

Lightness Perception Model for Natural Images

Xianglin Meng, Zhengzhi Wang

College of Mechatronics Engineering and Automation, National University of Defense Technology, Changsha, China.
Email: xlmeng@nudt.edu.cn

Received May 11th, 2010; revised May 28th, 2010; accepted June 12th, 2010.

ABSTRACT

A perceptual lightness anchoring model based on visual cognition is proposed. It can recover absolute lightness of natural images using filling-in mechanism from single-scale boundaries. First, it adapts the response of retinal photoreceptors to varying levels of illumination. Then luminance-correlated contrast information can be obtained through multiplex encoding without additional luminance channel. Dynamic normalization is used to get smooth and continuous boundary contours. Different boundaries are used for ON and OFF channel diffusion layers. Theoretical analysis and simulation results indicate that the model could recover natural images under varying illumination, and solve the trapping, blurring and fogging problems to some extent.

Keywords: Perception, Lightness Anchoring, Filling in, Adaptation, Multiplex Contrast

1. Introduction

Studies show that human could perceive a wide dynamic range of lightness from dim moonlight to dazzling sunlight. Human visual system has stable perceptual capability for scenes under variable illuminations, which is referred to as lightness constancy. First of all, two concepts must be clarified: luminance and lightness. Luminance values denote light intensities within the retinal image while lightness is related to our perceived world. Luminance values in the retinal image are a product, not only of the actual physical shade of gray of the imaged surfaces, but also, and even more so, of the intensity of the light illuminating those surfaces. The luminance of any region of the retinal image can vary by a factor of no more than thirty to one as a function of the physical reflectance of that surface. However, it can vary as a factor of a billion to one as a function of the amount of illumination on that surface. The net result is that any given luminance value can be perceived as literally any shade of gray, depending on its context within the image. Despite the challenge, human perceive shades of surface grays with rough accuracy [1]. Lightness represents the cortical perceptual result of retinal stimuli and the process of lightness perception can be understood as a mapping from natural image input to visual percept output.

BCS/FCS (Boundary Contour System/Feature Contour System) proposed by Grossberg *et al.* is representative of visual lightness perception model. Its extended versions have explained a mass of psychological experimental data [2,3]. BCS/FCS model discounts the illuminant to

obtain contrast information through retinal preprocessing. Further processing is made by visual cortex to get boundary contour and surface. However, such illumination discounting information can just estimate relative measurements of reflectance of the surface. Visual cortex needs to compute the absolute lightness values that exploit the full dynamic range to perceive effectively. That's just the so-called anchoring problem.

Grossberg believed that boundaries and surfaces are visual perceptual units, and proposed aFILM model in 2006 [4]. The model augments a lightness anchoring stage in the framework of BCS/FCS. Sepp *et al.* proposed a multi-scale filling-in model to reconstruct the image surface lightness [5]. It extends the confidence-based filling-in model to multi-scale processing, thus speeding up the filling-in process. A key aspect of visual cognition research is to process natural images effectively while explaining psychological data, which facilitates higher cognition process such as object recognition and provides inspirations to machine vision.

This paper presents a neural dynamic model to simulate the mapping process from luminance to lightness according to recent neurophysiological and psychological experimental findings. The proposed model could recover natural images under varying levels of illumination, and solve the trapping, blurring and fogging problems to some extent.

2. Model Description

After retinal photoreceptors receive the input image, re-

tinal adaptation occurs firstly. The light-adapted signal goes to multiplex coding and boundary detection. The contrast information from multiplex coding is used to recover the absolute surface lightness, and the boundary contour is to block activity diffusion between surfaces. Final perceptual lightness output is obtained through surface filling-in mechanism. The overall model architecture is depicted in **Figure 1**.

