A Psychophysical Experiment to Test the Efficient Stereo Coding Theory

Danmei Chen and Zhaoping Li

Department of Computer Science,
The Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong

Published in Theoretical aspects of neural computation K.M. Wong, I. King, and D.Y. Yeung (eds), page 225-235 Springer-verlag January 1998

Abstract. A theory of efficient stereo coding [2] predicts that, in a natural visual environment, where the ocular correlation of the input depends on stimulus orientations, the striate cortical cells are more likely binocular when selective to horizontal rather than vertical orientations. A psychophysical experiment was designed to test this prediction. The interocular transfers of simultaneous orientation contrast at near horizontal and vertical orientations were measured. This measure was used to access the binocularity of the underlying cells. It turned out that in the natural stereovisual environment, the transfers for horizontal orientations were larger than that for vertical case. And in the unnatural stereo experimental environment, where the binocular correlation is the same for vertical and horizontal orientations, the transfer for the horizontal orientation was almost equal to that in the corresponding vertical case. These results are consistent with the theoretical prediction.

1 Introduction

Recoding sensory inputs to remove the input redundancy has been advocated as a sensory preprocessing goal and is argued to have cognitive advantages[1]. An overwhelming source of redundancy in visual input is the pair-wise image pixel correlation, and the binocular correlation is one of them. One of the authors developed a theory[2] proposing that the striate cortex is concerned with, among other things, decorrelating the binocular inputs[2]. It predicts a distribution of monocular/binocular and disparity selective cells and their relationship with the receptive field sizes and orientations. Some of these predictions agree well with experimental observations and demonstrate that many neural coding properties can be understood from visual input statistics.

In addition, the theory is predicting neural properties not yet explored systematically in experiments, and motivating experimental tests. One of such predictions is that in a natural visual environment, the striate cortical cells are more likely binocular when selective to horizontal rather than vertical orientation. This prediction can be understood as follows. Since our eyes are aligned vertically, horizontal disparties are usually larger than vertical disparities , as one can see in Fig. 1. So visual inputs of horizontal orientation, which access the variations and disparities in the vertical direction, are more ocularly correlated compared to inputs of vertical orientation. In the stereo coding theory, the striate cortex is proposed to reduce binocular redundancy, which means that cells should be more binocular if they are selective to features that are more binocularly correlated. So cells selective to horizontal orientation will have higher binocularity than cells selective to vertical orientation.

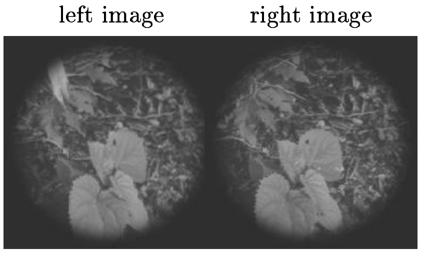


Fig. 1. A pair of the natural stereo images. It is very clear that the vertical disparities are mostly smaller than horizontal disparities.

In this paper, a psychophysical experiment to test the above prediction is reported. We used tilt illusion to find the correlation between binocularity and orientation. The tilt illusion refers to the misperceived orientation of a test line or grating when it is superimposed on, or intersected by, inducing lines or gratings of a different orientation. The magnitude of this illusion, which is called the induced orientation shift, is the difference between the perceived orientations of the test line with and without the inducing lines. This illusion can be transferred ocularly when the test and inducing lines are presented in different eyes, and this interocular transfer can be used to access the underlying binocularity of the neural basis. The interocular transfer is defined as the ratio of the dichoptically induced orientation shift when the test and inducing lines are presented in different eyes and monocular shift when the test and inducing lines are presented in the same eye. Since the transfer is mediated by binocularly driven cortical cells [3], it is a useful measure of cell binocularity. Higher binocularity of the underlying cortical cells will make the transfer closer to 100%. The transfers for both vertical and horizontal test lines were measured in the experiment. We intended to find whether there was a correlation between the degrees of interocular transfer and the orientation of the test lines. Here, simultaneous orientation contrast was used, that is, the test and inducing lines were presented simultaneously[4].

