Hierarchical processes of binocular vision and depth perception

by

HE Shufang

Student ID Number: 1196016

A dissertation submitted to the
Engineering course, Department of Engineering,
Graduate School of Engineering,
Kochi University of Technology,
Kochi, Japan

Dissertation Assessment Committee:

Supervisor: Hiroaki Shigemasu
Co-Supervisor: Kiyoshi Nakahara
Co-Supervisor: Hiroshi Kadota
Shinichi Yoshida
Yukinobu Hoshino

March 2018

ABSTRACT

Hierarchical processes of binocular vision and depth perception

HE Shufang

Graduate school of Engineering

Doctor of Engineering

Three-dimensional (3D) world is full of information. Human visual processing to the 3D scene can be divided into three stages: encoding, selection and decoding. It is quite important to clarify the underlying mechanisms. In this dissertation, two studies were implemented to investigate the binocular vision at different levels of brain processing and the mechanisms of depth perception.

In part I, it was reported that the vergence change could be used as a tool to detect the visual attention, visual memory and so on. To view the dichoptic inputs, there might be a perceptual rivalry over time between the left and right eyes. There was also a prediction that the visual information was processed in an efficient way to take advantage of the limited brain resources. Based on this prediction, an efficient coding theory was proposed that the correlated binocular inputs can be decorrelated into the binocular summation (S_+) and binocular difference (S_-) channels in V1 brain area for further processing. Based on this theory, previous study used special designed dichoptic stimuli with motion, tilt or color features as a tool to investigate the differences between the central and peripheral vision. In the study, for the monocular inputs, the percepts were ambiguous features; whereas for the S_+ and S_- channels, the percepts were unambiguous features. Participants' task was to judge the features (e.g., the motion direction, tilt orientation, or color). Whether the S_+ channel or S_-

channel was the dominant percept was analyzed. The results showed there was a bias towards S_+ percept at both the central and peripheral visions, suggesting the involvements of the analysis-by-synthesis computation and the prior knowledge that the binocular inputs were correlated. The results also showed that the bias towards S_+ at the central vision was larger than that at the peripheral vision, which suggested that the top-down feedback at the central view condition might be stronger than that at the peripheral view condition, indicating the functional difference between the central and peripheral vision. To have further investigations on the hierarchical processes of the binocular vision, this study continued the previous study by asking two questions: (1) whether the vergence eye movement is involved in the process? (2) Since the two eyes' inputs are dichoptic stimuli which may cause the rivalry, how are the temporal dynamics of S_+ and S_- percepts?

To this end, experiment 1 used the motion dichoptic stimuli as those in previous study and recorded the binocular eye information simultaneously. The changes in horizontal vergence, vertical vergence and pupil size were analyzed and the results showed that there was obvious change in vertical vergence at the central view condition; whereas no such tendency at the peripheral view condition. Matched sample t-test showed there was a significant difference of change in vertical vergence between the two view conditions at around $t > 700$ ms after the dichoptic stimuli onset. The stimuli used in this experiment were the summation or differencing of two horizontal gratings, which had independent random phases of each grating. To perceive the motion direction, the left and right eyes needed to integrate the ambiguous gratings into S_+ and S_- channels based on the top-down feedback. The integration might cause the change in vertical vergence. The significant difference had longer latency after the dichoptic stimuli onset, which might be because multi-cycles of the feedforward-feedback-verify-weight (FFVW) processes are involved before participants made their choices.

Experiment 2 investigated the temporal dynamics of the S_+ and S_- percepts by elongating the time duration of the dichoptic stimuli to 300 s. Participants' real-time behavioral responses (S_+, S_- and neither percepts) and binocular eye information were recorded simultaneously. The moments when the percept changing from S_+ to S_- and also from S_- to S_+ were extracted separately within a defined time window t = [-4000, 4000] ms. The change in horizontal vergence, vertical vergence and pupil size were analyzed, and the probabilities of the blink and (micro)saccade were calculated. The results show that: (1) the change in horizontal vergence is involved at the central vision when the percept changing from S_+ to S_- and also from S_- to S_+, which might indicate the involvement of the high level visual attention; (2) the temporal dynamics of the S_+ and S_- percepts at the central view condition show a tendency of competition between the S_+ and S_-. Since the visual recognition (S_+ or S_- percepts) at the central vision involves the feedback from the higher brain areas, the temporal dynamics of the S_+ and S_- percepts might indicate the hierarchy of the binocular rivalry, which is different from the direct competition between the visual inputs in the conventional binocular rivalry; (3) the neither percept at the peripheral view condition is significantly larger than that at the central view condition, suggesting the different underlying mechanisms, in which the former could tolerate for the binocular conflicts, whereas the latter could not tolerate for the binocular uncorrelated information, thus works in "winner-take-all" manner; (4) the temporal dynamics of the S_+, S_- and neither percepts and the quantities of percept flips might indicate the involvement of adaptation; (5) the change in pupil size can be explained by the existence of the LC-NE complex, which indicates the involvement of the visual attention when percept flips.

In part II, two experiments were implemented to investigate the mechanisms of depth adaptation of the disparity-specified sinusoidal corrugations. Experiment 3 investigated the phase-dependency and independency, and also the orientation-independency of depth

adaptation by using disparity-specified horizontal, vertical corrugations and plaids. The results suggest that the relatively early level is involved in depth adaptation. Experiment 4 used the plaids, noise-structure and horizontal corrugations as stimuli, and examined whether the different shapes of adaptor-probe pairs could cause the similar amounts of adaptation. The results suggest that both the disparity- and shape-level depth adaptation are involved. The continuous surface of horizontal corrugation adaptor causes the larger amount of depth adaptation than the noise-shape condition, suggesting the importance of surface structure in depth adaptation.

Keywords: hierarchical processes; vergence eye movement; top-down feedback; binocular vision; depth adaptation

ACKNOWLEDGMENTS

I want to give my deep thanks to my supervisor Dr. Shigemasu Hiroaki. Thanks Shigemasu sensei for accepting me in his laboratory. I want to thank Shigemasu sensei for creating an easy lab atmosphere and organizing many activities to shorten the distance of the lab members. I want to thank Shigemasu sensei for guiding on the study of depth perception, and also for supporting me to learn from Professor Li Zhaoping. I want to thank Shigemasu sensei for providing me many chances to join the conference, which were good for my research and also for broadening my view to the world.

I want to give my deep thanks to Prof. Li Zhaoping from University College London for her great supports on my study. I want to thank Prof. Li for suggesting the project idea for the Part I study in this thesis. Prof. Li also gave me a lot of useful suggestions and comments on how to do research in a scientific way, how to organize reports in a more logical way and how to do critical thinking. She always replied me long and detailed emails no matter it was already the late night of her time or weekend. Her attitudes towards the scientific research deeply influenced me.

I would like to express my great appreciation to my subsupervisors, Dr. Nakahara Kiyoshi, Dr. Kadota Hiroshi, Dr. Yoshida Shinichi, and Dr. Hoshino Yukinobu who gave me many useful and helpful suggestions and comments to my study.