2.1 Retinal Adaptation

The model retina calculates the steady-state of retinal adaptation to a given input image [4]. It adapts the response of photoreceptors to varying levels of incoming light, since otherwise the visual process could be desensitized by saturation right at the photoreceptor. Light adaptation, at the photoreceptor outer segment, protects each photoreceptor from saturation by using intracellular temporal adaptation that shifts the photoreceptor sensitivity curve [6]. As illustrated in **Figure 2**, the light-adapted signal is further processed at the photoreceptor inner segment where it gets feedback from a horizontal cell (HC) that is connected with other HCs by gap junctions [7], forming a syncytium that is sensitive to spatial contrast. HC inhibition further adjusts the sensitivity curve to realize spatial contrast adaptation. It is assumed that the permeability of gap junctions between HCs decreases as the difference of the inputs to HCs from coupled photoreceptors increases. The model retina hereby segregates and selectively suppresses signals in regions that have strong contrasts, such as a light source. HCs connections are not constrained to nearest neighbors but reach farther regions. Inhibition of the HC on the photoreceptor controls the output of the photoreceptor by modulating Ca^{2+} influx at the photoreceptor inner segment. This feedback prevents the output from being saturated by localized high-contrast input

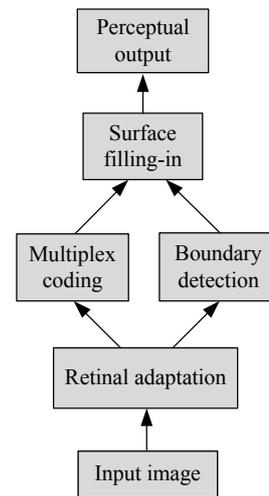


Figure 1. Model block diagram

signals.

The outer segment of retinal photoreceptor can be modeled by the equation:

$$s_{ij}(t) = I_{ij} \cdot g_{ij}(t) \tag{1}$$

where (i, j) denotes spatial position, $s_{ij}(t)$ represents the output of outer segment, I_{ij} is the input image, $g_{ij}(t)$ is an automatic gain control term simulating negative feedback mediated by Ca^{2+} ions and follows:

$$\frac{dg_{ij}}{dt} = (A_g - g_{ij}) - g_{ij}(B_l I_{ij} + B_{\bar{l}} \bar{l}) \tag{2}$$

where A_g, B_l and $B_{\bar{l}}$ are constants. \bar{l} denotes spatio-laverage of input image that approximates the effect of recent image scanning by sequences of eye movements. The second term describes the inactivation of $g_{ij}(t)$ by

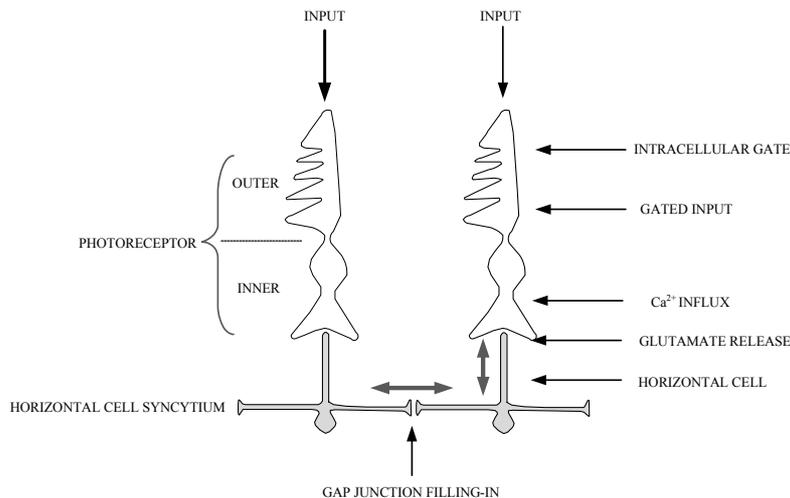


Figure 2. Circuit of retinal adaptation

the present I_{ij} and \bar{I} . The inner segment of the photoreceptor receives the signal s_{ij} from the outer segment and also gets feedback H_{ij} from the horizontal cell. HC modulation of the output of the inner segment is modeled by the equation:

$$S_{ij} = \frac{s_{ij}}{C_H \exp(H_{ij}) \cdot (C_s - s_{ij}) + 1} \quad (3)$$

where $C_s = A_g / B_I$, C_H is a constant. The relationship between H_{ij} and HC activity h_{ij} follows:

$$H_{ij} = \frac{a_H h_{ij}^2}{b_H^2 + h_{ij}^2} \quad (4)$$

where a_H and b_H are constants. The activity of an HC connected to its neighbors through gap junctions is defined as diffusion equation:

$$\frac{dh_{ij}}{dt} = -h_{ij} + \sum_{(p,q) \in N_{ij}} P_{pqij} (h_{pq} - h_{ij}) + S_{ij} \quad (5)$$

where P_{pqij} is the permeability between cells at (i, j) and (p, q) ,

$$P_{pqij} = 1 - \frac{1}{1 + \exp[-(|S_{pq} - S_{ij}| - \beta) / \lambda]} \quad (6)$$