Since the prediction holds only under natural visual environment, where there is a difference between the ocular correlations for the vertical and horizontal orientations. we adopted two approaches to test it. One was the experiment with natural scenes, the other was the experiment without natural scenes. According to the prediction, in the experiment with natural scenes, where we inserted stereo natural scenes during the experimental session to keep a natural visual environment for the subject, the transfer will be closer to 100% for near horizontal test lines than that for vertical test lines. In contrast, in the experiment without natural scenes, where the natural stereo scenes were removed and the subjects were intentionally adapted to an unnatural visual environment without the horizontal-vertical difference, the transfer should be similar for vertical and horizontal test lines. The experiment without natural scenes was used as a control to test the prediction.

2 Method

A schematic plan of the appratus is shown in Fig.2. The subject fixed his/her head in front of the middle of stereoscope. The left eye only viewed stimulus L and the right eye only viewed stimulus R. The positions of stimuli on the monitor screen were carefully adjusted so that the subject perceptibly viewed one image, i.e., the fusion of L and R.

For both the experiments with and without natural scenes, the tests were divided into 8 sections, 4 each for the vertical and horizontal cases, corresponding to 4 different stimulus patterns given by whether the left or right eye viewed test line and whether the test and inducing lines were presented monocularly or dichoptically. For the vertical case, two rows of inducing lines oriented 20 degrees away from vertical were displayed above and below the test line. The test line was about 47 min arc long, and the separation between the top of the test line and the bottom of the inducing lines was 24 min arc. The four sections were denoted by VLM, VLD, VRM, and VRD, respectively, with V denoting vertical orientation, L or R denoting the left or right eye for test line stimulus, and M or D denoting the test and inducing lines presented monocularly or dichoptically. Some examples are shown in the top row of Fig. 3. For the horizontal case, experimental sections, HLM, HLD, HRM and HRD, were designed analogously by rotating the stimulus patterns of the vertical case by 90^o. Some examples are shown in the bottom row of Fig. 3. The eight stimulus patterns for the experiments with and without natural scenes were the same.

Data taking is carried out in a normally illuminated room. The stimulus was displayed on the screen of a SGI graphics terminal driven by SGI/Indigo 2/EX. Test stimulus was exposed for 190 milliseconds(msec) in each trial. Additionally, we did the following to minimize vergence movement. First, two black fixation rings, as well as 4 black anchoring points right inside each ring, were present on the screen during the whole data taking session. Second, during 472 msec before the presentation of the test stimulus, a small black circle concentric to each fixation ring, was exposed in the first half of this period. This circle had a radius of 1/3 of the fixation ring and was much thinner in its circumference, as shown in Fig. 4. Third, before each data taking session, we let the subject view the stimulus pattern as show in Fig. 5(A). This stimulus pattern was presented for the same time duration and under the same conditions as the test stimulus pattern during data taking. We adjusted the positions of L and R on the screen until the subject saw the thin line in the left eye aligned with the central thin