I'd like to thank all the lab members for their friendships and supports on my study, especially to the visiting researcher Dr. Hatano Aya. I want to thank Hatano san for helping me on my Japanese, sharing the experiences of her Ph.D study, and also the daily communication.

I'd like to thank all the IRC members, especially to Professor Sakikawa Shinichiro. I

appreciated for all the friends in KUT, especially to Dr. Hou Shengwen, Zhang Qiang, Ruedeerat Keerativittayayut (Fai), and Phung Hai Thien An.

I appreciated for my friends Shi Haixia, Zhang Kaihe, Dr. Lian Yusheng and Li Yu in China. They gave me a lot of encouragements and supports on my study.

Last, I would like to thank my family. Thank my parents, brothers and other family members for always encouraging and supporting me to pursue knowledge.

Contents

Chapter 1 Outline 1
Chapter 2 Introduction 5
2.1 Efficient coding theory 5
2.2 Central and peripheral vision 7
2.3 Vergence eye movement 9
2.3.1 Physiology and neural mechanisms 9
2.3.2 Previous studies of vergence eye movement 10
2.4 The purpose of this study 11
Chapter 3 Experiment 1: Top-down feedback for visual recognition and vergence eye movements 13
3.1 Methods 14
3.1.1 Participants 14
3.1.2 Apparatus 14
3.1.3 Stimuli 18
3.1.4 Procedure 21
3.1.5 Data analysis 23
3.2 Results 24
3.2.1 Behavioral data 24
3.2.2 Vergence Eye movements 27
3.3 Discussion 45
3.3.1 Vergence eye movements is involved 45
3.3.2 The influence of relative contrast and temporal frequency on the vergence eye movements 47
Chapter 4 Experiment 2: temporal dynamics for visual recognition involves vergence eye movements and hierarchical processes 49
4.1 Method 50
4.1.1 Participants 50
4.1.2 Stimuli and apparatus 50
4.1.3 Procedure 50
4.2 Results 52

4.2.1 Behavioral data 52
4.2.2 Vergence eye movements, pupil size, blink and (micro)saccade 58
4.3 Discussion 63
4.3.1 Different amounts of the F_+, F_- and F_N suggest different underlying mechanisms between the central and peripheral visions 63
4.3.2 Temporal dynamics of the S_+ and S_- percepts at the central vision might indicate the hierarchy of binocular rivalry 64
4.3.3 Temporal dynamics and percept flips might indicate the involvement of adaptation 66
4.3.4 Vergence eye movement as a tool to detect the involvement of the higher brain areas 67
4.3.5 The change in pupil size might be explained by the existence of the LC-NE complex 69
4.4 Conclusion 70
Chapter 5 Introduction about depth adaptation 71
5.1 Depth adaptation and previous studies 71
5.2 The purpose of this study 74
Chapter 6 Experiment 3: Phase- and orientation-independency of depth adaptation 77
6.1 Experiment 3.1: retinal-dependency and independency of depth adaptation 77
6.1.1 Method 78
6.1.2 Results 82
6.1.3 Discussion 84
6.2 Experiment 3.2: Orientation-independency of depth adaptation 86
6.2.1 Method 86
6.2.2 Results 88
6.2.3 Discussion 89
Chapter 7 Experiment 4: Evaluation of disparity- and shape-level depth adaptation 92
7.1 Experiment 4.1: disparity- or shape-level depth adaptation 92
7.1.1 Methods 92
7.1.2 Results 97
7.1.3 Discussion 100
7.2 Experiment 4.2 depth adaptation by using surface and non-surface adaptors 102
7.2.1 Methods 102
7.2.2 Results 104
7.2.3 Discussion 106
7.3 General discussion 106

Chapter 8 General discussion 109
8.1 Dichoptic study 109
8.2 Depth adaptation study 110
References 113

Chapter 1 Outline

Three-dimensional (3D) world is full of information. Human vision transform the 3D information to the two dimensional (2D) images, and convey them to different levels of the brain areas. The visual processing can be divided into three stages: encoding, selection and decoding (Zhaoping, 2014). Encoding means transforming the input images, or to say the light signals, into retinal neuron activities. Selection means that only limited information can be selected for further processing, which is due to the limited brain resources. Decoding means the visual perception and how the brain recognize the selected information (Zhaoping, 2014).

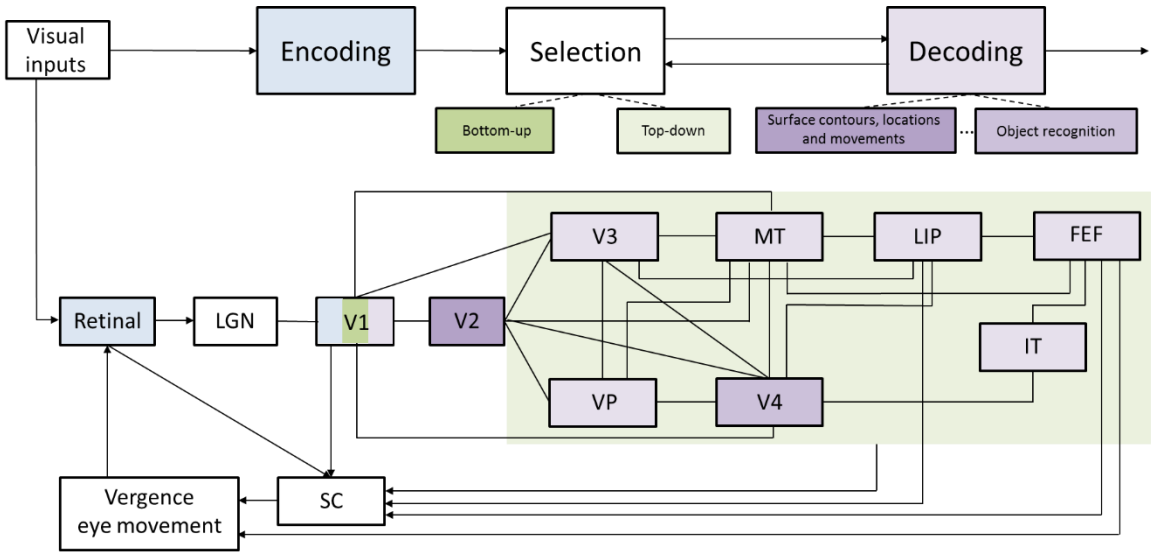


Fig. 1.1 Schematic of visual processing and the simplified physiological hierarchies. LGN is the lateral geniculate nucleus, MT (namely V5) is the middle temporal area, LIP is the lateral intraparietal area, FEF is the frontal eye field, IT is the inferotemporal cortex, and SC is the superior colliculus. The visual processing can be divided into three processes: encoding, selection and decoding. Physiologically, encoding mainly involves the retinal and V1 areas; selection can be divided into bottom-up selection which involves the V1 area, and top-down selection which involves the areas beyond V1; decoding links with wide brain areas. Signals from various brain areas can activate the SC (superior colliculus) neurons to control the eye movement (Zhaoping, 2014).

