Measuring the attentional effect of the bottom-up saliency map of natural images

Cheng Chen[1,3], **Xilin Zhang**[2,3], **Yizhou Wang**[1,3], **and Fang Fang**[2,3,4,5]

[1]National Engineering Lab for Video Technology and School of Electronics Engineering and Computer Science, [2]Department of Psychology
[3]Key Laboratory of Machine Perception (Ministry of Education)
[4]Peking-Tsinghua Center for Life Sciences
[5]IDG/McGovern Institute for Brain Research
Peking University, Beijing 100 1, P.R. China

textures because of their rich naturalistic low-level features that the human visual system is tuned to. Although these natural images were invisible, the difference of visual saliency (calculated by a famous computational saliency model [1]) between inside and outside a local region could attract attention to improve the

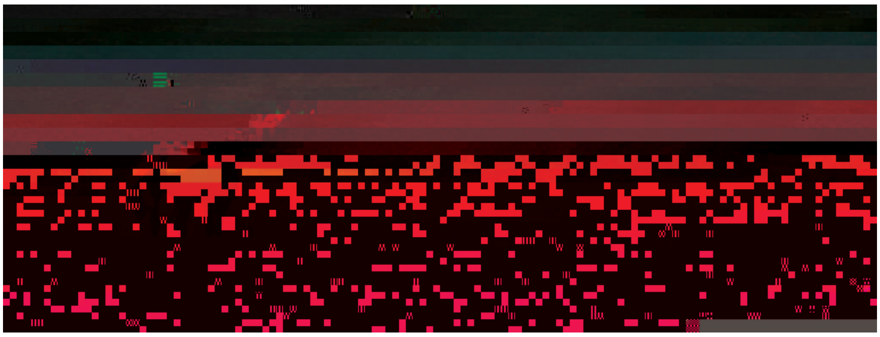
Fig. 1. An example of a color image (left) and its saliency map (right). White region in the right image indicate its salient region.

pixel in the saliency map ranges from 0 to 1, higher value correlated with more saliency. Itti *et al.* proposed a biologically-plausible saliency model based on a center-surround mechanism, by combining information from three channels: color, intensity and orientation [1]. According to the spectrum of natural images, Hou one subject who was one of the authors. They were given written, informed consent in accordance with the procedures and protocols approved by the human subjects review committee of Peking University.

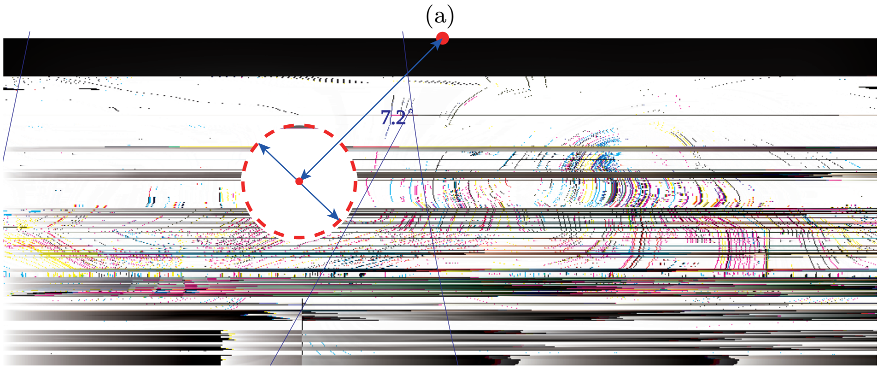
t

We collected a large number of grayscale images about natural scenes from the Internet, resized them into the same size (384$\times$1024 pixels), and decreased the luminance of these images to a low level (about 2.9 cd/m^2), Fig. 2 (a) shows a sample image. To quantitatively measure the attentional effect, we adopt a visual saliency model proposed by Itti *et al.* [1] and calculate the saliency map of each image. After that we selected 50 images, and each of them had a round salient region centered at about 7.2° eccentricity in the lower left quadrant(called left-salient images). The diameter of the salient region was about 4°. By flipping each image across its vertical midline, we can generate 50 new images, each of them had a local salient region in the lower right quadrant(called right-salient images). Notice that the content between the two groups of images were totally the same, the only difference between the two groups was the location of the salient region. The average saliency map of the 50 left-salient images can be seen in Fig. 2 (b).

Based on the bottom-up saliency, we classified all images into two groups: high salient images and low salient images. We proposed a salient index



(a)