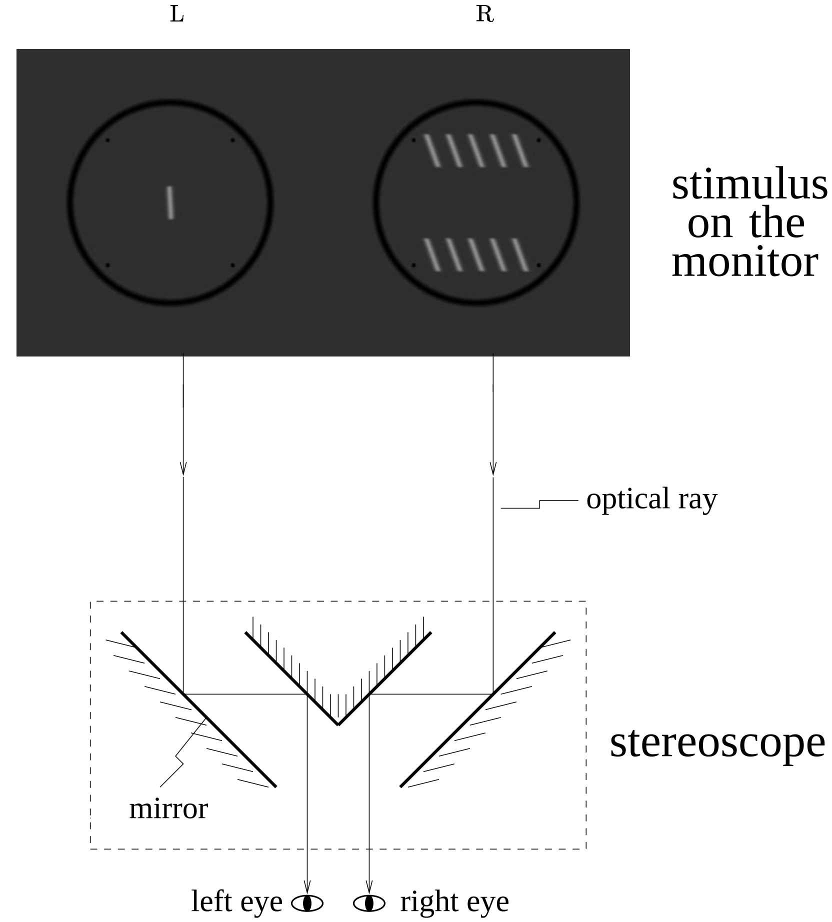
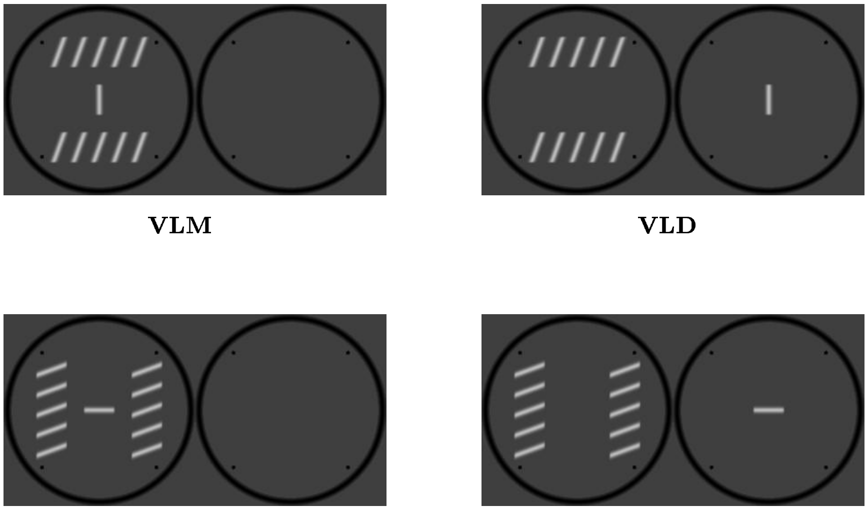


Fig. 2. A schematic view of the apparatus. L and R are the stimuli presented simultaneously on the monitor screen. They reach the subject's eyes through different optical paths, L reaches the left eye and R the right eye. The distance between the screen and the stereoscope is about 56 cm. By adjusting the positions of L and R on the screen, the subjects will get a good binocular fusion of them.

line in the right eye for several trials, as shown in Fig. 5(B). We estimate by this method that the vergence movement of the subjects was within 4.0 min arc on average during the data taking session.