Figure 1.1 shows the schematic of visual processing and the physiological hierarchies. In the figure, LGN is the lateral geniculate nucleus, MT (namely V5) is the middle temporal area, LIP is the lateral intraparietal area, FEF is the frontal eye field, IT is the inferotemporal cortex, SC is the superior colliculus. Physiologically, encoding mainly involves the retinal and V1 areas; selection can be divided into bottom-up selection which involves the V1 area, and top-down selection which involves the areas beyond V1; decoding links with wide brain areas. Signals from various brain areas can activate the SC neurons to control the eye movement (Zhaoping, 2014).

How these different brain areas connect and work together? In this dissertation, we conducted two studies to investigate the visual percepts at different brain hierarchies. The first study aimed to investigate the hierarchical processes of binocular vision, which contained the feedforward, top-down feedback, vergence eye movements and the temporal rivalry and involved different levels of the brain areas; the second study aimed to investigate the mechanisms of depth perception by using the adaptation paradigm, which was mainly related to the visual decoding process.

In the first study, when seeing a scene in the natural world, there is only subtle difference between the left and right images. For the visual processing, these images are quite redundant. With the prediction that the visual information is processed in an efficient way to take advantage of the limited brain resources, previous studies (Barlow, 1961; Li and Atick, 1994) proposed the efficient coding theory to reduce redundancy of binocular information. In the efficient coding theory, they decorrelated the left and right eye images into uncorrelated binocular summation (S_+) and binocular difference (S_-) channels. Based on this theory, Zhaoping (2017) investigated the S_+ and S_- percepts at both the central and peripheral view conditions, and reported that the bias towards S_+ percept at central vision might be due to the stronger top-down feedback when comparing to that at the peripheral vision. To have further

investigations on the hierarchical processes of the binocular vision, this study continued the previous study by asking two questions: (1) whether the vergence eye movement is involved in the process? (2) Since the two eyes' inputs are dichoptic stimuli which may cause the rivalry, how are the temporal dynamics of S_+ and S_- percepts?

To clarify above questions, chapter 2 generally introduces the concepts of the efficient coding theory, the central and peripheral vision, vergence eye movement, and the motivation of this study.

Chapter 3 investigates whether the top-down feedback at central view condition can cause the vergence eye movement based on the previous study (Zhaoping, 2017). In the experiment, a mirror stereoscope was used to observe the dichoptic stimuli and an eye tracking device was used to record the binocular eye information simultaneously. The results show there is significant difference of change in vertical vergence between the central and peripheral conditions, suggesting the vergence eye movement is involved in the visual process with the higher level feedback.

Chapter 4 clarifies the temporal dynamics of the S_+ and S_- percepts at the central and peripheral view conditions by increasing the stimulus time duration to 300 s. Participants' real-time percepts were reported by continuous key press, and the binocular eye information was recorded simultaneously. The results show there is a significant difference of change in horizontal vergence between the central and peripheral conditions when percept changing from S_+ to S_- and also from S_- to S_+; the hierarchy of the binocular rivalry might be involved at the central view condition. Moreover, both the temporal dynamics of the S_+, S_- and neither percepts and the quantities of percept flips might indicate the involvement of adaptation.

The second part mainly focuses on the depth perception, which is the decoding process and might involve the early and middle level brain areas.

Chapter 5 summarizes the previous studies, and clarifies the purpose of this study.

Chapter 6 investigates the phase and orientation-dependency of depth adaptation by using the horizontal, vertical and plaid stimuli. The results show that both phase dependency and independency are involved in the depth adaptation; whereas the orientation is independent of the depth adaptation. These results suggest the early level of depth adaptation.

Chapter 7, the sinusoidal corrugation, plaid and noise-pattern were used as adaptors to investigate whether the disparity or shape level of depth adaptation is involved. The results show that both the disparity- and shape-level depth adaptation are involved in the processes. Meanwhile, the continuous surface of horizontal corrugation adaptor causes the larger amount of depth adaptation than the noise-shape condition, suggesting the importance of surface structure in the depth adaptation.

Chapter 8 shows the general discussion about the two studies in this dissertation.

Part I: Binocular vision study

Chapter 2 Introduction

2.1 Efficient coding theory

When seeing the three dimensional (3D) world, there is only subtle difference between the left and right eye images. This difference is called the binocular disparity. For the visual processing, these similar images are quite redundant. This was a prediction that visual processing might transform information in an efficient way because of the limited brain resources (Barlow, 1961; Li and Atick, 1994; Zhaoping, 2014).

Based on this prediction, Li and Atick (1994) proposed an efficient stereo coding theory to reduce the binocular redundancy by eliminating the inter-ocular correlation and spatial correlation. To eliminate the inter-ocular correlation, they introduced two new variables S^+ and S^- shown as equation (2.1) and (2.2). In the equations, the $S^+ \; and \; S^-$ represent the ocular summation and difference respectively; and the $S^L \; and \; S^R$ represent the photoreceptor activities at the left and right retinal locations respectively.

$$S^+ = \frac{1}{\sqrt{2}}(S^L + S^R) \tag{2.1}$$

$$S^- = \frac{1}{\sqrt{2}}(S^L - S^R) \tag{2.2}$$

After above linear transformation, the $S^+ \; and \; S^-$ are uncorrelated.

To eliminate the spatial correlation, Li and Atick applied whitening filters $K^{\pm}(f)$ to the

signals $S^{\pm}$ since the power spectrum distribution of white noise was flat. The $K^{\pm}(f)$ are shown as equation (2.3), in which $R^{\pm}$ is the power spectra of $S^{\pm}$; f is the frequency, r is the correlation index.

$$K^{\pm}(f) \propto (R^{\pm}(f))^{-1/2} = |f|/\sqrt{1 \pm r} \quad (2.3)$$

After applying the whitening filters, the output $O^{\pm}(f) = K^{\pm}(f)S^{\pm}(f)$ became constant in frequency domain, indicating the reducing of spatial redundancy.

Then Li and Atick solved three issues before they used the model for prediction of binocular receptive fields: (1) based on the principle of the least deformation and least change to the original signal method, they solved the non-uniqueness problem during the decorrelation process; (2) defined a "low-pass" filter to solve the high-frequency noise problem; (3) and used lowpass and bandpass filters to solve the multiscale coding problem.

Finally, the receptive field distributions of the two eyes can be represented as equations (2.4), (2.5) and (2.6):

$$K_s^{L,R}(x) = \int df\ |K_s^{L,R}(f)|\cos(2\pi f x + \emptyset^{L,R}) \quad (2.4)$$

$$K_s^{+}(x) = K_s^{L}(x) + K_s^{R}(x) \quad (2.5)$$

$$K_s^{-}(x) = K_s^{L}(x) - K_s^{R}(x) \quad (2.6)$$

$\emptyset^{L,R}$ were the phases of the receptive fields.

