

# Enhanced Just Noticeable Difference Model for Images with Pattern Complexity

Jinjian Wu, Leida Li, Weisheng Dong, Guangming Shi, Senior Member, IEEE  
Weisi Lin, Fellow, IEEE and C.-C. Jay Kuo, Fellow, IEEE

**Abstract**—The just noticeable difference (JND) in an image, which reveals the visibility limitation of the human visual system (HVS), is widely used for visual redundancy estimation in signal processing. To determine the JND threshold with the current schemes, the spatial masking effect is estimated as the contrast masking, and this cannot accurately account for the complicated interaction among visual contents. Research on cognitive science indicates that the HVS is highly adapted to extract the repeated patterns for visual content representation. Inspired by this, we formulate the pattern complexity as another factor to determine the total masking effect: the interaction is relatively straightforward with limited masking effect in a regular pattern, and is complicated with strong masking effect in an irregular pattern. From the orientation selectivity mechanism in the primary visual cortex, the response of each local receptive field can be considered as a pattern; therefore, in this work, the orientation that each pixel presents is regarded as the fundamental element of a pattern, and the pattern complexity is calculated as the diversity of the orientation in a local region. Finally, taking both pattern complexity and luminance contrast into account, a novel spatial masking estimation function is deduced, and an improved JND estimation model is built. Experimental results on comparing with the latest JND models demonstrate the effectiveness of the proposed model, which performs highly consistent with the human perception. The source code of the proposed model is publicly available at <http://web.xidian.edu.cn/wjj/en/index.html>.

**Index Terms**—Just Noticeable Difference, Spatial Masking, Pattern Complexity, Orientation Selectivity Mechanism

## I. INTRODUCTION

It is well known that the visual sensitivity of the human visual system (HVS) is limited. The HVS can only sense the content change which is larger than a certain threshold, and such a threshold is termed as just noticeable difference (JND) [1, 2]. Since the JND describes the visibility of the HVS on visual contents, its modeling is widely used for visual redundancy estimation in image/video coding and transmission [3], information hiding [4], visual quality assessment [5], and so on.

Jinjian Wu, Weisheng Dong, and Guangming Shi are with School of Electronic Engineering, Xidian University, Xi'an, China. E-mail: jinjian.wu@mail.xidian.edu.cn; wsdong@mail.xidian.edu.cn; gmshi@xidian.edu.cn.

Leida Li is with School of Information and Electrical Engineering, China University of Mining and Technology, Xuzhou, China. E-mail: lileida@cumt.edu.cn.

Weisi Lin is with the School of Computer Engineering, Nanyang Technological University, Singapore. E-mail: wslin@ntu.edu.sg.

C.-C. Jay Kuo is with Department of Electrical Engineering, University of Southern California, Los Angeles, USA. E-mail: cckuo@sipi.usc.edu.

This work was supported by the Major State Basic Research Development Program of China (973 Program, No.2013CB329402), the National Natural Science Foundation of China (Nos. 61401325, 61632019, 61379143, 61621005), and “the Qing Lan Project” of Jiangsu Province.

A number of JND models have been proposed during the past decade. According to the domain for the JND threshold being calculated, the existing models can be divided into two categories. One category is the subband-domain JND models, for which an image is firstly transferred into a subband domain (e.g., discrete cosine transformation (DCT) domain) [6–8]. The other category is the pixel-wise JND models, for which the JND is directly calculated on the original visual content [1, 9–12]. Since the pixel-wise JND does not need the transformation to subbands, it is more convenient and cost-effective to provide a JND profile for many applications, such as image/video enhancement, quality assessment, and motion estimation [7, 9, 10].

For pixel-wise JND models, the masking effects from the background luminance and the edge contrast are usually considered. In [1], an early JND model (Chou et al.’s) was built with luminance adaptation and contrast masking. Since Chou et al.’s model overestimates the JND threshold for the edge region, a canny edge protector was employed in [9] to extract the edge region. In order to estimate the JND threshold of the texture region, an image was decomposed into texture and non-texture portions for JND estimation, respectively [10]. Moreover, the color information was considered for JND estimation in [4]. Recently, Wu et al. modeled the fact that the HVS is insensitive to the irregular visual content [11], and introduced a structure uncertainty based JND model [13].

In this work, we investigate into more detail on the mechanisms of the human visibility. The work on cognitive science states that the HVS presents an active internal generative mechanism [14], with which the HVS is good at summarizing rules of an input scene and is highly adapted to extract the repeated visual contents for understanding [15, 16]. As an invariant feature of images, a visual pattern represents the repeatable components of visual contents [17, 18]. According to the input visual contents and the priori knowledge, the HVS will automatically encode visual patterns for content prediction [19]. Generally, a regular image patch presents an intuitional organization rule, which can be encoded with a simple pattern; while an irregular image patch presents an obscure organization rule, which requires a complex pattern for encoding. Moreover, the interaction within a regular pattern is straightforward, and the visual masking in such a pattern is weak; while the interaction within an irregular pattern is complicated, which involves with strong visual masking effect, as illustrated in [13]. Thus, the complexity of a visual pattern is an effective measurement for the extent of visual masking.

Since a visual pattern changes with the contents of an



Fig. 1: Active prediction of the HVS for scene perception, where (a) is the original image and (b) is the blocked image.

input image, how to calculate the complexity of a visual pattern is a research problem. It has been shown that the HVS presents obvious orientation selectivity in the primary visual cortex, within which the orientation information is extracted for structure representation [20, 21]. Further studies on visual patterns indicate that each pattern is composed with several elements [17]: the more different elements that a pattern has, the more complex that a pattern presents. In general, a complex pattern contains more different orientations, and a simple pattern has limited distribution of orientations. Thus, we calculate the pattern complexity as the distribution of orientations.

With the formulated pattern complexity, a novel spatial masking function is deduced in this work. With the proposed spatial masking function, the masking effect of different visual contents is effectively estimated. Finally, an improved JND model in pixel domain is proposed for more efficient and effective human visibility threshold estimation.

The rest of this paper is organized as follows. The motivation of the formulation for pattern complexity on visual sensitivity is firstly illustrated in Section II. Next, the computational model for the pattern complexity is built in Section III. With Section IV, an improved pixel-wise JND model is proposed. The performance of the proposed JND model is demonstrated in Section V. Finally, the conclusion is drawn in Section VI.

## II. MOTIVATION OF RESEARCH

In this section, we firstly give a brief introduction about the visual perception of the HVS, especially the internal generative mechanism (IGM) for visual content prediction. Next, visual patterns for scene perception are analyzed, and the idea of pattern complexity on visual information perception is illustrated. Finally, the method to determine the pattern complexity is demonstrated.

The human brain possesses an IGM for visual content perception [14]. With the help of the prior knowledge, the HVS optimizes its processing to minimize the content uncertainty and actively predicts an input scene [15]. As an example, when we looking at the *DOG* image as shown in Fig. 1, even though parts of the contents are blocked, we can still fully understand Fig. 1(b): 1) when a part of the stream of water is blocked with  $B_1$ , the HVS knows that the stream of water is always continuous, and according to such a rule it can predict the blocked information according to the other parts of the stream of water; 2) when a part of the grassland is

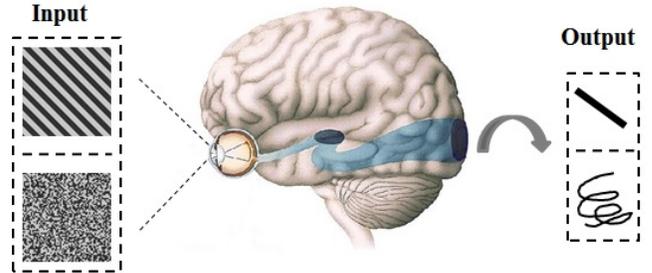


Fig. 2: Active prediction on different patterns with different complexities.

blocked with  $B_2$ , with the help of the human common sense and the correlation with its surrounding content (i.e., a similar rule), the HVS can imagine the contents under block  $B_2$ . This is an illustration that the HVS is adaptive to summarize the organization rule (regularity) of input visual contents for active prediction [22, 23].