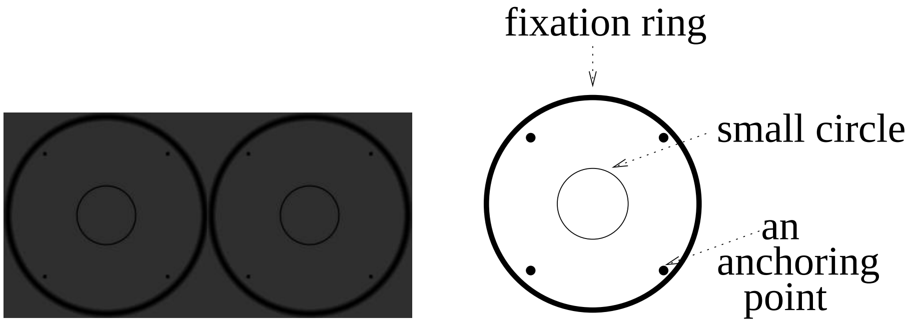
For each trial, the test line orientation was randomly selected from one of the 5 near vertical (or horizontal in the horizontal case) orientations 1 degree apart from each other. The subject was required to make binary judgements about the orientation of the test line by pressing the left or right button after each trial, to indicate whether the test line was perceived as tilted left or right at the top (for vertical case), or titled up or down at the right end (for the horizontal case).



VLM **VLD**

HLM **HLD**

Fig. 3. Some examples of stimulus patterns. VLM denotes that the test line is nearly vertically oriented and stimulates the left eye of subject, and the inducing lines are also viewed by the left eye . VLD is the same as VLM, except that the inducing lines are viewed by the right eye. HLM and HLD are the horizontal counterparts of VLM and VLD, respectively, which can be obtained by rotating VLM and VLD 90 degrees. The transfers with left eye viewing test line in the vertical and horizontal cases are then the ratio between the shifts in VLD and VLM, and ratio between the shifts in HLD and HLM, respectively. Then we can compare the transfers in vertical and horizontal cases to find whether there is a correlation between the interocular transfer and the orientation of the test line. The transfers with the right eye viewing the test line can be obtained analogously.



A B

Fig. 4. A: Two fixation rings with inner rings to minimize vergence movement. B: Schematics of A.

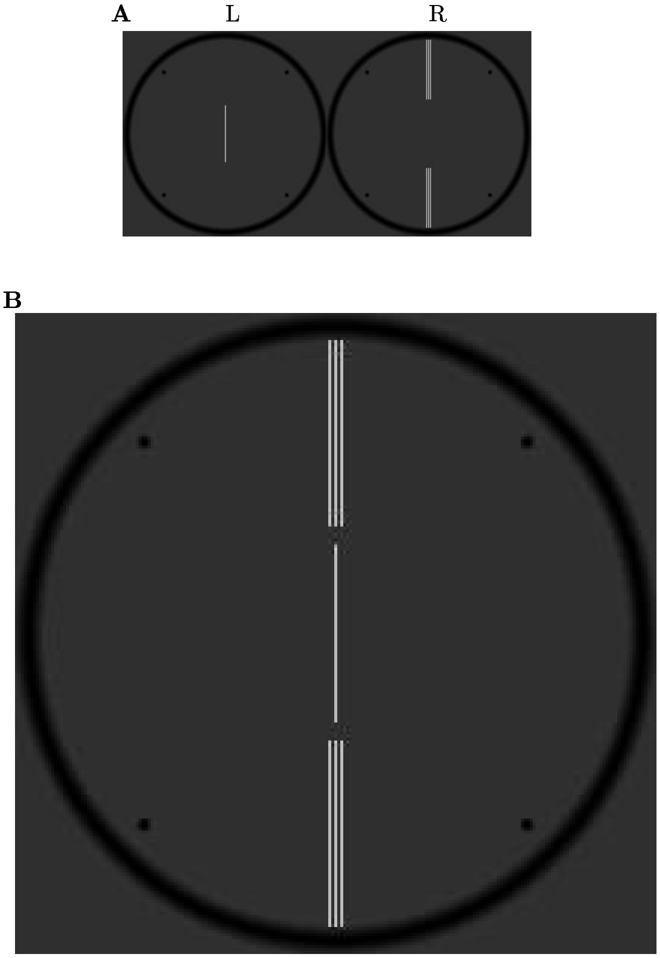


Fig. 5. A: The stimulus pattern used to adjust the positions of stimuli on the screen. B: The perceptual fusion of L and R in A, shown enlarged, when L and R are adjusted to the desired positions on the screen.