β, λ are constants. N_{ij} is the neighborhood of size r to which the HC at (i, j) is connected:

$$N_{ij} = \{(p, q) : \sqrt{(i-p)^2 + (j-q)^2} \leq r, (p, q) \neq (i, j)\} \quad (7)$$

2.2 Multiplex Contrast Code

Lightness anchoring model generally incorporates an extra luminance-driven channel to recover absolute lightness in addition to retinal contrast channels. Maybe multi-scale band-pass filters are used to get contrast and luminance information, such as aFILM model which acquires contrast information through small scale filtering and obtains luminance information through large scale filtering [4]. There is evidence showing that a larger disinhibitory surround exists outside of the classical receptive field of retinal ganglion cells. Accordingly, a multiplex retinal code is proposed to solve the anchoring problem. The code is composed of retinal contrast responses, where contrast responses are locally modulated by brightness (ON cell) or darkness (OFF cell). The modulation is implemented by an extensive disinhibitory surround or outer surround (OS), an annulus which is situated beyond the classical center-surround receptive of retinal ganglion cells as shown in **Figure 3** [8]. The classical receptive field is sensitive to contrast information and outer surround is sensitive to local luminance. Then

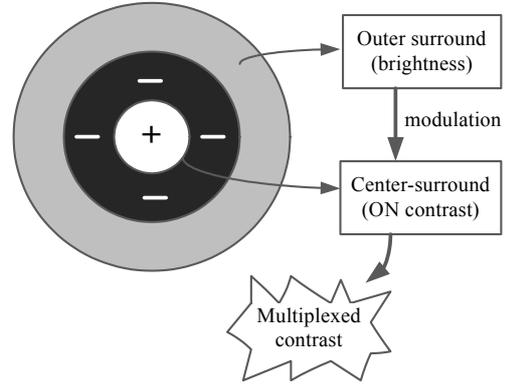


Figure 3. Multiplex retinal code

luminance-correlated contrast information can be obtained through multiplex encoding without additional luminance channels. Also, it is plausible from a neuro-physiological point of view.

Due to the asymmetry phenomenon of ON cell and OFF cell responses for the classical center-surround receptive field of retinal ganglion cells, a self-inhibition mechanism is adopted:

$$\frac{dx_{ij}^+(t)}{dt} = -\alpha x_{ij}^+ + I_{ij}^c - I_{ij}^s - [I_{ij}^c - I_{ij}^s]^+ \cdot x_{ij}^+ \quad (8)$$

$$\frac{dx_{ij}^-(t)}{dt} = -\alpha x_{ij}^- + I_{ij}^s - I_{ij}^c - [I_{ij}^s - I_{ij}^c]^+ \cdot x_{ij}^- \quad (9)$$

where x_{ij}^+ and x_{ij}^- are ON cell and OFF cell responses respectively, $[\bullet]^+ \equiv \max(\bullet, 0)$, and α is a decay factor. $I_{ij}^c = S_{ij}$ is center input. $I_{ij}^s = (S \otimes G_s)_{ij}$ is surround input. G_s is a Gaussian kernel. “ \otimes ” denotes convolution operator. The last term of the right side of above equations denotes self-inhibition which is used to solve the response asymmetry problem [9].

Let m^+ and m^- denote local luminance-correlated multiplexed contrast response of ON and OFF channels respectively. The activity of the outer surround is computed by convolution with a Gaussian kernel: $O = \text{Norm}[S] \otimes G_{\sigma_o}$. $\text{Norm}[\bullet]$ implements the normalization operator which maps the input into $[0, 1]$. Outer surround activity acts to multiplicatively gate the classical center-surround responses of retinal ganglion cells. So the multiplexed contrast responses are defined as

$$m_{ij}^+ = \frac{O_{ij}}{\beta_o + O_{ij}} \cdot [x_{ij}^+]^+, \quad m_{ij}^- = \frac{1 - O_{ij}}{\beta_o + 1 - O_{ij}} \cdot [x_{ij}^-]^+ \quad (10)$$

where β_o is a saturation constant.