Visual patterns represent the repeatable visual components, and each pattern is with a discernible regularity of the input scene [24]. In other words, each pattern represents a kind of organization rule that an input scene possesses. The HVS is good at extracting these patterns for visual content understanding [19].

Furthermore, different patterns have different compositions and with different complexities for visual perception. As an example, the orderly pattern in Fig. 2 (the first input pattern) is composed by oblique bars, namely, such a pattern possesses only one element. For such an orderly pattern with low complexity, the HVS is easy to summarize its regularity and can fully understands its visual contents. When it comes to a disorderly pattern as shown in Fig. 2 (the second input pattern), which is composed by uncertain elements (random noise), the HVS can hardly summarize its regularity and knows little about the organization rule of such a complex pattern. Therefore, the more different elements that a pattern possesses, the higher complexity that it has [17].

Pattern complexity is directly related to the visual masking effect. For a pattern with low complexity, the interactions among elements in such a pattern are straightforward. As the regular pattern shown in Fig. 2, the HVS can easily extract the oblique bars, and the masking effect on such kind of pattern is weak. While for a pattern with high complexity, the interactions among elements are complicated. As for an irregular pattern (as shown in Fig. 2), even for a case where the contents of the pattern is reorganized, the HVS can hardly sense the change. That is because with extremely high complexity, the HVS knows nothing about its organization rule. Therefore, higher pattern complexity corresponds to stronger visual masking.

During the past decades, many models have been designed to encode the visual patterns [25, 26]; however there exists no rigorous definition for elements in different patterns. As an example, when we looking at the pattern  $P_1$  in Fig. 3, it is easy to find out the element that  $P_1$  contains, namely,  $P_1$  is composed by three triangles. When we change  $P_1$  with a

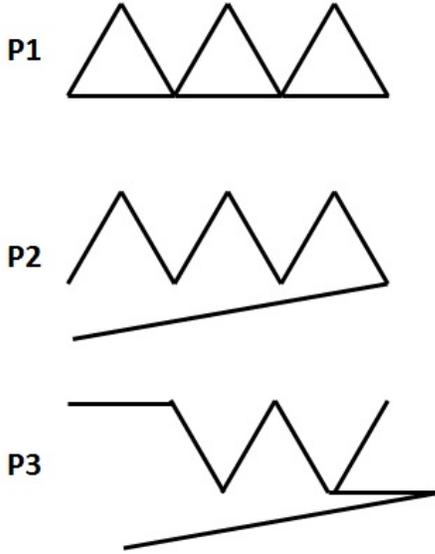


Fig. 3: Challenge on describing the elements for different patterns (i.e.,  $P_1$ ,  $P_2$ , and  $P_3$ ).

rotation on the bottom line, a new pattern  $P_2$  is created. For  $P_2$ , a triangle is no longer be its composed element, and the sawtooth and the oblique line become the new elements. When we continue change  $P_2$  into  $P_3$ , the elements for  $P_2$  are also no longer be the elements for  $P_3$ . Therefore, it is a challenge to describe the elements for different patterns.

Neuroscience research on visual cognition indicates that the primary visual cortex possesses a substantial orientation selectivity mechanism [27, 28]. When receiving different visual stimuli from the thalamic inputs, cortical cells present different orientations according to the stimuli that they received [29]. Cortical cells with different orientations present different interactions, and there are two distinguishing interactions among cortical cells in a local receptive field: cells with similar orientations are more likely to response as excitation, while cells with dissimilar orientations are more likely to response as inhibition [30]. The arrangement of these excitatory and inhibitory cells in a local receptive filed causes the orientation selectivity [31]. Moreover, the orientation selectivity mechanism is highly related to the visual pattern extraction for scene perception [32], with which the HVS is extremely sensitive to edge/boundary regions.

Inspired by the orientation selectivity mechanism, we consider the visual stimuli in a local receptive field as a pattern. Generally, if cortical cells in a local receptive field receive similar stimuli, these cells will represent similar orientations and the interactions among them are unitary (with excitatory interaction). From the perspective of visual contents, patterns with homogenous contents cause similar responses (i.e., similar orientations) in the cortical cells, and such kinds of patterns are simple and regular, and the masking effect for such a pattern is weak. While for cortical cells receiving dissimilar stimuli, cells in the local receptive field will represent different orientations. For such a condition, the interactions among cells are much complex and there exist extensive inhibitions. From the perspective of visual contents, patterns with irregular

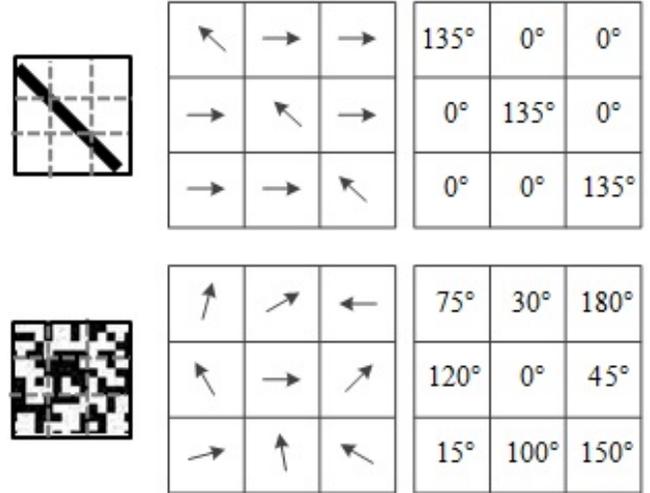


Fig. 4: Pattern complexity with orientation consideration

contents induce different responses (i.e., different orientations) in the cortical cells, and such kinds of patterns are irregular and possess strong masking effect.

Therefore, we can take the orientations that the cells response in a local receptive field as the fundamental element that a pattern possesses, and the dissimilarity among orientations that a pattern possess is highly related to the pattern complexity. As an example, the pattern of the oblique bar shows in the first row of Fig. 4 is regular and easy to be perceived. By analyzing its orientation composition, we can see that there are only two kinds of orientations, i.e.,  $0^\circ$  (a homogenous region) and  $135^\circ$  (the bar) for each region indicated by the dashed lines. While for the irregular pattern shown in the bottom row of Fig. 4, it possesses many different orientations.

### III. PATTERN COMPLEXITY

As introduced in the previous section, visual patterns, which present the discernible regularity of visual contents, play an important role in visual content prediction; the orientation that a cortical cell presents is directly related to the stimulus that it receives from the thalamic input. For any image  $\mathcal{F}$ , we can take a local region  $\mathcal{R} \in \mathcal{F}$  with radius  $r$  as a local receptive field, and the orientation for each pixel  $x \in \mathcal{R}$  relates to the gradient. Thus, in this work, the gradient direction of  $x$  is taken as its orientation  $\theta$ ,

$$\theta(x) = \arctan \frac{G_v(x)}{G_h(x)}, \quad (1)$$

where  $G_v$  means the gradient change along the vertical direction, and  $G_h$  means that along the horizontal direction. The changes along the two directions can be calculated with the edge filters,

$$G_v = \mathcal{F} * K_v, \quad (2)$$

$$G_h = \mathcal{F} * K_h, \quad (3)$$

where  $K_v$  and  $K_h$  are the edge kernels along vertical and horizontal directions, and in this work we choose the Prewitt kernels as shown in Fig. 5.

