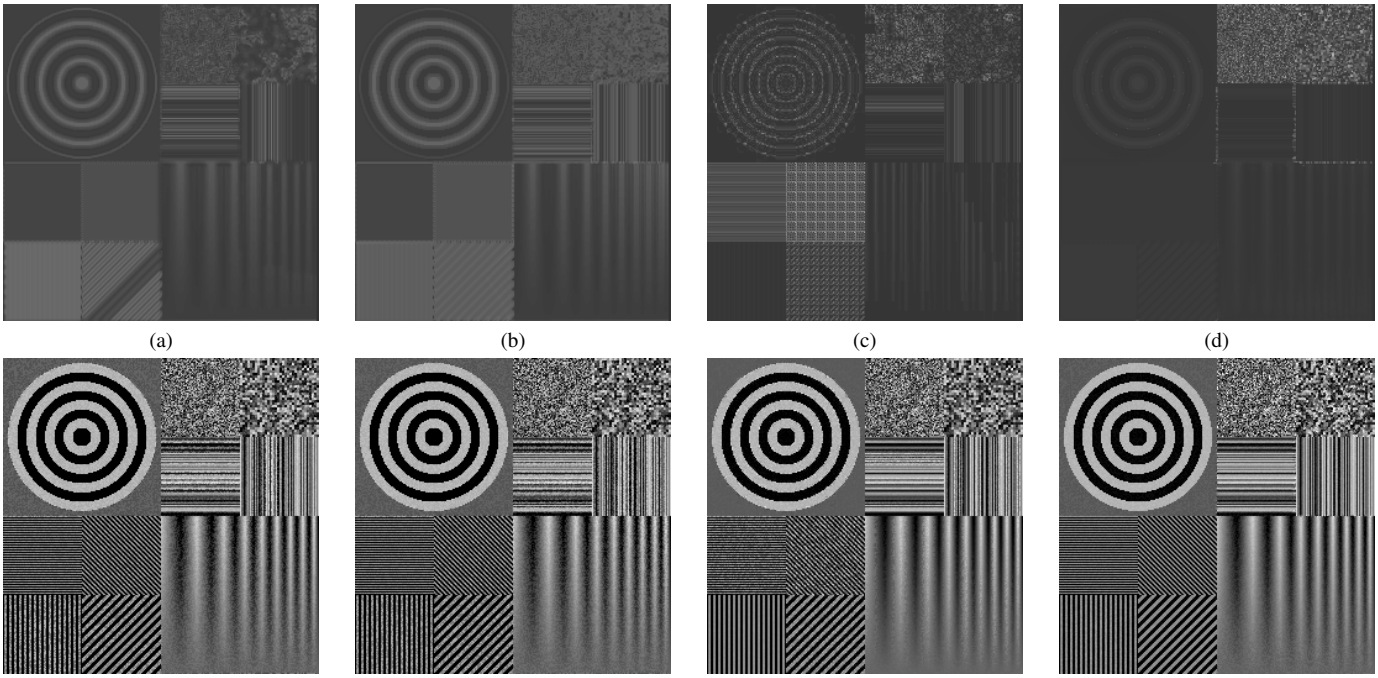

Fig. 4: JND models comparison on a representative image(with size 256×256). (a)-(d) The JND mask maps, and (e)-(h) The contaminated images with a same level of JND noise (MSE $= 160$). From left to right, they are the JND threshold of Yang et al.'s model [9], Liu et al.'s model [13], Zhang et al.'s model [16], and the propose model.

concealment effect, so much more noise is injected into these disorderly regions. However, there exist several cases that the proposed model is less effective than other models (i.e., on the Barbara, Indian, and Lena). With further analysis we have found that the three original images are with salient faces, and are slightly contaminated by noise. The proposed model computes the noise as uncertainty. In addition, in our current implementation, α is simply set as a fixed value, which is a little bigger for face regions since face regions are very sensitive to the HVS and little noise are allowed in these regions. Therefore, the proposed method performs less well for images with faces, and our further work is to optimize the value of α (e.g., make the value of α adaptive to the content of the image). On the whole, most of the subjective viewing scores and the overall average scores against the three models are positive, which further confirms that the proposed JND model outperforms these existing models.