2.3 Boundary Detection

The same retinal contrast information is used in both boundary detection and filling-in mechanism in BCS/

FCS model. In this paper, we use different contrast information. Boundary detection is completed through a dynamic normalization network [10]. First, we define an operator $\mathfrak{R}_\lambda[\bullet]$:

$$\mathfrak{R}_\lambda[x] = \frac{1 + e^{-|\lambda|}}{1 + e^{-\lambda \cdot x}} x \quad (11)$$

We can get

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \mathfrak{R}_\lambda[x] &= \lim_{\lambda \rightarrow \infty} \frac{1 + e^{-|\lambda|}}{1 + e^{-\lambda \cdot x}} x \\ &= \begin{cases} x, & x > 0 \\ 0, & x = 0 \\ 0, & x < 0 \end{cases} = \max(x, 0) \end{aligned} \quad (12)$$

similarly,

$$\lim_{\lambda \rightarrow -\infty} \mathfrak{R}_\lambda[x] = \min(x, 0) \quad (13)$$

$$\mathfrak{R}_\lambda[x]|_{\lambda=0} = x \quad (14)$$

for simplicity we denote:

$$\mathfrak{R}_\infty = \lim_{\lambda \rightarrow \infty} \mathfrak{R}_\lambda, \quad \mathfrak{R}_{-\infty} = \lim_{\lambda \rightarrow -\infty} \mathfrak{R}_\lambda, \quad \mathfrak{R}_0 = \mathfrak{R}_\lambda|_{\lambda=0}.$$

Subsequently, we define nonlinear diffusion operators:

$$\Phi^+ f_{ij} = \sum_{(p,q) \in N_{ij}^4} \mathfrak{R}_\infty[f_{pq} - f_{ij}] \quad (15)$$

$$\Phi^- f_{ij} = \sum_{(p,q) \in N_{ij}^4} \mathfrak{R}_{-\infty}[f_{pq} - f_{ij}] \quad (16)$$

where $N_{ij}^4 = \{(i+1, j), (i-1, j), (i, j-1), (i, j+1)\}$ is a four nearest neighborhood. The dynamic normalization equations are:

$$\frac{da_{ij}(t)}{dt} = \kappa \cdot \Phi^- a_{ij}(t) + S_{ij} \delta(t - t_0) \quad (17)$$

$$\frac{db_{ij}(t)}{dt} = \kappa \cdot \Phi^+ b_{ij}(t) + S_{ij} \delta(t - t_0) \quad (18)$$

$$\frac{dc_{ij}(t)}{dx} = b_{ij}(0 - c_{ij}) - a_{ij}(1 - c_{ij}) + S_{ij} \quad (19)$$

$$\frac{dd_{ij}(t)}{dx} = b_{ij}(1 - d_{ij}) - a_{ij}(0 - d_{ij}) - S_{ij} \quad (20)$$

κ is a diffusion coefficient, $a_{ij}, b_{ij}, c_{ij}, d_{ij}$ are min-syncytium, max-syncytium, normalized ON type and OFF type cell activity respectively. δ denotes Dirac's delta function. Seen from (17), a cell a_{ij} of the min-syncytium may only decrease its activity along with time, as long as there exists any activity gradient between this cell and its four nearest neighbors. The result of diffusion is that a_{ij} decreases to the global minimum value. In an analogous fashion, a cell b_{ij} of the max-syncytium finally gets the global maximum value of activity. We can get smooth and continuous boundary contour by ear-

ly dynamics of the dynamic normalization network. We denote the resultant responses as y^+, y^- . ON contour and OFF contour representing diffusion barriers are defined as

$$w_{ij}^+ = \frac{\text{thresh}_{\Theta_w}(y_{ij}^+)}{\beta_w + \text{thresh}_{\Theta_w}(y_{ij}^+)} \quad (21)$$

$$w_{ij}^- = \frac{\text{thresh}_{\Theta_w}(y_{ij}^-)}{\beta_w + \text{thresh}_{\Theta_w}(y_{ij}^-)} \quad (22)$$

where β_w is a saturation constant,

$$\text{thresh}_{\Theta_w}(x) = [\text{Norm}[x] - \Theta_w]^+, \quad \Theta_w \in [0, 1].$$