$\frac{1}{3}$	0	$-\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$
$\frac{1}{3}$	0	$-\frac{1}{3}$	0	0	0	0
$\frac{1}{3}$	0	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$

Fig. 5: The Prewitt edge kernels along vertical and horizontal directions

The complexity of each pattern  $\mathcal{R}$  is directly related to the interactions among these  $\theta(x)$ . Moreover, if a pattern  $\mathcal{R}$  possesses more dissimilar  $\theta(x)$ , it is more likely to be an irregular pattern. For such kind of patterns, there exist complicated interactions and presents strongly masking effects. When a pattern  $\mathcal{R}$  possesses more similar  $\theta(x)$ , it should be a regular pattern and its masking effect is weak. Thus, the complexity of a pattern can be regarded as the diversity of all  $\theta(x) \in \mathcal{R}$ .

As an effective representation of the distribution, the histogram of  $\theta(x)$  is employed to describe the diversity of a pattern. In order to build a histogram, all  $\theta(x)$  require to be quantified into a limited bin number  $N$ . Subjective masking experiments indicate that the interaction between two oriented gratings is quite strong when their orientation difference is smaller than  $12^\circ$ , and the interaction is rapidly decreasing when the difference is larger than that [28]. Thus, we quantify  $\theta(x)$  with the interval  $T=12^\circ$  and acquire the quantified  $\hat{\theta}(x)$  for the histogram  $\mathcal{H}$  generation,

$$\mathcal{H}_k(x) = \sum_{x \in \mathcal{R}(x)} \delta(\hat{\theta}(x), k), \quad (4)$$

where  $\delta(\cdot)$  is a pulse function, for which

$$\delta(\hat{\theta}(x), k) = \begin{cases} 1 & \text{if } \hat{\theta}(x) = k \\ 0 & \text{if } \hat{\theta}(x) \neq k \end{cases} \quad (5)$$

Since a regular pattern possesses elements with similar  $\theta(x)$ , the diversity of its corresponding histogram is small. While for an irregular pattern which possesses elements with dissimilar  $\theta(x)$ , the diversity of its corresponding histogram is large. In other words, the histogram for a regular pattern is sparse, while that for an irregular pattern is dense. Therefore, the pattern complexity  $\mathcal{C}_p$  is calculated as the sparsity of its corresponding histogram,

$$\mathcal{C}_p(x) = \sum_{k=1}^N \|\mathcal{H}_k(x)\|_0, \quad (6)$$

where  $\|\cdot\|_0$  denotes the  $L_0$  norm. An example of pattern complexity calculation is shown in Fig. 6. For a regular image (Fig. 6 (a)), its corresponding complexity is low (shown as Fig. 6 (c) with low intensity values). And for an irregular image (Fig. 6 (b)), its complexity (as shown in Fig. 6 (d)) is much larger than that of the regular one.

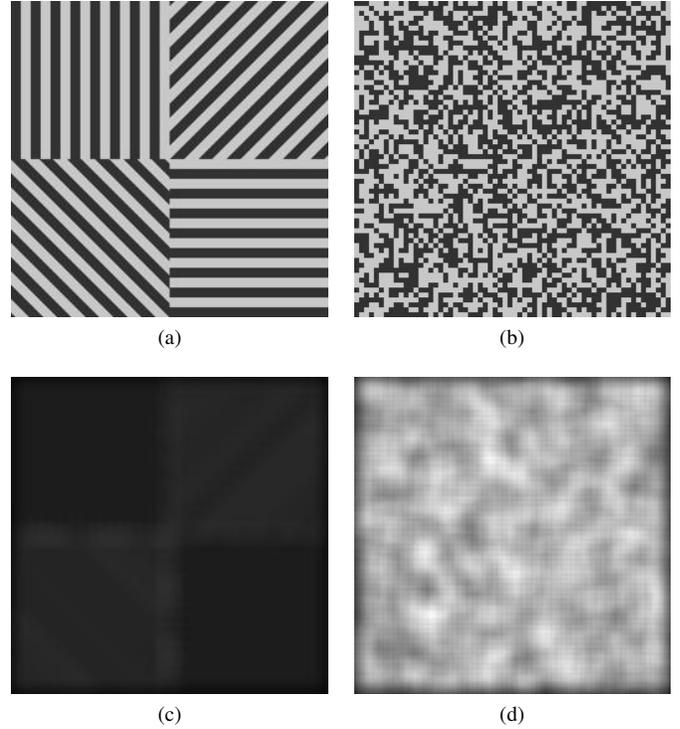


Fig. 6: Pattern complexity demonstration: (a) is a regular image and (c) is its corresponding complexity map; (b) is an irregular image and (d) is its corresponding complexity map.

#### IV. JND MODELING

The pattern masking effect is determined by both luminance contrast and pattern complexity. For example, a bar region possesses high luminance contrast (as shown in Fig. 6 (a)), whose spatial masking effect is much stronger than that of a uniform region. But when comparing with an irregular region (as shown in Fig. 6 (b)) with both high luminance contrast and large pattern complexity, the spatial masking effect of the bar region is much weaker than that of the irregular region. Therefore, we take both luminance contrast  $\mathcal{C}_l$  and pattern complexity  $\mathcal{C}_p$  into account for the pattern masking estimation.

The pattern complexity  $\mathcal{C}_p$  has been acquired with Eq. (6), and the luminance contrast  $\mathcal{C}_l$  can be calculated as the gradient magnitude,

$$\mathcal{C}_l(x) = \sqrt{G_v^2 + G_h^2}, \quad (7)$$

where  $G_v$  and  $G_h$  are the gradient magnitudes along the vertical and horizontal directions, as used in the previous section.

The pattern masking effect increases with the increase of the luminance contrast  $\mathcal{C}_l$  and the pattern complexity  $\mathcal{C}_p$ . With  $\mathcal{C}_l$  from eq. (9) and  $\mathcal{C}_p$  from eq. (6), the pattern masking effect is modeled as,

$$\mathcal{M}_P = f(\mathcal{C}_l) \cdot f(\mathcal{C}_p). \quad (8)$$

With the subjective viewing tests on pattern masking, we have found that the increasing rate of  $\mathcal{M}_P$  decreases with the

increase of  $\mathcal{C}_l$  ( $\mathcal{C}_p$ ). Moreover, the increasing effect of  $\mathcal{C}_l$  on eq. (8) follows logarithmic form,

$$f(\mathcal{C}_l) = \log_2(1 + \mathcal{C}_l). \quad (9)$$

And the increasing effect of  $\mathcal{C}_p$  on eq. (8) follows the nonlinear transducer as that of the gain control in [33]. Thus,  $f(\mathcal{C}_p)$  can be modeled as,

$$f(\mathcal{C}_p) = a_1 \cdot \frac{\mathcal{C}_p^{a_2}}{\mathcal{C}_p^2 + a_3^2}, \quad (10)$$

where  $a_1$  is a constant of proportion,  $a_2$  is an exponential parameter which determines the shape of the nonlinear transducer (a larger  $a_2$  corresponds to a faster gain), and  $a_3$  is a small constant which is used to avoid the denominator as zero. With subjective test, we set  $a_1=0.8$ ,  $a_2=2.7$ , and  $a_3=0.1$  in this work.