Based on the relative sizes of $|K_s^{+}|, |K_s^{-}|$ and their relative angle, they predicted the disparity cells, like the tuned excitatory/inhibitory, near/far cells, monocular cells and so on.

Furthermore, May, Zhaoping and Hibbard (2012) and May and Zhaoping (2016) implemented several dichoptic experiments to investigate the motion and tilt percepts by adapting to static and un-tilted images. Both studies further demonstrated the efficient coding theory.

2.2 Central and peripheral vision

When directly seeing a point, the central visual area (also called fovea) is used. It takes up only about 0.1% in human visual field. The left 99.9% area of the visual field is known as periphery (Snowden et al., 2012; Rosenholtz, 2016). The central and peripheral vision are mainly different from following aspects.

Spatial resolution. Compared with central vision, when seeing with peripheral vision, it is harder to see the high spatial frequency gratings, while little influence to the low spatial frequency ones. This difference of spatial resolution between the central and peripheral vision is caused by the different size of their receptive fields, and different amounts of brain resources assigned. The central visual area (fovea) has small receptive fields, higher density of photoreceptors and retinal ganglion cells; whereas the peripheral visual area with large receptive fields, relatively lower density of retinal neurons (Snowden et al., 2012; Rosenholtz, 2016; Zhaoping, 2017). Moreover, the central visual area has different cortical magnification factor from the peripheral area. The former is about 100 times larger than the latter (Daniel and Whitteridge, 1961). Because of these differences, the peripheral vision has lower acuity to see fine things. To compensate for the differences, the stimulus size for the peripheral vision needs to be doubled around every 2.5 degree eccentricity angle if we use the eye chart as an example (Snowden et al., 2012).

Visual attention. The visual inputs to human eyes are massive and at about 10^9 bits/second speed. However, the limited brain resources can only do further processing to the visual inputs at about 100 bits/second speed (Zhaoping, 2014). Thus, only very small amount of information can be selected. How the brain selects and brings the related information for further processing is quite important. Based on Zhaoping (2014) and Zhaoping (2017), at least

two stages as “looking” and “seeing” are involved in this processing, which link with the peripheral and central vision respectively. When participants search the target oblique bar from many horizontal, vertical and oblique oriented bars, at first they need to move their attention to the peripheral targets by several saccades. Although they may get close to the target, they do not see it at this stage. Then the next saccade moves the attentional spotlight to the central vision, participants can see the target (Zhaoping, 2014; Zhaoping, 2017). These two stages demonstrate the different roles of peripheral and central vision at visual attention. The former takes responsibility to select the targets for further processing, whereas the latter focuses on the decoding by using the selected information.

Visual recognition. Another significant difference between central and peripheral vision is their different performance at visual recognition. Even compensated for the lower spatial resolution (i.e. enlarging the size of peripheral visual stimuli), the peripheral vision still has lower visual recognition ability (Strasburger, Rentschler, and Juttner, 2011; Zhaoping, 2017). One of the important reasons is crowding. Crowding means the relative low ability to recognize object with peripheral vision where the target nearby has surrounding stimuli (Levi, 2008). Comparing with spatial resolution, crowding has larger influence on the degradation of the peripheral vision at various eccentricities (Rosenholtz, 2016). Making clear the mechanisms of crowding could help us to understand what the bottleneck of peripheral vision is in object recognition and how the features are integrated into objects in brain processing (Levi, 2008). Recently some computational models were proposed (Balas, Nakano, and Rosenholtz, 2009).

Moreover, the peripheral vision might have more important roles than what we thought before (Rosenholtz, 2016). For example, the audiovisual processing was reported to be processed in and/or prefer to peripheral vision (Chen et al., 2017).

2.3 Vergence eye movement

Vergence eye movement is a kind of disjunctive eye movement, meaning that the left and right eyes do not rotate at the same direction, and can be used in the study of the binocular depth perception. For example, to perceive a near object, the two eyes will move in towards the nasal side and cause the convergence; and to perceive a far object, they will move out towards the temporal side and cause the divergence.

2.3.1 Physiology and neural mechanisms

Physiologically, the third, fourth and sixth cranial nerves send innervation to the three pairs of extraocular muscles to drive the eye movement (Porter et al, 1983; Howard, 2012). There were many reports in the subcortical control of vergence that the Paramedian Pontine Reticular Formation (PPRF), the Mesencephalic Reticular Formation (MRF), the cerebellum, the Nucleus Reticularis Tegmenti Pontis (NRTP) and the superior colliculus are related to vergence eye movements (Baker and Highstein, 1975; Gamlin et al., 1996; Gamlin and Clarke, 1995; Gamlin, 2002; Judge and Cumming, 1986; Keller, 1989; King et al., 1994; May, 1984; May et al., 1986; Moschovakis, 1995; Ohtsuka and Nagasaka, 1999; Sylvestre et al., 2003; Walton and May, 2005; Zhang et al., 1991, 1992; Zhang and Gamlin, 1998).

In the cortical control of vergence, the brain areas V1 and V2 were reported as preprogram for changes in vergence (Thomas et al., 2002). Cells in the suprasylvian area of the parieto-occipital cortex (Bando et al., 1984, 1996; Toyama et al., 1986; Toda et al., 1991, 2001, 2006), Medial Superior Temporal cortex (MST) (Sakata et al., 1983; Takemura et al., 2001, Akao et al., 2005b), LIP (Gnadt and Mays, 1995; Gnadt and Beyer, 1998), were also involved in vergence control. Moreover, the frontal eye fields of the cerebral cortex and NRTP were involved in the voluntary eye movements, saccades and pursuit eye movements

(Gamlin et al., 1996; Gamlin, 2002; Fukushima et al., 2002; Akao et al., 2005; Howard, 2012).

2.3.2 Previous studies of vergence eye movement

Recent studies reported the visual attention related vergence eye movements. For example, Sole Puig et al. (2013a) adopted cue/no-cue paradigm to investigate whether there was any difference of the vergence eye movement between these two conditions. In the cue condition, before the presentation of the test stimuli, a tilted bar at the central fixation position was used as a cue to orient to the position of the target stimuli that located at the peripheral vision (7.5 degree eccentricity); in the no-cue condition, there was no such information. In both conditions, participants were asked to focus on the central fixation and judge the orientation (leftwards or rightwards) of the target stimuli which located at the peripheral vision. The binocular eye information were recorded simultaneously and the results showed the change in vergence at the cue condition was significantly larger than that at the no-cue condition. Since the target stimuli located at the peripheral visual field might involve the covert attention in the cue condition, they concluded that the vergence eye movements might have relationship with the visual attention from the higher brain areas.

Similarly, Sole Puig et al. (2013b) accessed whether there was any link between the vergence eye movements and the spatial visual attention. To do so, they divided participants into the local field independent (FI) group and global filed dependent (FD) group and instructed them to do the cue/no-cue task as the previous study (Sole Puig et al., 2013a). The FI group means those who were not influenced by the surrounding stimuli when they did some tests before experiment, but the FD group were influenced. Results showed both the FI and FD groups had larger eye vergence at the cue condition than the no-cue condition, which was consistent with the previous study (Sole Puig et al., 2013a). Moreover, their results showed that the FI group had higher accuracy, faster speed to the task, and also larger

vergence eye movements than the FD group. As Sole Puig et al. (2013b) proposed, before the stimuli onset the vergence eye movements might do some preprocess by conveying signals across cortical areas for the incoming visual stimuli.