The detected boundaries are always discontinuous due to noise or other factors, which allows for activity exchange between adjacent surface representations. Consequently, perceived luminance contrasts between surfaces decrease, because they eventually adopt the same value of perceptual activity. It's the so-called fogging problem. In order to counteract fogging, an interaction zone around contours is defined. Within this zone, brightness activity and darkness activity undergo mutual inhibition, leading to a slow-down of diffusion rate at boundary leaks. Thus fogging is decelerated and surface edges will appear blurry at boundary gaps. Let

$$z_{ij} = \text{thresh}_{\Theta_z}(w_{ij}^+ + w_{ij}^-) \quad (23)$$

with a threshold value Θ_z . Interaction zone activity Z is defined as

$$Z = \frac{z}{\beta_z + z} \otimes G_{\sigma_z} \quad (24)$$

where β_z is a saturation constant, G_{σ_z} is a Gaussian kernel with standard deviation σ_z .

2.4 Surface Filling-in

The multiplex contrast responses diffuse within regions formed by boundary contours to form perceptual surface representations. Diffusion layers are computed by dynamic equations:

$$\begin{aligned} \frac{df_{ij}^+(t)}{dt} &= \gamma w_{ij}^-(E_{in} - f_{ij}^+) + \sum_{(p,q) \in N_{ij}^4} P_{ijpq}^+ \mathfrak{R}_\infty[f_{pq}^+ - f_{ij}^+] \\ &\quad + \delta(t - t_0) m_{ij}^+ \end{aligned} \quad (25)$$

$$\begin{aligned} \frac{df_{ij}^-(t)}{dt} &= \gamma w_{ij}^+(E_{in} - f_{ij}^-) + \sum_{(p,q) \in N_{ij}^4} P_{ijpq}^- \mathfrak{R}_{-\infty}[f_{pq}^- - f_{ij}^-] \\ &\quad + \delta(t - t_0) m_{ij}^- \end{aligned} \quad (26)$$

where $f_{ij}^{+/-}$ represent brightness and darkness activity. γ is a constant, E_{in} is an inhibitory reversal potential. Diffusion coefficients:

$$P_{ijpq}^+ = \frac{1}{1 + \varepsilon(Z_{ij}[f_{ij}^-]^+ + Z_{pq}[f_{ij}^-]^+)} \quad (27)$$

$$P_{ijpq}^- = \frac{1}{1 + \varepsilon(Z_{ij}[f_{ij}^+]^- + Z_{pq}[f_{ij}^+]^-)} \quad (28)$$

where ε is a constant. From equations above we can see that OFF contours are used to block brightness activity diffusion while ON contours are employed to block darkness activity diffusion. In this way, we can alleviate the activity trapping problem. The perceptual activity of surface representations

$$p_{ij} = \frac{g_{leak}V_{rest} + [f_{ij}^+]^+ - [f_{ij}^-]^+}{g_{leak} + [f_{ij}^+]^+ + [f_{ij}^-]^+} \quad (29)$$

where g_{leak} is leakage conductance, and V_{rest} is the resting potential.

3. Simulation Results

In order to validate the effectiveness of the proposed model, we first evaluate the performance of each stage, then the lightness perception performance of the whole model is tested using natural images. The size of test images in simulation is 256×256 .

3.1 Retinal Adaptation

Firstly, we evaluate the performance of retinal adaptation. The simulation results are illustrated in **Figure 4**. The original input image, HC activity and retinal adaptation output are shown in **Figures 4(a)-(c)** respectively. The input image has intensive contrast, thus we can hardly see the dark scene. Through retinal adaptation processing, we cannot only see the clouds in sky of bright region, but also see the sand of dark region, even small rocks on the ground. It's the result of light adaptation and spatial contrast adaptation of the retina.