The contrast masking function  $\mathcal{M}_C$  is fitted with the data of visibility thresholds from masking experiment on luminance contrast [1, 13], which is calculated as,

$$\mathcal{M}_C = 0.115 \times \frac{\alpha \cdot \mathcal{C}_l^{2.4}}{\mathcal{C}_l^2 + \beta^2}, \quad (11)$$

where the two parameters are set as  $\alpha=16$  and  $\beta=26$  in [13] (by fitting Eq. (11) with the subjective visibility thresholds from contrast masking experiment).

We take both the contrast masking and pattern masking into account for spatial masking. In general, the two types of masking effects are concurrent, while one of them may play a dominant role in some situations. As an example, for a regular edge region, the contrast masking is the dominating effect; while for irregular region, the pattern masking becomes the dominating one. In other words, the stronger one plays the dominating role, and therefore we take the stronger one to represent the final spatial masking effect. With the contrast masking from Eq. (11) and the pattern masking from Eq. (8), the total spatial masking effect is calculated as,

$$\mathcal{M}_S(x) = \max\{\mathcal{M}_P(x), \mathcal{M}_C(x)\}. \quad (12)$$

Furthermore, the HVS presents different sensitivity to different background luminance, and thus luminance adaptation needs to be considered as well. It is well known that the sensitivity of the human eyes is low in the dark environment, and increases with a suitable light condition. Therefore, the visibility thresholds for different luminance backgrounds are different. The visibility threshold of the luminance adaptation  $\mathcal{L}_A$  is modeled as [1]

$$\mathcal{L}_A(x) = \begin{cases} 17 \times (1 - \sqrt{\frac{\mathcal{B}(x)}{127}}) & \text{if } \mathcal{B}(x) < 127 \\ \frac{3 \times (\mathcal{B}(x) - 127)}{128} + 3 & \text{if } \mathcal{B}(x) \geq 127 \end{cases}, \quad (13)$$

where  $\mathcal{B}(x)$  is the background luminance, which is calculated as the mean luminance value of a surrounding region that  $x$  located.

Finally, taking both the spatial masking effect and the luminance adaptation into account, the JND threshold for each pixel is calculated using the nonlinear additivity model for masking (NAMM) [9] as

$$\mathcal{T}_{\text{JND}}(x) = \mathcal{L}_A(x) + \mathcal{M}_S(x) - C \cdot \min\{\mathcal{L}_A(x), \mathcal{M}_S(x)\}, \quad (14)$$

where  $C$  is the gain reduction parameter determined by the overlapping between  $\mathcal{L}_A$  and  $\mathcal{M}_S$ , and here we set  $C=0.3$  (the same as in [9]).

## V. EXPERIMENTAL RESULT

In this section, the efficiency of the novel pattern complexity based spatial masking model (using Eq. (12)) is firstly illustrated by the comparison with the case of evaluation of contrast masking alone (using Eq. (11)). Next, the proposed JND model is compared with the latest JND models to demonstrate the accuracy and efficiency of these estimated visibility thresholds. Then, a subjective viewing test has been conducted to further verify the effectiveness of the proposed model. Finally, the proposed model is applied to image compression for perceptual redundancy reduction.

### A. Pattern Complexity based Spatial Masking Evaluation

Since the spatial masking effect is a complicated psychological and physiological phenomenon, existing JND models usually take the contrast masking as a simple representation of spatial masking. In this work, we suggest to take both pattern masking and contrast masking into account for spatial masking estimation.

In order to illustrate the effectiveness of the novel pattern complexity based spatial masking model, a comparison example against the contrast masking on a concept image (composed by both regular and irregular patches) is given in Fig. 7. The visibility threshold  $\mathcal{T}_v$  is firstly calculated with the contrast masking function and the spatial masking function (using Eqs. (11) and (12) respectively). Then, random noise is injected into the image with the guidance of  $\mathcal{T}_v$ ,

$$\hat{\mathcal{F}}(x) = \mathcal{F}(x) + \mathcal{E} \cdot R(x) \cdot \mathcal{T}_v(x) \quad (15)$$

where  $\hat{\mathcal{F}}$  is the contaminated image,  $R(x)$  is the random noise controller which takes  $-1$  or  $+1$  randomly, and  $\mathcal{E}$  is the noise level adjuster which keeps the injected noise from different models at a same level. For the pattern complexity based model, we set the pattern size as  $3 \times 3$  for the local receptive field  $\mathcal{R}$  during all of the experiments.

With the help of Eq. (15), the same level of noise (with PSNR=26.65 dB) is injected into the concept images under the guidance of contrast masking and the proposed spatial masking, as shown in Fig. 7 (a) and (b). Though with a same level of noise, the perceptual qualities of the two contaminated images are obviously different: Fig. 7 (b) has a better quality than Fig. 7 (a). With further analysis, we have found that the model with contrast masking alone mainly responses to the edge region, and its response amplitude is determined by the edge height. The response map of the contrast masking function is shown in Fig. 7 (c), for which all of the edge regions are highlighted. Moreover, the edge height for all of the edge regions in the concept image are the same, and thus the response amplitudes for them are the same (with a same intensity values for the response map as shown in Fig. 7 (c)). In the proposed spatial masking model, both luminance contrast and pattern complexity are considered. As a result, the irregular textural region (as the left part shown in Fig. 7