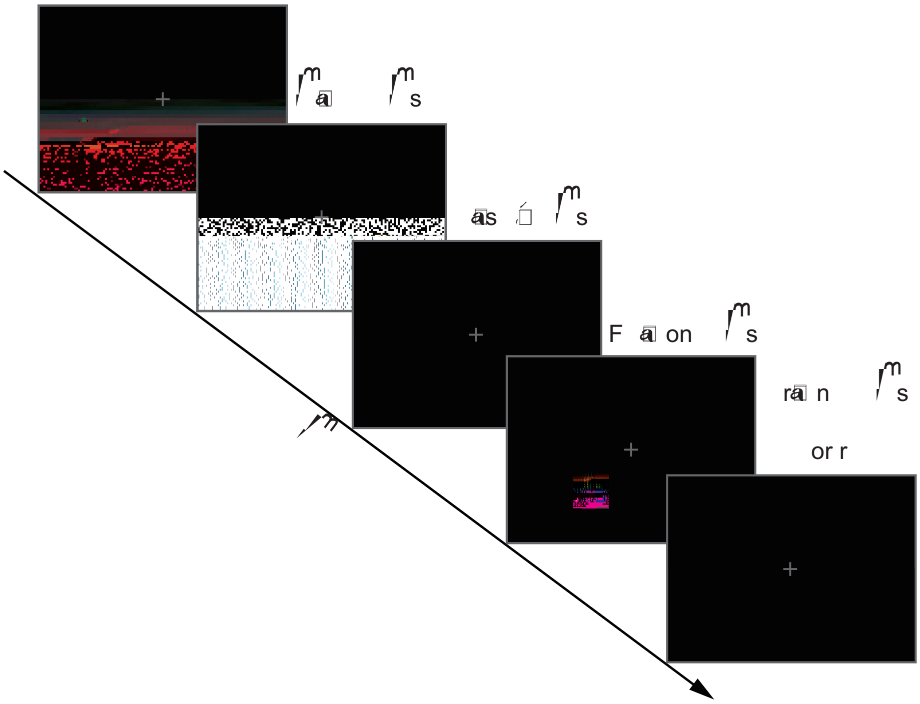


Fig. 3. The procedure of our experiment.

were asked to press one of the two keys to indicate the orientation of the grating. The duration of each trial was 2s, Fig. 3 shows the procedure of our experiment.

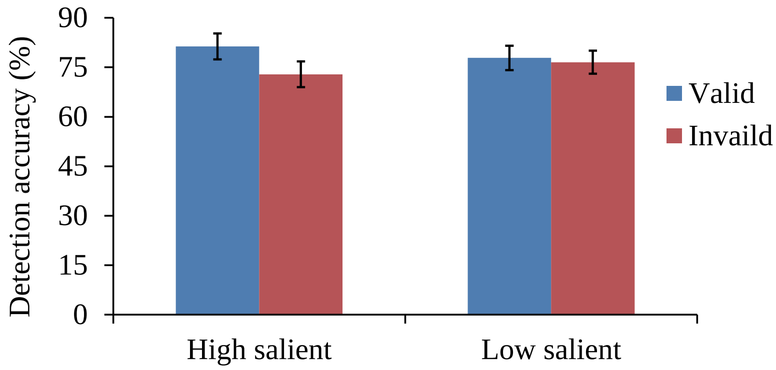
The experiment consisted of 10 blocks. Each block contained 100 trials with

3 Experimental Result

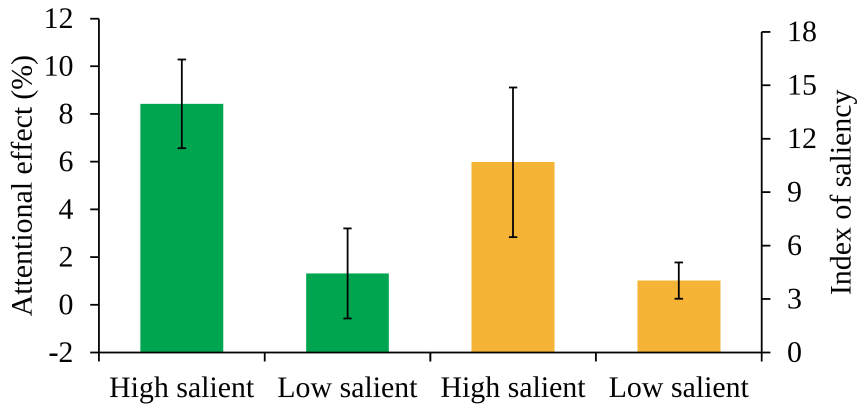
.1 s n s t

The purpose of the 2AFC experiment was to evaluate whether those natural images used as the cue in the attentional experiment were indeed invisible. High salient images and low salient images were counterbalanced in this task. Subjects had to report whether they can see an image before the mask (details can be found in Section 2.3.

We found that percentages of correct detection (mean $\pm$ std) were 48.6 $\pm$ 6.0% and 50.9 $\pm$ 5.7% for high salient and low salient images respectively. Paired t-test results showed that the percentages of correct detection were statistically



(a)



(b)

Fig. 4. Results of our experiment. (a) The performance of the grating orientation discrimination task for high salient images and low salient images. (b) The left two green bars indicate the attentional e ect of bottom-up saliency maps in high salient and low salient groups. The right two yellow bars indicate the predication of the attentional e ect in two groups.

. Posner, M.I., Snyder, C.R.R., Davidson, B.J. Attention and the detection of signals. J. Exp. Psychol. 10 , 1 0C-1 (1 0)
. Eckstein, M.P., Schimozaki, S.S., Abbey, C.K. The footprints of visual attention in the Posner cueing paradigm revealed by classification images. J. Vis. 2, 2 – (2002)
10. Hou, X., Zhang, L. Saliency detection, A spectral residual approach. IEEE Computer Vision and Pattern Recognition. (200)
11. Wang W., Wang, Y., Huang, Q., Gao, W. Measuing visual saliency by site entropy rate. IEEE Computer Vision and Pattern Recognition. (2010)
12. Itti, L., Koch, C., Niebur, E., Computational Modelling of Visual Attention. Nat. Rev. Neurosci. 2, 1 –20 (2001)
1 . Li, Z., Contextual influences in V1 as a basis for pop out and asymmetry in visual search. Proc. Natl. Acad. Sci. USA. , 10 0–10 (1)
1 . Li, Z., A saliency map in primary visual cortex. Morgan Kaufmann. Trends Cogn. Sci. , –1 (2002)
1 . Allman, J., Miezin, F., McGuinness, E., Stimulus specific responses from beyond the classical receptive field, neurophysiological mechanisms for localCglobal comparisons in visual neurons. Annu. Rev. Neurosci. , 0 – 0 (1)
1 . Jiang, Y., Costello, P., Fang, F., Huang, M., He, S. A gender-and sexual orientation-dependent spatial attentional e ect of invisible images. Proc. Natl. Acad. Sci. USA. 10 , 1 0 C-1 0 2 (200)