3.2 Boundary Contour Detection

In this section, we compare the boundary detection performance of dynamic normalization with that of classical laplacian linear filtering method. **Figures 5(a) and (b)** are ON and OFF boundaries of Lena image obtained by laplacian method. **Figures 5(c) and (d)** are detection results from dynamic normalization. As we can see from **Figure 5**, the boundary contours obtained by dynamic normalization are smoother and more continuous, which will block the diffusion between different regions more effectively in later filling-in processing, thus alleviating

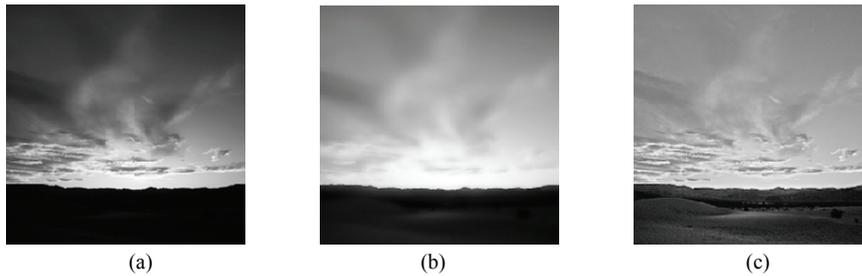


Figure 4. Retinal adaptation simulation. (a) stimulus; (b) HC activity; (c) adaptation output

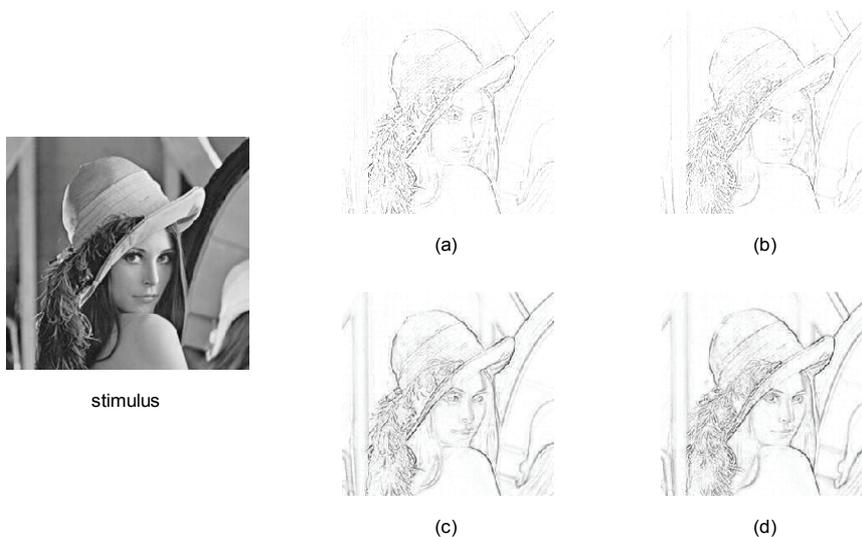


Figure 5. Boundary contour detection

the fogging problem.

3.3 Multiplex Contrasts

For the synthetic stimulus in **Figure 6(a)**, one-dimensional luminance staircase is shown in **(b)-(d)** by the black solid lines. The red solid lines denote ON responses while the blue dashed lines correspond to OFF responses. **Figure 6(b)** shows the responses of classical center-surround receptive field. OFF responses are always higher than ON responses around edges, and both decrease with luminance increase. The asymmetry problem of ON and OFF responses is solved by self-inhibition mechanism illustrated in **Figure 6(c)**, where ON and OFF responses are only sensitive to contrast and insensitive to luminance. Therefore, we could modulate ON responses with local brightness and OFF with local darkness, thus getting luminance-correlated multiplex

contrast responses. As seen in **Figure 6(d)**, ON responses increase while OFF responses decrease as the luminance increases.