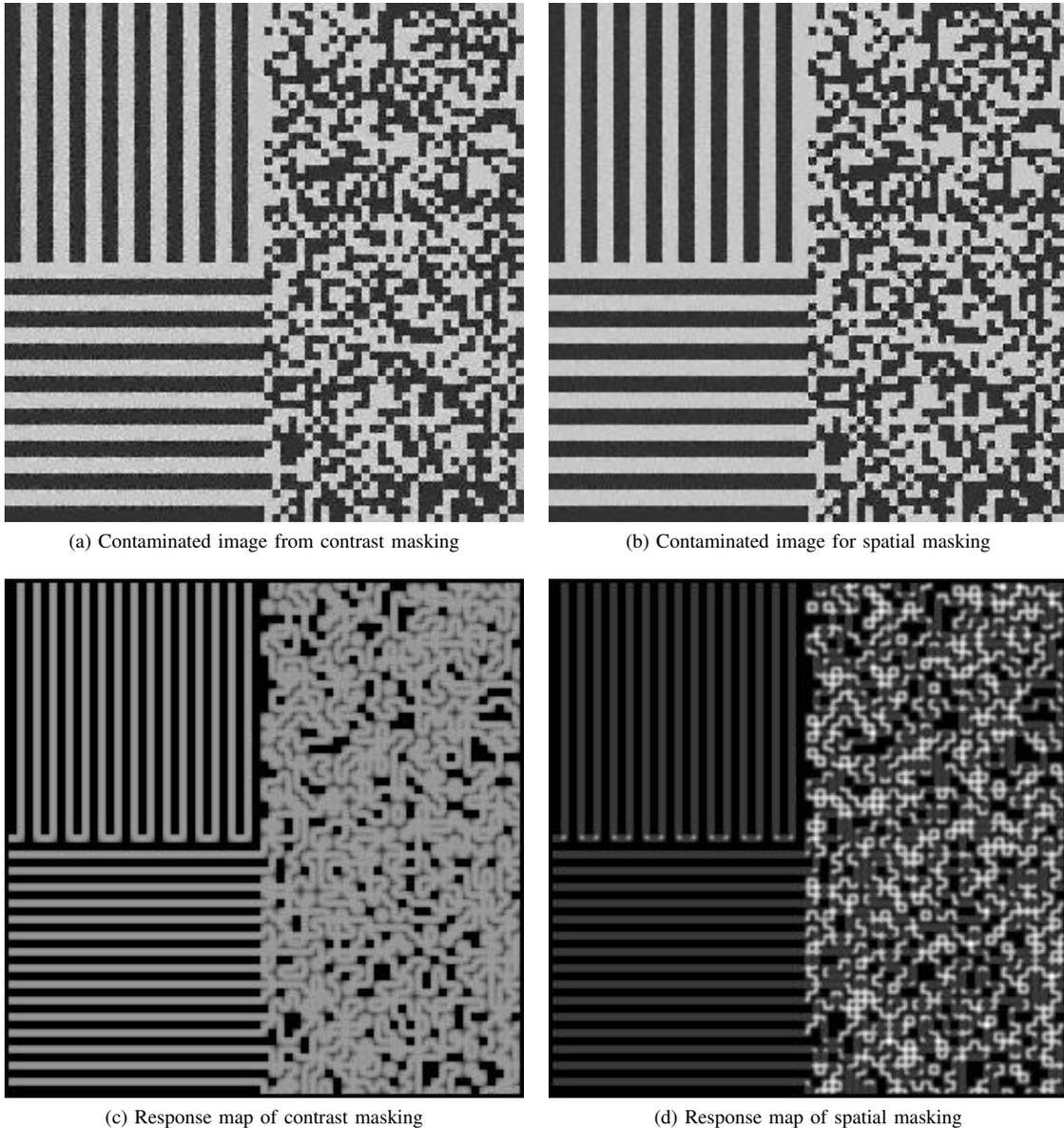


Fig. 7: Comparison between the contrast masking and the proposed spatial masking on a concept image (the size of the image is  $256 \times 256$ , and the injected noise in the two contaminated images are the same: with PSNR=26.65dB)

(b), which possesses both large edge height and high pattern complexity) has much larger response than the edge region (as the right part shown in Fig. 7 (b)), as shown in Fig. 7 (d). With the analysis above, we can see that though with the same level of noise, the proposed spatial masking model guides less noise into the edge region, where the HVS is highly sensitive to and can easily find out the noise, while it guides more noise into the irregular texture region, where the HVS is insensitive due to the strong masking effect.

To further demonstrate the effectiveness of the pattern complexity on visual masking, a natural *horse* image is adopted. With the guidance of the contrast masking and the proposed spatial masking, respectively, a same level of noise (with PSNR=26.65dB) is injected into the images. As shown in

Fig. 8 (c), the contrast masking model mainly guides the noise into the boundary of the horse (which possesses large edge heights), and guides limited noise into the grassland (which possesses small edge height but high pattern complexity). As a result, the boundary of the horse is obviously distorted, as shown in Fig. 8 (a). The proposed spatial masking model guides more noise into the grassland due to the effect of pattern masking, and less noise into the boundary of the horse, as shown in Fig. 8 (d). By comparing the two contaminated images, though they have a same level of noise, we can see that Fig. 8 (b) has a better quality than Fig. 8 (a).

The analysis on both the concept and the natural images above demonstrates that pattern complexity is important in determining the visibility threshold, and the proposed spatial



Fig. 8: Comparison between the contrast masking and the proposed spatial masking on a natural *horse* image (the size of the image is  $200 \times 150$ , and the injected noise in the two contaminated images are the same: with  $\text{PSNR}=26.65\text{dB}$ ).

masking model performs more consistently with the HVS than the case with contrast masking alone.

### B. JND model comparison

The proposed JND model is compared with four relevant JND models, i.e., Zhang et al.'s model (*Zhang2008*) [7], Wei et al.'s model (*Wei2009*) [8], Liu et al.'s model (*Liu2010*) [10], and Wu et al.'s model (*Wu2013*) [13], with the noise for different JND models being adjusted into a same level with the help of Eq. (15).

An example of comparison is shown in Fig. 9, where Fig. 9 (a) is the original image, and Fig. 9 (b)-(f) are the contaminated images by different JND models, with the noise level for all of the contaminated images at  $\text{PSNR} = 26.65\text{dB}$ . Though with a same level of noise, the perceptual quality for the five contaminated images are different. Since *Zhang2009* and *Wei2009* are subband-domain JND models, degradations on the DCT coefficients may cause ringing artifact at the edge regions. By overestimating the JND thresholds of the regions with high edge height, the two subband-domain JND models generate noticeable distortions on the edge regions, such as the boundaries of the building and the fencing as shown in Fig. 9 (b) and (c). In *Liu2010* (a pixel-wise model),

the image is firstly classified into textural and non-textural regions for separative JND estimation [10]. However, this method still has difficulty in classifying texture from non-texture, due to the nature of lack of appropriate definition of texture; moreover, texture is not the same as irregularity, and the masking for regular texture is also weak. As shown in Fig. 9 (d), though most of the boundaries of buildings are protected (with less noise injected), the fencing regions (with regular texture/structure) are highlighted with too much noise. As a result, the degradation on the fencing regions is obvious with *Liu2010*. In *Wu2013*, the disorder extent of of visual contents is firstly considered for masking estimation [13]. With a complicated procedure for disorder computation, *Wu2013* injects less noise into the boundaries of buildings, and more noise into the disorderly grassland. However, *Wu2013* still cannot accurately estimate the extend of disorder for the fencing regions, and injects too much noise into such regions to cause visible distortion, as shown in Fig. 9 (e). With the help of the pattern complexity, the proposed JND model can correctly highlight the irregular grassland and wall regions, and injects much noise into such irregular regions. At the meanwhile, the proposed model injects less noise into the boundaries of building and fencing. As a result, the noise in



(a) Original Image

(b) Zhang2008

(c) Wei2009



(d) Liu2010

(e) Wu2013

(f) Proposed

Fig. 9: JND comparison among the proposed model and the latest four JND models. The noise levels on the five contaminated images (with the size of  $256 \times 384$ ) are the same, with  $\text{PSNR} = 26.65\text{dB}$ .

TABLE I: Computational complexity analysis on different JND models

Models	<i>Zhang2008</i>	<i>Wei2009</i>	<i>Liu2010</i>	<i>Wu2013</i>	Proposed
Time	2.064s	9.022s	0.406s	3.560s	0.418s



Fig. 10: Image set for subjective viewing test, From top left to bottom right, they are named as I1-I12.

Fig. 9 (f) is almost invisible.

As can be seen, the proposed model injects much noise into the irregular regions where the HVS is insensitive to, and little noise into the regular (with high edge height but low pattern complexity) regions where the HVS is sensitive to. With a same level of noise, the proposed model enables a better perceptual quality than the other compared JND models, as shown in Fig. 9.

Furthermore, the computational complexity of these JND models is also analyzed. All of the five JND models are running with an unoptimized MATLAB programs on a personal computer (2 cores Intel i7-3790K CPU at 4.00GHz and 8GB RAM). A testing image (the *horse* image, as shown in Fig. 9 (a)) with size of  $256 \times 384$  is chosen in the comparing experiment. The execution time for all of the five models on this image is listed in Table I. As can be seen, the proposed model (with 0.418s) is slightly slower than *Liu2010* (with 0.406s), but much faster than the other three JND models.