The next trial started after the subject responded. The test line orientation for which the subject pressed the left and right button equally probable was determined as the apparent vertical orientation for the subject. In principle, the difference between the apparent vertical (or horizontal) orientations with and without the inducing lines is the induced orientation shift. In practice, to avoid the systematic error from the possible observer bias, we ran each section with equal and opposite inducing line tilts in an interdigitated fashion, and the shift was measured as half the difference between the apparent vertical (or horizontal) orientations under the two opposite inducing line tilts. The interocular transfer was then the ratio between the dichoptic and monocular shifts for the particular test line orientation and particular eye viewing the test line.

With the same test stimuli, the experiments with and without natural scenes differed from each other in two points. First, between two trials, in the experiment with natural scenes, 8 stereo images selected randomly from 21 natural scenes were presented randomly one after another to the subjects for 6.14 sec to keep a natural visual environment. An example of the stereo images is given in Fig.6. So for the experiment with natural scenes, the test and inducing lines appeared roughly once every 9.14 sec, and for the experiment without natural scenes, the stimulus appeared roughly once every 3 sec. Second, in the beginning of the experiment without natural scenes, the subjects experienced roughly 600 sec adaptation to an unnatural visual environment. The adaptation procedure was the same as the following test procedure except that the inducing lines were removed, and the test line was presented binocularly. Two examples of adaptation stimulus pattern are given in Fig. 7. The subjects were also required to make binary judgements about the orientation of test line to concentrate their attention. The schematics of experiment with and without natural scenes are shown in Fig. 8 and 9, respectively.

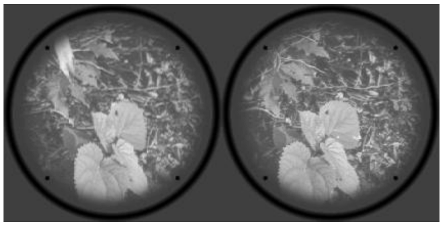
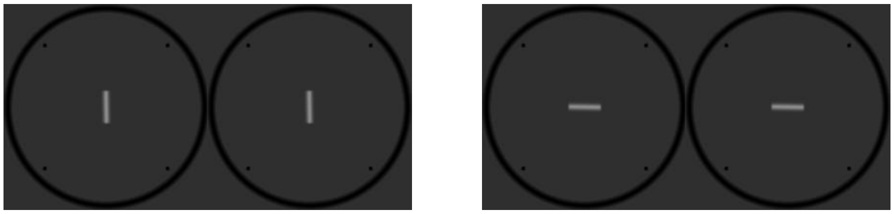


Fig. 6. An example of the natural stereo images , which were inserted between two trials in the experiment with natural scenes to keep a natural visual environment for the subjects.



Adaptation Stimulus in Vertical Case

Adaptation Stimulus in Horizontal Case

Fig. 7. The examples of stimulus pattern for adaptation to an unantural visual environment in the experiment without natural scenes. The left example was used in the vertical case and the right one in horizontal case.

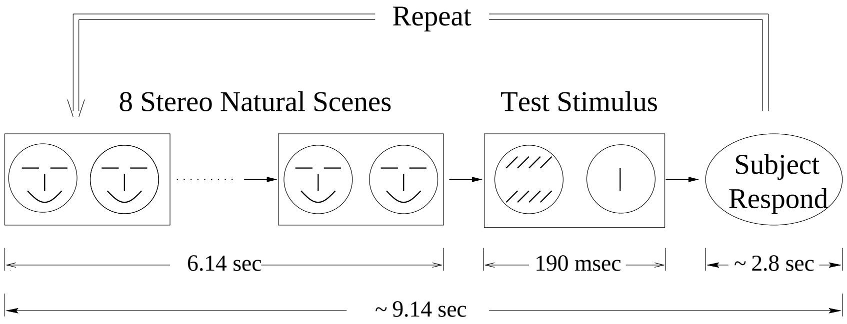


Fig. 8. The schematics of experiment with natural scenes, where natural stereo scenes are substituted by smiling faces.