3.4 Surface Filling-in

Nonlinear diffusion is used to implement surface filling-in and recover absolute perceived lightness. Most filling-in models have blurring, trapping and fogging problems such as confidence-based filling-in. As described previously, in this paper, different boundaries are used for the brightness and the darkness diffusion layers. As a consequence, trapping can be weakened to some extent. Besides, the adoption of interaction zone could alleviate fogging. **Figure 7(a)** shows the result of confidence-based filling-in, demonstrating serious fogging problem. **Figure 7(b)** is the filling-in result of the proposed model. It can be seen that absolute lightness is recovered while

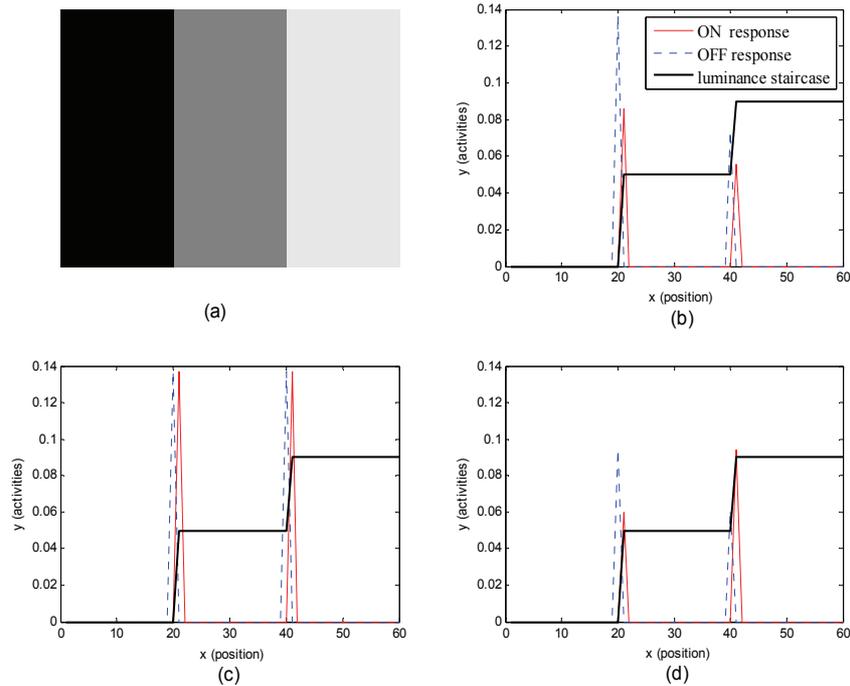


Figure 6. Multiplex contrast responses



Figure 7. Surface filling-in results. (a) confidence-based filling-in; (b) nonlinear diffusion

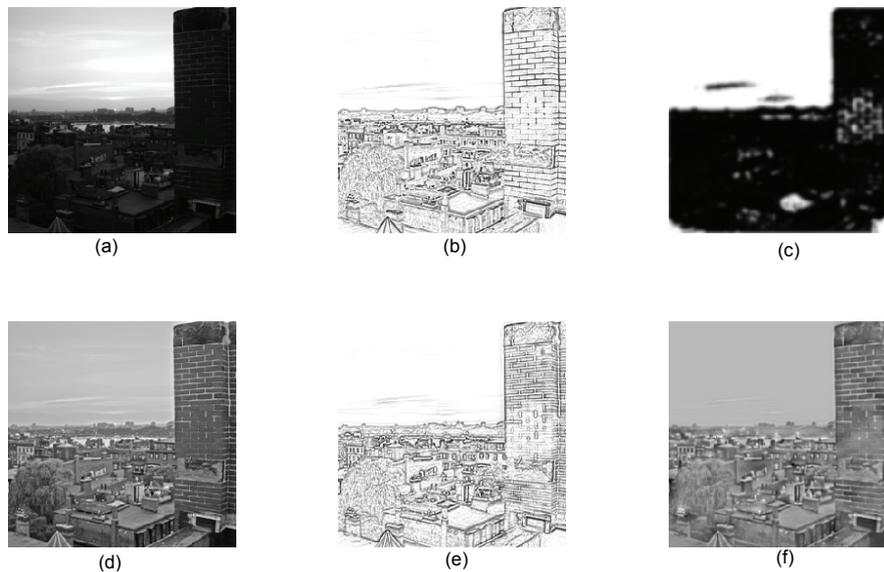


Figure 8. Lightness perception of a real-world image. (a) stimulus; (b) ON contour; (c) interaction zone; (d) retinal adaptation; (e) OFF contour; (f) perception result

reserving contrast, without fogging problem.