### C. Subjective Viewing Test

For a more comprehensive analysis, a set of images (as shown in Fig. 10, for which I1-I6 are often used in JND estimation [9–11] and I7-I12 are often used in image quality assessment [34] are selected for comparison among different

TABLE II: The standard of evaluation for quality comparison between two images

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

JND models. A subjective viewing test experiment on perceptual quality comparison has been conducted. During the experiment, two JND noise-contaminated images on a same scene (one is with the guidance of the proposed JND model, and the other one is with the guidance of another existing JND model) are randomly juxtaposed on the right or left part of a screen. Subjects are asked to decide the image which has better quality, and how much better it is (the standard for evaluation is shown in Table II). The viewing environment and the viewing condition are set with the guidance of ITU-R BT.500-11 standard [35].

In this experiment, forty nine subjects (all of them have normal or corrected-to-be-normal eyesight) are invited. The statistical results on the testing image set (I1-I12) are listed in Table III, where ‘Mean’ refers to the mean of the quality values given by the subjects and ‘Std’ refers to their standard deviation.

By comparing with the two subband-domain JND models (i.e., *Zhang2008* and *Wei2009*), we can see that the proposed models outperforms them (with positive mean values as the second and the fourth columns in Table III) on all of these images, especially for I6, I11, and I12 (the three images are with both regular and irregular regions as shown in Fig. 10). For comparison with *Liu2010*, the proposed model performs better on most of the images, especially for I4, I6 and I12 (all of these three images are with a large irregular region); the proposed model performs just a little bit worse than *Liu2010* on I2 and I5 (with negative values -0.33 and -0.10). For the comparison with *Wu2013*, the two models performs similarly (since both models take the disorderness/irregularity into account) except for I2 and I3 (the proposed method has better performance on these 2 images than *Wu2013*). With further analysis we have found that although *Wu2013* uses a self-similarity procedure for content complexity computation, it is sensitive to small disturbance (it highlights the freckle of the faces in I2 and I3).

Besides, standard deviation values (i.e., ‘Std’) are a little bit large for most of the quality comparison results as listed in Table III, which means the results from all of the subjects are not so consistent. Actually, it is extremely hard to get highly consistent quality assessment results from different people in this test, the reason may as follows: 1) During the subjective viewing test, the level of JND noise is no so high, and the quality degradation for both contaminated images are always not so obvious. As a result, it is not so easy for the subjects to give a confident quality score. 2) All of the subjects invited for this experiments are naive (rather than expert) in image processing, and they will judge the quality according to their own hobbies.

In general, as the last row shown in Table. III, the proposed model performs better than *Zhang2008*, *Wei2009*, *Wei2009* and *Wu2013* (with average scores of 1.33, 1.17, 0.75, and

TABLE III: Quality comparison between two contaminated images from the proposed model and one of the four latest JND models

Images	Methods		VS. Zhang2008		VS. Wei2013		VS. Liu2010		VS. Wu2013	
	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
I1	1.03	0.72	1.08	0.86	0.12	1.11	0.04	0.21		
I2	1.68	0.32	1.42	0.59	-0.33	0.82	1.35	0.61		
I3	0.83	1.37	0.65	1.21	0.00	1.11	1.53	0.84		
I4	0.54	1.16	0.73	0.94	1.85	0.88	-0.17	0.96		
I5	0.96	0.84	0.88	1.25	-0.10	0.97	0.02	1.64		
I6	1.89	0.67	1.43	0.51	1.57	0.93	0.38	1.17		
I7	0.96	0.85	1.08	0.77	0.42	1.21	0.12	1.35		
I8	1.38	0.25	1.06	0.39	1.25	0.67	0.48	1.33		
I9	0.85	1.31	0.77	1.08	0.25	1.25	0.08	1.06		
I10	1.59	0.68	1.33	0.83	1.25	0.95	0.12	0.88		
I11	1.95	0.77	1.64	0.93	1.19	0.48	0.33	1.25		
I12	2.32	0.85	1.96	0.91	1.54	0.84	0.25	1.35		
Average	1.33	-	1.17	-	0.75	-	0.37	-		

0.37, respectively). Therefore, the proposed pattern complexity based JND model is more consistent with the subjective perception, because it is able to more accurately represent the sensitivity of the HVS on different image regions.

#### D. Perceptual Redundancy Reduction

Since the JND threshold reveals the visibility of the HVS, it is often used to preprocess the visual signal for optimizing compression [9, 11]. In this experiment, the proposed JND model is applied to visual redundancy reduction during compression. Generally, the smoothing operation can reduce the signal variance, which makes image compression easier. However, blindly smoothing operations will always jeopardize image quality. Thus, we use the JND threshold for smoothing guidance, with which we can optimize the smoothing operation with limited perceptual distortion. For a given image  $\mathcal{F}$ , the JND guided smoothing operation for each pixel is as

$$\tilde{\mathcal{F}}(x) = \begin{cases} \mathcal{F}(x) + \mathcal{T}_{\text{JND}}(x), & \text{if } \mathcal{F}(x) - \bar{\mathcal{F}}_B < -\mathcal{T}_{\text{JND}}(x), \\ \bar{\mathcal{F}}_B, & \text{if } |\mathcal{F}(x) - \bar{\mathcal{F}}_B| \leq \mathcal{T}_{\text{JND}}(x), \\ \mathcal{F}(x) - \mathcal{T}_{\text{JND}}(x), & \text{if } \mathcal{F}(x) - \bar{\mathcal{F}}_B > \mathcal{T}_{\text{JND}}(x), \end{cases} \quad (16)$$

where  $\bar{\mathcal{F}}_B$  is the mean value of the block that  $\mathcal{F}(x)$  belongs to during the compression (e.g., the divided  $8 \times 8$  block which  $\mathcal{F}(x)$  located at during JPEG compression).

With the help of Eq. (16), the visual redundancy can be removed to optimize the compression. An example is shown in Fig. 11, for which the airplane image is directly compressed (with compression QP=1) by the JPEG algorithm (as shown in Fig. 11-a) and is preprocessed with JND threshold for JPEG compression (as shown in Fig. 11-b). As can be seen, though less bit rate is required (Fig. 11-a and b with 0.687 bpp and 0.589 bpp, respectively), the perceptual quality of Fig. 11-b is almost equal to Fig. 11-a.