Moreover, Sole Puig et al. (2017) studied the relationship between the memory task and eye vergence, and found a correspondence between them. However, they did not find a clear correspondence between the pupil size and eye vergence. From the neural level, the frontal, parietal regions of the cerebral cortex and cerebellum, might be involved in the vergence eye movement. Hence, their results suggested the vergence eye movement was a factor in the high level visual attention.

Beside, Rambold et al. (2010) used two 1-D horizontal sinusoidal gratings with ¼ wavelength difference in phase as dichoptic stimuli (one grating for each eye). The stimuli were produced by modulating the contrast of a high-frequency carrier. They added various amounts of luminance for both eyes, and found the vertical vergence changes were dependent on the added luminance, all with latencies less than 150 ms, suggesting the bottom-up process.

2.4 The purpose of this study

Based on the efficient coding theory, Zhaoping (2017) investigated binocular summation (S_+) and binocular suppression (S_-) percepts at the central and peripheral view conditions, and reported that the bias towards the binocular summation (S_+) percept at the central vision might be due to the high level top-down feedback. However, this feedback was weaker or even absent at the peripheral vision. This study aimed to have further investigations on (1) whether the vergence eye movement is involved in the process? (2) Since the two eyes' inputs are dichoptic stimuli which may cause the rivalry, how are the temporal dynamics of S_+ and S_-

The purpose of this study

percepts?

Chapter 3 Experiment 1: Top-down feedback for visual recognition and vergence eye movements

Previous study (Zhaoping, 2017) reported the bias towards binocular summation (S_+) perception at the central vision was due to the higher level top-down feedback. However, this feedback was weaker or even absent at the peripheral vision. There were reports that the disparity neurons in V1 brain area might serve not only the disparity information, but also to guide for vergence eye movement (Cumming and Parker, 1997; Poggio, 1995). Previous studies (Zhaoping, 2014; Zhaoping, 2017) demonstrated that the feedback from the higher level to V1 mediated the V1's bottom-up selection and also cause the vergence eye movement via the mediation of the superior colliculus. The schematic was shown as in Figure 1.1 (Zhaoping, 2014).

Experiment 1 was implemented to investigate whether the vergence eye movement was involved in the process. The experimental setup and parameters were the same as in Zhaoping (2017), except that a four-reflective-mirror stereoscope and a head-mounted eye tracking device were used in this study. The eye tracking device can record the binocular eye positions and pupil sizes simultaneously.

The hypothesis is that the signals from the top-down feedback may cause the V1 neurons to control the superior colliculus for mediating the vergence eye movement. Since the top-down feedback is involved, the vergence eye movement might have certain latency (> 150 ms after the stimuli onset).

3.1 Methods

3.1.1 Participants

Twenty-six students aged 19-34 years (18 male, mean age: 22.5) from Kochi University of Technology were recruited as participants. All of them had normal or corrected-to-normal vision, and were tested for their stereo acuity and motion acuity. Participants were naïve to the aims of the experiments and were compensated for their time. The author did not serve as participant. All experiments and procedures were approved by the Research Ethics Committee of Kochi University of Technology and conformed to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to experiments.

3.1.2 Apparatus

Stimuli were presented on a 22-inch CRT color display (RDF223H; Mitsubishi, Tokyo, Japan; 1024 × 768 pixels, 85 Hz frame refresh rate). The luminance of the display was measured using a CS-100A colorimeter (Minolta, Japan) and linearized using a look-up table method. The program was created by using Matlab (Mathworks, Natick, MA, USA) with PsychToolbox Version 3 to present the experimental stimuli (Brainard, 1997; Pelli, 1997). During the experiments, participants sat in a dark room fronto-parallel to the surface of the display and observed the stimuli through a mirror stereoscope (Edmund Optics, USA). A chin rest was used to prevent head movement. A head-mounted eye tracking device (Eyelink II, SR Research, Canada) was used to record the eye information with 250 Hz sampling rate.

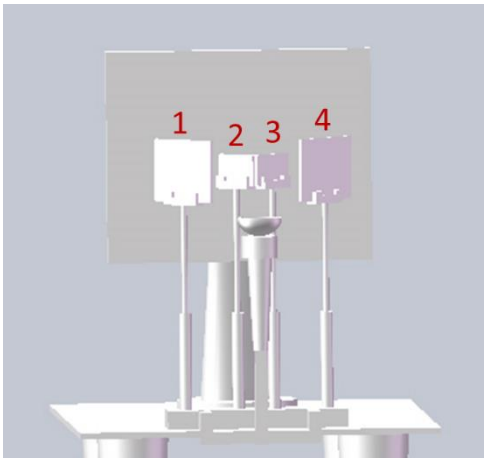
3.1.2.1 The mirror stereoscope

The mirror stereoscope was assembled with four 4 - 6λ first-surface coated reflective mirrors. The parameters of the mirrors are shown in table 3.1.

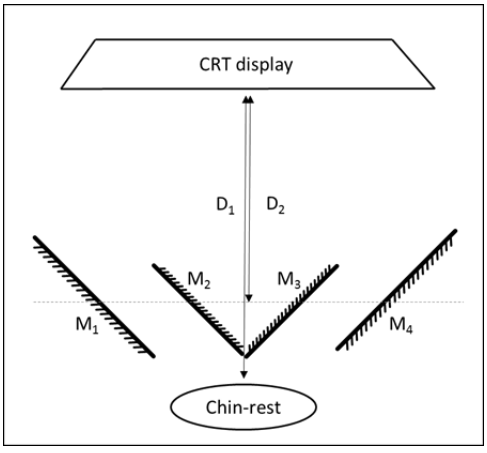
Table 3.1 Parameters of the mirrors

Mirror	Size (width × height)	Thickness	Reflective rate	Coating
1, 4	12.7 cm × 10.2 cm	0.6 cm	> 85% @ 400-700 nm	Protected Aluminum (400-700 nm)
2, 3	7.6 cm × 5.1 cm	0.6 cm	> 85% @ 400-700 nm	Protected Aluminum (400-700 nm)

The 3D model of the mirror stereoscope system was built up with Solidworks 2017 (Dassault Systemes S.A, USA) as shown in Figure 3.1 (A), and the top view of the schematic was shown in Figure 3.1 (B).



(A) The 3D model of the mirror stereoscope



(B) The top view of the mirror stereoscope

Fig. 3.1: The 3D model and top view of the mirror stereoscope system. (A) The 3D model and (B) the top view of the mirror stereoscope. The mirrors 1 and 2 (also mirrors 3 and 4) were parallel; and the mirrors 2 and 3 (also mirror 1 and 4) were symmetric to the center line of this system. The distance between the chin-rest and the CRT

display was D_1 = 50 cm, and the distance between the mirror stereoscope and the CRT display was D_2 = 38 cm.