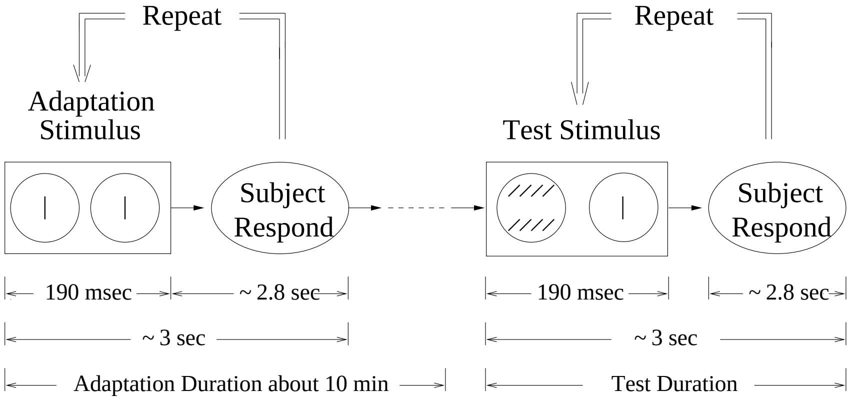


Fig. 9. The schematics of experiment without natural scenes as a control experiment.

3 Results and Discussion

Experiment results of two subjects(T.Z. and Z.Y.) are given in Fig. 10. As is seen in the figure, in the experiment with natural scenes, the transfers in vertical case are all smaller than their corresponding transfers in horizontal case. This result is consistent with the prediction. However, in the experiment without natural scenes, the transfers in vertical case are all approximately equal to their corresponding transfers in horizontal case. In a controlled way, this result also agrees with the prediction. It shows that natural and unnatural visual inputs will cause quite different neural properties in binocularity so as to keep the efficiency of stereo coding. This work strengthens our belief in efficient coding as a framework for predicting and understanding neural processing in the cortex.

Also this work is an example where a computational theory leads to predictions that can be experimentally tested, and the experiment in turn confirms the prediction. A directer illustration of the result is shown in Fig. 11.

Subject: T. Z.