In order to give a more comprehensive demonstration, the 12 images (as shown in Fig. 10) are chosen for JPEG compression comparison. Each image is with two types of compression: 1) directly compressed with the JPEG algorithm; and 2)

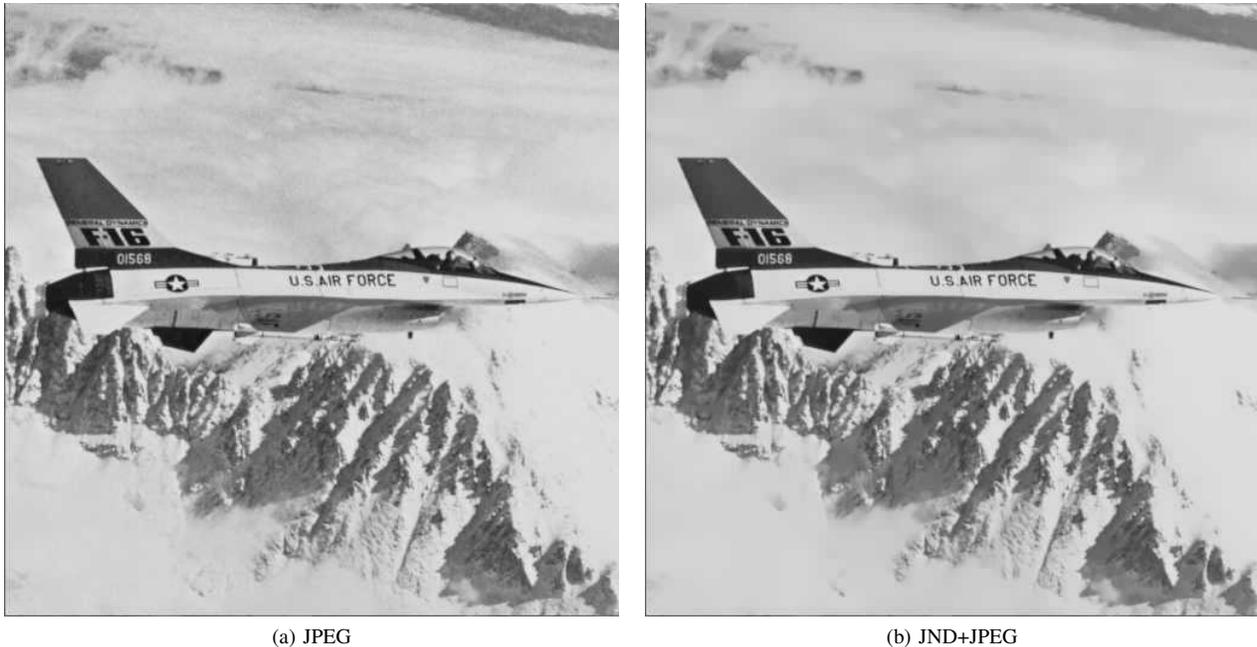
TABLE IV: Compression result comparison between JPEG algorithm and the JND+JPEG.

Image	JPEG (bpp)	JND+JPEG (bpp)	Bit Rate Saving	Subjective Quality	
				Mean	Std
I01	0.647	0.593	8.3%	0.812	1.033
I02	0.936	0.827	11.6%	0.353	1.342
I03	0.960	0.765	20.3%	0.137	1.458
I04	0.947	0.813	14.1%	-0.080	1.727
I05	1.258	1.035	17.7%	0.102	1.853
I06	0.916	0.721	21.3%	-0.137	1.887
I07	1.977	1.848	6.5%	1.125	0.857
I08	3.835	3.514	8.4%	0.237	1.137
I09	4.314	3.512	18.6%	-0.118	1.895
I10	4.971	4.125	17%	0.375	1.028
I11	2.930	2.569	12.3%	0.368	1.249
I12	3.543	3.021	14.7%	0.275	0.977
Average	-	-	14.3%	0.288	-

compressed with the guide of JND threshold (i.e., JND+JPEG) under a same compression QP. Then, a subjective quality compression test is designed to compare the compression results from the two types of compression (the setting of the environment follows the ITU-R BT.500-11 standard [35] as described in Subsection V-C). The comparison result is shown in Tab. IV. With the help of the JND guided smoothing operator, many bit rates can be saved by visual redundancy reduction (an average of 14.3% bit rate is saved). At the meantime, the qualities of these images are almost maintained (actually, the proposed model turns out a slightly better quality due to the reduction of disturbance with the smoothing operator).

## VI. CONCLUSION

In this work, we have introduced an improved pixel-wise JND estimation model for images based upon evaluation of pattern complexity. The pattern complexity is important in determining the visual signal masking; however, it has



(a) JPEG

(b) JND+JPEG

Fig. 11: The JND guided smoothing operation on the Airplane image (its size is  $512 \times 512$  for JPEG compression, where (a) is directly compressed with the JPEG algorithm (0.687 bpp) and (b) is preprocessed with the JND guided model for JPEG compression (0.589 bpp).

not been considered in the existing JND models before this work. The characteristic of the related visual perception, i.e., the active prediction on the visual pattern with the internal generative mechanism in the HVS, has been firstly illustrated. And then, by imitating the response of the local receptive field in the primary visual cortex, a visual pattern has been represented by the distribution of orientations in a local region. Finally, the pattern complexity has been computed for the interaction among different types of orientations.

The resultant spatial masking effect is determined by both luminance contrast and pattern complexity. Combining with the luminance adaptation, a new pixel-wise JND estimation model has been built. Experiments for the proposed model against the relevant existing JND models have demonstrated the accuracy and efficiency of the proposed model.

#### REFERENCES

- [1] 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.
- [2] W. Lin, *Computational Models for Just-noticeable Difference, Chapter 9 in Digital Video Image Quality and Perceptual Coding*. CRC Press, 2005.
- [3] H. R. Wu, A. Reibman, W. Lin, F. Pereira, and S. S. Hemami, "Perceptual visual signal compression and transmission," *Proceedings of The IEEE*, vol. 101, no. 9, pp. 2025–2043, 2013.
- [4] C. H. Chou and K. C. Liu, "A perceptually tuned watermarking scheme for color images," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2966–2982, Nov. 2010.
- [5] 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.
- [6] Y. Jia, W. Lin, and A. A. Kassim, "Estimating just-noticeable distortion for video," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 16, no. 7, pp. 820–829, July 2006.
- [7] 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.
- [8] Z. Wei and K. Ngan, "Spatio-temporal just noticeable distortion profile for grey scale image/video in dct domain," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 19, no. 3, pp. 337–346, Mar. 2009.
- [9] 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.
- [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 Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1648–1652, Nov. 2010.
- [11] 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.
- [12] S. Wang, L. Ma, Y. Fang, W. Lin, S. Ma, and W. Gao, "Just noticeable difference estimation for screen content