The centers of the CRT display and the four mirrors were at the same height (the distance to the table was 41.5 cm). The height of the chin-rest was adjustable in a range of 28.0 ~ 33.0 cm. The distance between the chin-rest and the CRT display was D_1 = 50 cm, and the distance between the center of the mirror stereoscope and the CRT display was D_2 = 38 cm.

The mirrors M_1 and M_2 (also mirrors M_3 and M_4) were parallel. Mirrors M_2 and M_3 (also mirror M_1 and M_4) were symmetric to the center line of this system with an adjustable angle. Before the experiment, the distances and angles between mirrors were adjusted to make sure the images for the left and right eyes were visible through the mirror stereoscope; then all parameters were fixed the same to all the participants during the experiment. All participants could fuse well with the mirror stereoscope, and no one reported the un-fusion problem.

3.1.2.2 The Eyelink II system and monocular calibration method

The Eyelink II system consisted of an EyelinkII host PC, a display PC, headband camera as shown in Figure 3.2. The EyelinkII host PC held the high speed eye tracking card, connection interfaces to the headband, functional software and provided powers for the four infrared markers. The display PC was used for experimental application, with which we could configure the Eyelink tracker, present experiment stimuli and so on. The host PC and display PC connected via an Ethernet cable to transfer data or send commands. The headband had two cameras and could track the binocular information simultaneously. It connected with the host PC via a headband cable.

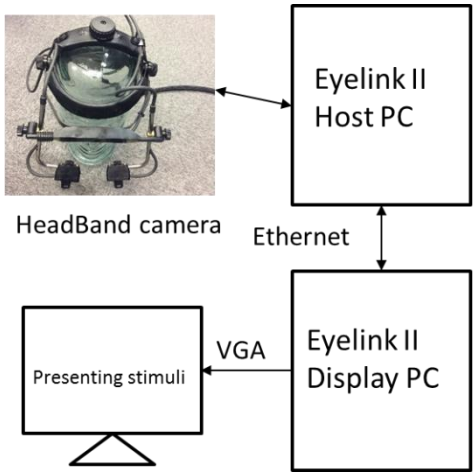


Fig. 3.2: Schematic of the Eyelink II system. The EyelinkII host PC held the high speed eye tracking card, connection interfaces to the headband, functional software and provided powers for the four infrared markers. The display computer was used for experimental application, with which we could configure the Eyelink tracker, present experimental stimuli and so on. The host PC and display PC connected via an Ethernet cable to transfer data or send commands. The headband had two cameras and can track the binocular information simultaneously. It connected with the host PC via a headband cable.

A 9-point calibration was implemented for the left and right eyes separately with our customized targets which were presented at the left and right halves of the CRT display respectively. During the experiment, the standard binocular tracking mode was used to record the pupil size and eye gaze information. The procedure was as below:

At first, enable (disable) the left (right) eye from the Eyelink host PC, present the calibration targets at the left side of the CRT display, and do the calibration and validation for the left eye.

Then, enable (disable) the right (left) eye and do the calibration and validation in the same way as previous step by using the targets shown within the outer frame at the right side of the screen. The schematic of the calibration targets was shown in Figure 3.3.

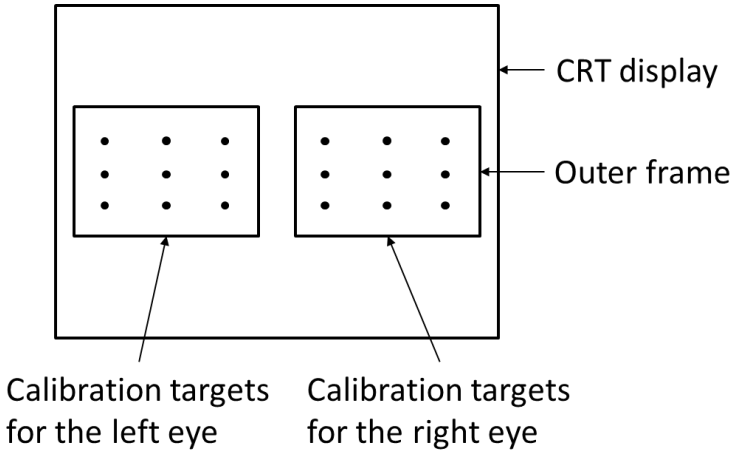


Fig. 3.3: Schematic of the calibration targets for the left and right eye

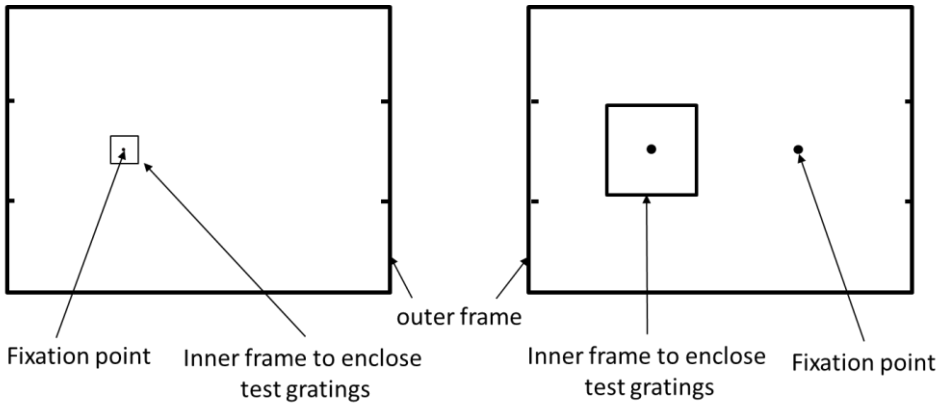
Finally, enable both eyes and record the binocular eye positions and pupil sizes simultaneously.

3.1.3 Stimuli

In experiment, the display was divided into the left and right halves to present the stimuli for the left and right eyes respectively. For each half of the display, an outer frame was drawn with 0.22° thickness lines. Different from Zhaoping (2017), a smaller outer frame was used to fit the size of our mirror stereoscope. The width and height were 19° and 15.86° respectively. But the sizes and positions of the inner frame and stimuli still kept the same as in Zhaoping (2017). At the two vertical lines of the outer frame, four spikes were located at the 1/3 and 2/3 positions to help for vergence. There were two kinds of view conditions: central and peripheral conditions. For the central view condition, an inner square frame was drawn with side size L = 1.13°; whereas for the peripheral condition, the side size was $L = 1.13° \cdot (1 + \frac{e}{e_2})$ with $e_2 = 3.3°$, e = 7.2° (left eccentricity). The line thicknesses for both inner frames were L/25.