TABLE II: Subjective viewing test results (the proposed model vs. three existing JND models, respectively) on contaminated images with JND noise.

Image \ Our vs.	Mean			Std		
	Yang	Liu	Zhang	Yang	Liu	Zhang
Barbara	0	-0.250	0.188	0.866	0.750	1.073
Tank	0.688	0.625	0.563	0.916	0.696	1.171
Huts	0.438	0.063	0	0.788	0.658	0.791
Boat	0.188	0.125	0	0.882	0.781	1.275
Port	1	0.625	1	0.935	0.696	1.225
Indian	-0.250	0.188	0.125	1.031	1.014	1.053
Mandrill	0.250	0.188	0.250	1.031	0.634	0.750
Lena	-0.250	-0.313	0.563	1.031	0.845	0.998
Airplane	-0.063	0.188	0.313	1.088	1.014	1.210
Couple	0.563	0.063	0.875	0.864	0.658	0.857
Average	0.263	0.194	0.388	0.926	0.788	1.040

B. Application in Image Compression

As the JND threshold reveals the minimum sensory signal strength to the HVS, an appropriate JND model can effectively remove perceptual redundancy and improve the performance of perceptual coding algorithms [11], [24]. To further demonstrate the proposed JND model, we apply it to image compressing for perceptual redundancy reduction.

With the guidance of JND threshold, we preprocess an input image by removing its perceptual redundancy according to the coding scheme of [24], which tries to smooth the original image within the constraint of JND threshold. Then we compress both the original image and the JND-processed image for comparison. The same 10 images as used in Section IV-A are compressed (under the same QP) with the JPEG and the proposed (i.e., preprocessing + JPEG) algorithms, respectively. Meanwhile, two different levels of QP, namely, QP=1 where the compression result has high perceptual quality and QP=2 where the compression is slightly perceptually loss, are chosen. A subjective viewing test (following the ITU-R BT.500-11 standard [23] which is introduced in the previous subsection) is performed on these images. Sixteen subjects are invited in the test, and the results are shown in Table III. With the help of the JND threshold, less bit rates are needed for all of these images, especially for these images with abundant of

TABLE III: Compression result comparison between JPEG algorithm and the proposed model.

Image (QP=1)	JPEG (bpp)	Proposed (bpp)	Bit Rate Saving	Quality	
				Mean	Std
Barbara	0.936	0.861	9.4%	0.500	1.118
Tank	0.773	0.565	26.8%	-0.125	1.317
Huts	0.894	0.746	16.5%	-0.188	1.285
Boat	0.894	0.773	13.6%	0	1.118
Port	0.941	0.861	8.6%	0.313	1.210
Indian	0.960	0.833	13.2%	0.563	1.368
Mandrill	1.395	1.303	6.6%	0.125	0.992
Lena	0.644	0.544	15.5%	0.813	1.184
Airplane	0.687	0.600	12.8%	0	1.248
Couple	0.826	0.725	12.2%	0.125	1.218
Average	0.891	0.781	13.4%	0.213	1.206
Image (QP=2)	JPEG (bpp)	Proposed (bpp)	Bit Rate Saving	Quality	
				Mean	Std
Barbara	0.620	0.562	9.4%	0.688	1.285
Tank	0.447	0.299	33.1%	0.063	1.177
Huts	0.558	0.430	22.9%	-0.313	0.992
Boat	0.457	0.399	12.7%	0.313	1.368
Port	0.592	0.504	14.9%	0.125	0.857
Indian	0.621	0.521	16.1%	0.438	1.177
Mandrill	0.884	0.749	15.3%	0.250	1.273
Lena	0.420	0.362	13.8%	1.250	0.759
Airplane	0.459	0.400	12.8%	0.063	1.138
Couple	0.538	0.454	15.6%	-0.125	1.118
Average	0.560	0.468	16.4%	0.275	1.114

disorderly places (e.g., Tank and Huts). At the meantime, their subjective quality is similar.