- images,” *IEEE Transactions on Image Processing*, 2016, accepted.
- [13] J. Wu, W. Lin, G. Shi, X. Wang, and F. Li, “Pattern masking estimation in image with structural uncertainty,” *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4892–4904, 2013.
- [14] K. Friston, “The free-energy principle: a unified brain theory?” *Nat Rev Neurosci*, vol. 11, no. 2, pp. 127–138, Feb. 2010.
- [15] 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.
- [16] J. Wu, W. Lin, G. Shi, L. Li, and Y. Fang, “Orientation selectivity based visual pattern for reduced-reference image quality assessment,” *Information Sciences*, vol. 351, pp. 18–29, Jul. 2016.
- [17] R. J. v. L. Peter A van der Helm, “Serial pattern complexity: irregularity and hierarchy,” *Perception*, vol. 21, no. 4, pp. 517–44, 1992.
- [18] C. Yuan, X. Sun, and R. Lv, “Fingerprint liveness detection based on multi-scale LPQ and PCA,” *China Communications*, vol. 13, no. 7, pp. 60–65, Jul. 2016.
- [19] P. A. Van Der Helm and E. L. J. Leeuwenberg, “Accessibility: A criterion for regularity and hierarchy in visual pattern codes,” *J. Math. Psychol.*, vol. 35, no. 2, pp. 151–213, May 1991.
- [20] R. Ben Yishai, R. L. Bar-Or, and H. Sompolinsky, “Theory of orientation tuning in visual cortex,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 92, no. 9, pp. 3844–3848, 1995.
- [21] D. Ferster and K. D. Miller, “Neural mechanisms of orientation selectivity in the visual cortex,” *Annual review of neuroscience*, vol. 23, pp. 441–471, 2000.
- [22] P. R. Roelfsema, A. van Ooyen, and T. Watanabe, “Perceptual learning rules based on reinforcers and attention,” *Trends in cognitive sciences*, vol. 14, no. 2, pp. 64–71, Feb. 2010.
- [23] Z. Pan, P. Jin, J. Lei, Y. Zhang, X. Sun, and S. Kwong, “Fast reference frame selection based on content similarity for low complexity HEVC encoder,” *Journal of Visual Communication and Image Representation*, vol. 40, Part B, pp. 516–524, Oct. 2016.
- [24] U. Grenander and M. Miller, *Pattern Theory: From Representation to Inference*. New York, NY, USA: Oxford University Press, Inc., 2007.
- [25] M. Leyton, “A theory of information structure II. a theory of perceptual organization,” *Journal of Mathematical Psychology*, vol. 30, no. 3, pp. 257–305, Sep. 1986.
- [26] F. Restle, “Theory of serial pattern learning: Structural trees,” *Psychological Review*, vol. 77, no. 6, pp. 481–495, 1970.
- [27] D. H. Hubel and T. N. Wiesel, “Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex,” *The Journal of Physiology*, vol. 160, no. 1, pp. 106–154, 1962.
- [28] F. W. Campbell and J. J. Kulikowski, “Orientational selectivity of the human visual system,” *The Journal of Physiology*, vol. 187, no. 2, pp. 437–445, 1966.
- [29] E. L. Bienenstock, L. N. Cooper, and P. W. Munro, “Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex,” *The Journal of neuroscience: the official journal of the Society for Neuroscience*, vol. 2, no. 1, pp. 32–48, 1982.
- [30] D. Ferster, “Spatially opponent excitation and inhibition in simple cells of the cat visual cortex,” *The Journal of Neuroscience*, vol. 8, no. 4, pp. 1172–1180, 1988.
- [31] T. W. Troyer, A. E. Krukowski, N. J. Priebe, and K. D. Miller, “Contrast-invariant orientation tuning in cat visual cortex: Thalamocortical input tuning and correlation-based intracortical connectivity,” *The Journal of Neuroscience*, vol. 18, no. 15, pp. 5908–5927, 1998.
- [32] O. Yizhar, L. E. Fenno, M. Prigge, F. Schneider, T. J. Davidson, D. J. OShea, V. S. Sohal, I. Goshen, J. Finkelstein, J. T. Paz, K. Stehfest, R. Fudim, C. Ramakrishnan, J. R. Huguenard, P. Hegemann, and K. Deisseroth, “Neocortical excitation/inhibition balance in information processing and social dysfunction,” *Nature*, vol. 477, no. 7363, pp. 171–178, 2011.
- [33] A. B. Watson and J. A. Solomon, “Model of visual contrast gain control and pattern masking,” *Journal of the Optical Society of America A*, vol. 14, no. 9, pp. 2379–2391, 1997.
- [34] H. R. Sheikh, K. Seshadrinathan, A. K. Moorthy, Z. Wang, A. C. Bovik, and L. K. Cormack. (2006) Image and video quality assessment research at live. [Online]. Available: <http://live.ece.utexas.edu/research/quality/>.
- [35] “Method for the subjective assessment of the quality of television pictures,” ITU, Document ITU-R BT.500-11, 2002, geneva, Switzerland.



**Jinjian Wu** received his B.Sc and Ph.D from Xidian University, Xi’an, China, in 2008 and 2013, respectively. From September 2011 to March 2013, he was a research assistant in Nanyang Technological University, Singapore. From August 2013 to August 2014, he was a postdoctoral research fellow in Nanyang Technological University. From July 2013 to June 2015, he was a lecture in Xidian University. Since July 2015, he has been an associated professor with the School of Electronic Engineering, Xidian University. His research interests include visual perceptual modeling, saliency estimation, quality evaluation, and just noticeable difference estimation. He has served as TPC member for ICME2014-2015, PCM2015-2016, ICIP2015, and QoMEX2016. He awarded the best student paper of ISCAS 2013.



**Leida Li** received the B.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 2004 and 2009, respectively. From February to June 2008, he was a visiting Ph.D. student in the Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Taiwan. From January 2014 to January 2015, he was a Visiting Research Fellow in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Currently, he is a Full Professor in the School of Information and Control Engineering, China University of

Mining and Technology, China, and a Senior Research Fellow in the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. His research interests include multimedia quality assessment, information hiding and image forensics.



**Weisheng Dong** received the B.S. degree in electronic engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2004, and the Ph.D. degree in circuits and system from Xidian University, Xian, China, in 2010. From 2009 to 2010, he was a Research Assistant with the Department of Computing, Hong Kong Polytechnic University, Hong Kong. In 2010, he joined the School of Electronic Engineering, Xidian University, as a Lecturer, where he has been a Professor since 2016. His research interests include inverse problems

in image processing, sparse and low-rank representation. He was a recipient of the Best Paper Award at the SPIE Visual Communication and Image Processing (VCIP) in 2010. He currently serves as an associate editor of IEEE Transactions on Image Processing.



**Guangming Shi** (SM'10) received the B.S. degree in automatic control, the M.S. degree in computer control, and the Ph.D. degree in electronic information technology from Xidian University, Xi'an, China, in 1985, 1988, and 2002, respectively. He had studied at the University of Illinois and University of Hong Kong. Since 2003, he has been a Professor with the School of Electronic Engineering, Xidian University. He awarded Cheung Kong scholar Chair Professor by ministry of education in 2012. He is currently the Academic Leader on circuits and

systems, Xidian University. His research interests include compressed sensing, brain cognition theory, multirate filter banks, image denoising, low-bitrate image and video coding, and implementation of algorithms for intelligent signal processing. He has authored or co-authored over 200 papers in journals and conferences. He served as the Chair for the 90th MPEG and 50th JPEG of the international standards organization (ISO), technical program chair for FSKD06, VSPC 2009, IEEE PCM 2009, SPIE VCIP 2010, IEEE ISCAS 2013.



**Weisi Lin** (F'16) received his Ph.D. from Kings College, London University, U.K. He served as the Lab Head of Visual Processing, Institute for Infocomm Research, Singapore. Currently, he is an Associate Professor in the School of Computer Engineering. His areas of expertise include image processing, perceptual signal modeling, video compression, and multimedia communication, in which he has published 170 journal papers, 230+ conference papers, filed 7 patents, and authored 2 books. He is an AE for IEEE Trans. on Image Processing, and IEEE

Trans. Circuits and Systems for Video Tech. He has been a Technical Program Chair for IEEE ICME 2013, PCM 2012, QoMEX 2014 and IEEE VCIP 2017. He has been an invited/panelist/keynote/tutorial speaker in 20+ international conferences, as well as a Distinguished Lecturer of IEEE Circuits and Systems Society 2016-2017, and Asia-Pacific Signal and Information Processing Association (APSIPA), 2012-2013. He is a Fellow of IEEE and IET, and an Honorary Fellow of Singapore Institute of Engineering Technologists.



**C.-C. Jay Kuo** received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1980, and the M.S. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively. He is currently the Director of the Media Communications Laboratory and Dean's Professor of Electrical Engineering at the University of Southern California, Los Angeles. His research interests include digital image/video analysis and modeling, multimedia data compression,

communication and networking and machine learning. He is a coauthor of about 250 journal papers, 900 conference papers, 30 patents, and 14 books.