At the center of each inner frame, a black disk was drawn with radius L/20. At the peripheral condition, the fixation was 7.2° right to the center of the inner frame, so both disks

shifted 3.6° from the center of the outer frame; at the central condition, the fixation was the center point of the inner frame, and located at the same position as the center disk of the inner frame at the peripheral condition. Figure 3.4 (A) and (B) represented the schematics of two binocular fixation images at central and peripheral conditions respectively (redrawn from (Zhaoping, 2017)).



(A) Central view condition; (B) Peripheral view condition

Fig. 3.4: Schematics of the two binocular fixation images as an example. (A) Central view condition; (B) Peripheral view condition. At the peripheral view condition, the fixation was 7.2 degree right to the center of the inner frame, so both disks shifted 3.6° from the center of the outer frame; at the central view condition, the fixation was the center point of the inner frame, and located at the same position as the center disk of the inner frame at the peripheral view condition (redrawn from (Zhaoping, 2017)).

As described in section 2.1, based on the efficient coding theory, the correlated binocular inputs can be decorrelated into binocular summation channel (S_+) and binocular suppression channel (S_-).

To do so, as described in Zhaoping (2017), the ambiguous gratings S_L and S_R were used as stimuli. The equations were (3.1) and (3.2).

$$S_L = \bar{S}[1 + (C_+ S_q + C_- S_{q'})/2] \quad (3.1)$$

$$S_R = \bar{S}[1 + (C_+ S_q - C_- S_{q'})/2] \quad (3.2)$$

In which, $\bar{S}$ was the background luminance. It was 20.8 cd/m^2 view through the mirror stereoscope in this experiment.

C_+ and C_- were the contrast of the two gratings. C_+ = 0.3. C_- = 0.66, 0.48, 0.3, 0.12.

$$S_q = \cos\left[k \cdot \left(y \mp 2\pi \frac{w}{k} t\right) + \emptyset_q\right] \quad (3.3)$$

$$S_{q'} = \cos[k \cdot \left(y \pm 2\pi \frac{w}{k} t\right) + \emptyset_{q'}] \quad (3.4)$$

$$k = \frac{4\pi}{L} \quad (3.5)$$

k was the spatial frequency, and w was the temporal frequency. $\emptyset_q$ and $\emptyset_{q'}$ were the phases for gratings S_q and $S_{q'}$ respectively. They were independent with each other, and uniformly distributed across all trials in one block. y

The gratings were presented with w = 2.5, 5 and 10 Hz. Based on equation (3.1) and (3.2), for each eye the stimulus was the summation of two gratings, so the appearance was ambiguous. Figure 3.5 was an example of S_L based on equation (3.1). The S_R was similar, thus omitted here.

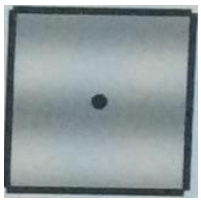


Fig. 3.5 An example of the ambiguous gratings S_L

However, for the binocular summation S_+ ($S_+ = (S_L + S_R)/2$) and binocular suppression S_- ($S_- = (S_L - S_R)/2$), based on equations (3.1) and (3.2), they could be perceived as the same direction of S_q and $S_{q'}$ respectively, which were unambiguous.

3.1.4 Procedure

Figure 3.6 showed the procedure of the experiment 1.

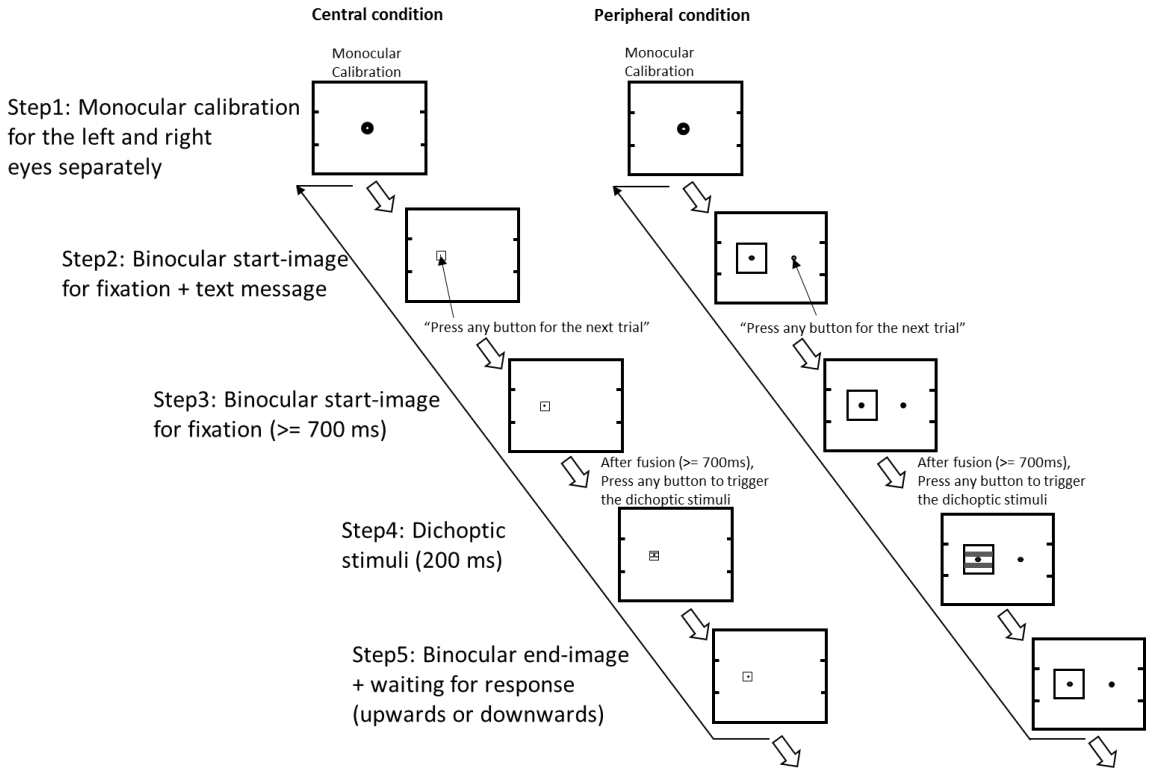


Fig. 3.6: Procedure of the experiment 1. The left and right images show the procedures at the central and peripheral view conditions respectively. At both conditions, the procedures contain: the monocular calibration for the left and right eyes separately at beginning of each trial (step 1); binocular start-image for fixation and the text message "press any button for the next trial" to instruct participants to trigger for the next trial (step 2); binocular start-image for fixation, after fusion (>= 700 ms) participants were instructed to press a button to trigger the dichoptic stimuli (step 3); presentation of dichoptic stimuli (time duration 0.2 s) (step 4); participants were asked to judge the motion direction (upwards or downwards) by key press (step 5).

At the beginning of each block, participants were asked to put their heads on the chin-rest, and check monocularly and binocularly whether they could see the whole outer frame through the mirror stereoscope. The purpose of checking was to make sure all the calibration dots were in the range of view. After checking, monocular calibration for the left and right eyes separately was done by using the Eyelink II system (Step 1). The calibration method was described in section 3.1.2.2.