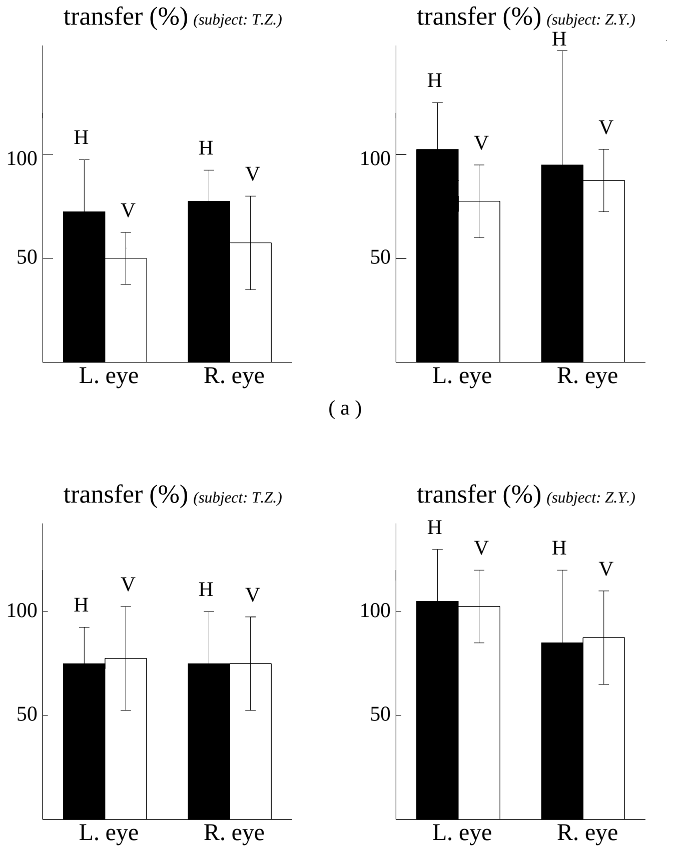
Stimulus Pattern	Experiment with Natural Scenes		Experiment without Natural Scenes	
	Shift (degrees)	Transfer	Shift (degrees)	Transfer
VLD	0.62± 0.11	51% ± 12%	0.63 ± 0.15	79% ± 24%
VLM	1.22± 0.18		0.80 ± 0.14	
HLD	0.80± 0.20	74% ± 26%	0.83 ± 0.15	75% ± 19%
HLM	1.08± 0.28		1.10 ± 0.19	
VRD	0.25± 0.06	58% ± 23%	0.69 ± 0.15	74% ± 22%
VRM	0.43± 0.13		0.93 ± 0.19	
HRD	0.81± 0.14	78% ± 17%	0.59 ± 0.21	74% ± 34%
HRM	1.04± 0.14		0.80 ± 0.24	

Subject: Z. Y.

Stimulus Pattern	Experiment with Natural Scenes		Experiment without Natural Scenes	
	Shift (degrees)	Transfer	Shift (degrees)	Transfer
VLD	0.81± 0.13	79% ± 17%	0.86 ± 0.08	101% ± 17%
VLM	1.02± 0.13		0.85 ± 0.12	
HLD	0.81± 0.08	101% ± 22%	0.60 ± 0.10	103% ± 25%
HLM	0.80± 0.15		0.58 ± 0.10	
VRD	0.79± 0.11	88% ± 15%	0.65 ± 0.12	88% ± 23%
VRM	0.90± 0.08		0.74 ± 0.13	
HRD	0.33± 0.11	94% ± 55%	0.43 ± 0.18	86% ± 34%
HRM	0.35± 0.16		0.50 ± 0.03	

Fig. 10. The results of two subjects both agree with the prediction.

In addition, this experiment shows that early visual coding is adaptive in a rather fast time scale, in the order of minutes. This supports the idea that sensory coding is ecological, and can be derived from the statistics of the sensory environment.



(a)

(b)

Fig. 11. Plots depicting the magnitude of transfers. The heights of filled and open bars show the size of transfer in the horizontal and vertical cases respectively. The eye viewing the test line is shown below the histogram. (a)Results from the experiment with natural scenes; (b)Results from the experiment without natural scenes.

Acknowledgement

We would like to thank Prof. Westheimer for conversation and encouragement for the experiment. Thank Dr. Liu Zhengyou, Mr. Qian Tiezheng and Dr. Wu Si for their patience and serious attitude to the experiment.

References

1. Barlow, H.B. 1961, "Possible principles underlying the transformation of sensory messages." Sensory Communication ed W. A. Rosenblith (Cambridge, MA: MIT Press), and "Unsupervised learning", *Neural Comp.*, **1**, 295-311, 1989.

2. Li, Zhaoping and Atick, J.J., "Efficient stereo coding in the multiscale representation", *Network: Computations in Neural Systems* **5**, 1-18, 1994.
3. Mitchell, D.E. and Ware, C., "Interocular transfer of a visual after-effect in normal and stereoblind humans", *J. Physiol.* **236**, 707-721, 1974.
4. Westheimer, G., "Simultaneous orientation contrast for lines in the human fovea", *Vision Res.* **30**, 1913-1921, 1990.

This article was processed using the LaTeX macro package with LLNCS style