In summary, under the same perceptual quality (in fact, the proposed method is slightly better than the JPEG algorithm with positive average subjective viewing scores of 0.213 and 0.275), the proposed method saves on average 13.4% and 16.4% of bit rates against the JPEG algorithm on the two different QP levels, respectively. Therefore, with the help of the proposed JND model, an effective image compression algorithm to represent the input image with high perceptual quality and low bit rate is achieved.

V. Conclusion

In this paper, a novel JND model has been proposed based on the latest unified brain theory [1]. The free-energy principle indicates that the brain works with an internal generative mechanism for image perception and understanding, which avoids uncertainty with a disorderly concealment effect. We suggest that such disorderly concealment effect is important for JND estimation, and propose a procedure to handle it. With the proposed procedure, we can accurately estimate the JND threshold of the texture region which is underestimated by the existing models. The subjective experiment confirms that the proposed JND model outperforms the existing JND models. Furthermore, we have used the proposed JND model to improve the performance of image compression. With the aid of the proposed JND model, it has been demonstrated that we save about 15% of bit rate under the same perceptual quality (confirmed by the subjective viewing test).

[1]The source code is released at http://web.xidian.edu.cn/wjj/en/index.html

References

[1] K. J. Friston, J. Daunizeau, and S. J. Kiebel, "Reinforcement learning or active inference?" *Public Library of Science One*, vol. 4, p. e6421, Jul. 2009.

[2] R. L. Gregory, "Perceptions as hypotheses," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 290, pp. 181–197, Jul. 1980.

[3] R. Sternberg, *Cognitive Psychology*, 3rd ed. CA: Thomson Wadsworth, Aug. 2003.

[4] K. Friston, J. Kilner, and L. Harrison, "A free energy principle for the brain," *Journal of Physiology, Paris*, vol. 100, no. 1-3, pp. 70–87, Sep. 2006.

[5] K. Friston, "The free-energy principle: a unified brain theory?" *Nature Reviews Neuroscience*, vol. 11, no. 2, pp. 127–138, Feb. 2010.

[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] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free energy principle," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 41–52, 2012.

[8] 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, Jul. 2005.

[9] 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.

[10] C. Chou and K. Liu, "Colour image compression based on the measure of just noticeable colour difference," *IET Image Processing*, vol. 2, pp. 304–322, 2008.

[11] 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.

[12] E. Peli and J. Lubin, "A visual discrimination model for imaging system design and evaluation," in *Vision Models For Target Detection And Recognition*. World Scientific, May 1995, pp. 245–283.

[13] 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.

[14] A. N. Netravali and B. Haskell, *Digital Pictures: Representation and Compression*. New York: Plenum, 1988.

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

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

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

[18] D. C. Knill and R. Pouget, "The bayesian brain: the role of uncertainty in neural coding and computation," *Trends in Neuroscience*, vol. 27, pp. 712—719, 2004.

[19] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[20] D. Gao, S. Han, and N. Vasconcelos, "Discriminant saliency, the detection of suspicious coincidences, and applications to visual recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 6, pp. 989–1005, 2009.

[21] M. Vasconcelos and N. Vasconcelos, "Natural image statistics and Low-Complexity feature selection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 228–244, Feb. 2009.

[22] N. Ponomarenko, V. Lukin, A. Zelensky, K. Egiazarian, M. Carli, and F. Battisti, "Tid2008 - a database for evaluation of full-reference visual quality assessment metrics," *Advances of Modern Radioelectronics*, vol. 10, pp. 30–45, 2009.

[23] "Method for the subjective assessment of the quality of television pictures," ITU, Document ITU-R BT.500-11, 2002, geneva, Switzerland.

[24] 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– 752, Jun. 2005.