Finally, we test the performance of the whole model using a real-world image. Simulation results of different processing stages are shown in **Figure 8**. We can find that the proposed model could recover absolute lightness of natural images under varying illuminating conditions. Comparing with aFILM model, it can get smoother and more continuous contours than BCS, and explicitly weakens blurring, trapping and fogging problems to some degree. Moreover, the proposed model has the advantage of illuminating adaptation in comparison with Sepp's brightness perception model. It provides an alternative approach for lightness perception.

4. Conclusions

In this paper, a perceptual lightness anchoring model based on visual cognition is proposed. It can recover absolute lightness of real-world images using filling-in mechanism from single-scale boundaries. Further, it can weaken trapping, blurring and fogging to some extent. Reasonable perceptual results could be obtained for natural images under varying illumination conditions. The proposed model could be applied to image enhancement and image reconstruction in machine vision, and facilitate robust processing of higher levels such as object recognition. However, there are not normative objective measurements to evaluate the performance of visual cognition model. Most measurements are qualitative, and it makes no exception for this paper. Additionally, this model recovers absolute lightness from single scale channel. So its ability of noise suppression is weaker

than multi-scale processing. We will study further in these aspects.

5. Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 60835005) and Scientific Research Plan Projects of NUDT (No. JC09-03-04). The authors are very grateful to the reviewers for their valuable remarks and constructive comments on this paper.

REFERENCES

- [1] A. Gilchrist, C. Kossyfidis, F. Bonato and T. Agostini, "An Anchoring Theory of Lightness Perception," *Psychological Review*, Vol. 106, No. 4, 1999, pp. 795-834.
- [2] M. Anton-Rodriguez, F. J. Diaz-Pernaz, J. F. Diez-Higuera, M. Martinez-Zarzuola, D. Gonzalez-Ortega and D. Boto-Giralda, "Recognition of Coloured and Textured Images through a Multi-Scale Neural Architecture with Orientational Filtering and Chromatic Diffusion," *Neurocomputing*, Vol. 72, No. 16-18, 2009, pp. 3713-3725.
- [3] A. Fazl, S. Grossberg and E. Mingolla, "View-Invariant Object Category Learning: How Spatial and Object Attention are Coordinated Using Surface-Based Attentional Shrouds," *Cognitive Psychology*, Vol. 58, No. 1, 2009, pp. 1-48.
- [4] S. Grossberg and S. Hong, "A Neural Model of Surface Perception: Lightness, Anchoring, and Filling-in," *Spatial Vision*, Vol. 19, No. 2-4, 2006, pp. 263-321.
- [5] W. Sepp and H. Neumann, "A Multi-Resolution Filling-in Model for Brightness Perception," *9th International Conference on Artificial Neural Networks*, 1999, pp. 461-466.

- [6] M. Vanleeuwen, I. Fahrenfort, T. Sjoerdsma, R. Numan and M. Kamermans, "Lateral Gain Control in the Outer Retina Leads to Potentiation of Center Responses of Retinal Neurons," *The Journal of Neuroscience*, Vol. 29, No. 19, 2009, pp. 6358-6366.
- [7] H. Jouhou, K. Yamamoto, M. Iwasaki and M. Yamada, "Acidification Decouples Gap Junctions but Enlarges the Receptive Field Size of Horizontal Cells in Carp Retina," *Neuroscience Research*, Vol. 57, No. 2, 2007, pp. 203-209.
- [8] M. S. Keil, G. Cristobal, T. Hansen and H. Neumann, "Recovering Real-World Images from Single-Scale Boundaries with a Novel Filling-In Architecture," *Neural Networks*, Vol. 18, No. 10, 2005, pp. 1319-1331.
- [9] R. D. S. Raizada and S. Grossberg, "Context-Sensitive Binding by the Laminar Circuits of V1 and V2: A Unified Model of Perceptual Grouping, Attention, and Orientation Contrast," Boston University, 2000.
- [10] M. S. Keil, "Local to Global Normalization Dynamic by Nonlinear Local Interaction," *Physica D*, Vol. 237, No. 6, 2008, pp. 732-744.