After calibration, the binocular start-image for fixation and text message "press any button

for the next trial" were presented. Meanwhile, the Eyelink II started to record the binocular eye positions and pupil sizes simultaneously. The text message located below the fixation to reminder participants moving their eyes to the fixation point and starting the next trial. Once participants pressed a button to go to the binocular start-image for fixation step, they needed to gaze at the fixation for at least 700 ms. When participants confirmed that they were well fused, they were instructed to press a button to trigger the dichoptic stimuli, which would be presented for 200 ms.

Finally, the binocular end-image stimuli were presented, and waiting for participants key press. Participants' task was to judge whether the perceived gratings moved upwards or downwards based on two-alternative-forced-choice (2AFC). There was no time limit for their response. Once they made their choice by key press, the binocular start-image for fixation and text message for the next trial would be triggered and presented automatically. The Eyelink II marked all the step information and kept recording during the whole procedure.

During the experiment, participants could close their eyes for a short rest or blink at the "binocular start-image for fixation + text message" step before their key press. Once they started the "binocular start-image for fixation", they were asked to try their best not to blink or close their eyes till the next "binocular start-image for fixation" appeared. But the unconscious blink was allowed during the whole experiment. The purpose was to make sure that the Eyelink II device could record all the necessary eye information at each trial.

In one block, participants were asked to keep their heads as steady as possible, and keep the eyes on the fixation; nevertheless, there was a drift correction every two or three minutes to compensate for the possible head movements. After each trial, participants were asked to take their heads off the chin-rest, remove the Eyelink II and have a short break. This was done to avoid the fatigue to the eyes and necks caused by the high concentration during the experiment or the headache caused by the weight of the head-mounted device. After break,

participants were asked to redo the calibration before the next trial. For each participant, there were two sessions which needed to be done at different days. Each session had four blocks, with 112 trials in one block. For all participants, they had following conditions as shown in table 3.2 at the central and peripheral view conditions separately.

Table 3.2. Experimental conditions. For conditions 1-6, we could get 64 trials data from each participant; for conditions 7-8, we could get 32 trials data from each participant.

Condition	C-	Temporal Frequency (Hz)
1	0.3	5
2	0.3	2.5
3	0.3	10
4	0.48	5
5	0.48	2.5
6	0.48	10
7	0.12	5
8	0.66	5

For conditions 1-6, 64 trials data were collected from each participant; for conditions 7-8, 32 trials data were collected for each participant. All conditions were randomly presented within one block. No one could predict which condition would be shown.

3.1.5 Data analysis

The Matlab R2012a and DataViewer software (SR Research, Canada) were used to analyze the behavioral and eye tracking data.

For the eye tracking data, we calculated the horizontal vergence in the following steps: (1) define a time window t = [t_1, t_2] and extract the horizontal positions of the left and right eye positions within this period (e.g., denote as x(l, t) and x(r, t) respectively, t = [t_1, t_2]); (2) subtract the horizontal right eye position from the horizontal left eye position to get the horizontal vergence, x(hor, t) = x(l, t) – x(r, t)); (3) normalize the result in step (2) at t = t0 ($t_1 \leq t_0 \leq t_2$). e.g., x(hor_norm, t) = x(hor, t) - x(hor, t_0), t_0 is the time point for normalization; (4) convert the unit “pixel” into “degree” (based on the pixel per degree of the display); (5)

calculate the mean of the normalized horizontal vergence across all trials and all subjects. Because the horizontal vergence is calculated by subtracting the right eye position from the left eye position, the negative value means divergence.

Similarly, we also calculated the vertical vergence based on the above five steps, except that in the step 2, we subtracted the vertical position of the right eye from that of the left eye, e.g., y(ver, t) = y(l, t) – y(r, t). Hence, the negative value of the vertical vergence means the right-sursumvergence (the gaze position of the right eye is higher than the position of the left eye at the vertical direction).

Moreover, the mean pupil size was calculated by: (1) averaging the pupil sizes of the left eye (Pu(l, t)) and right eye (Pu(r, t)) within the same time window t = [t_1, t_2]; e.g., Pu(t) = (Pu(l, t) + Pu(r, t))/2; (2) normalizing at t = t_0; e.g., Pu(t_norm) = Pu(t) - Pu(t_0); (3) calculating the mean of the normalized pupil size across all trials and all subjects.

3.2 Results

Dichoptic Stimuli were presented in a random order with eight conditions and two view locations. Participants' behavioral and eye tracking data were recorded simultaneously at each trial.

3.2.1 Behavioral data

Fractions (F_+) of seeing S_+ drift direction were calculated at C_- = 0.3, w = 5 Hz for both central and peripheral conditions by using 26 participants' data. Figure 3.7 shows the results. The error bars are the standard errors of the means. Matched sample t-test is used to analyze the significance *p* between the two view conditions.

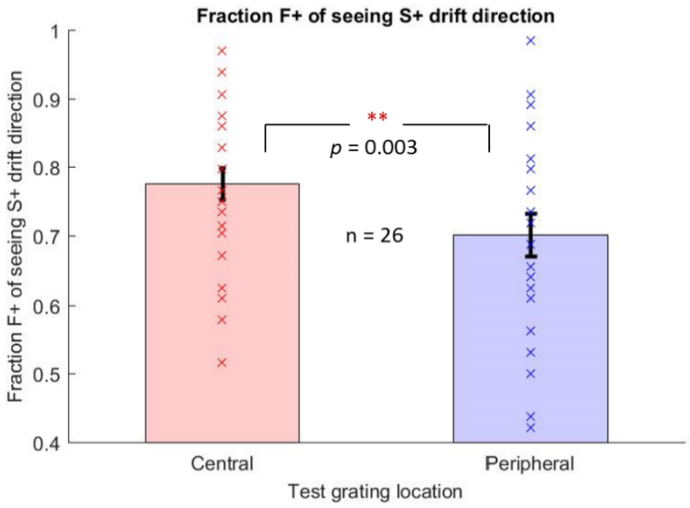


Fig. 3.7: Faction (F_+) of seeing binocular summation (S_+) drift direction is calculated at $C_- = 0.3$, $w = 5$ Hz at both the central and peripheral view conditions. The error bars are the standard errors of the means. The red asterisk mean there is a significant difference, with the marker as following meaning: $+p < .10$, $*p < .05$, $**p < .01$, $***p < .001$. The p value is calculated by using the matched sample t-test.

From the result, the fractions (F_+) of seeing the binocular summation (S_+) drift direction at the central and peripheral conditions are about 78% and 70% respectively, which match with the previous results as shown in (Zhaoping, 2017). There is a significant difference of fraction (F_+) of seeing binocular summation (S_+) drift direction between the central and peripheral conditions by using the matched sample t-test ($p = 0.003$, $t(25) = 0.072$).

Figure 3.8 shows the fraction (F_+) of seeing the S_+ drift direction as psychometric functions of the relative contrast $C_+/(C_+ + C_-)$ at $w = 5$ Hz temporal frequency. Error bars are the standard errors of the means.

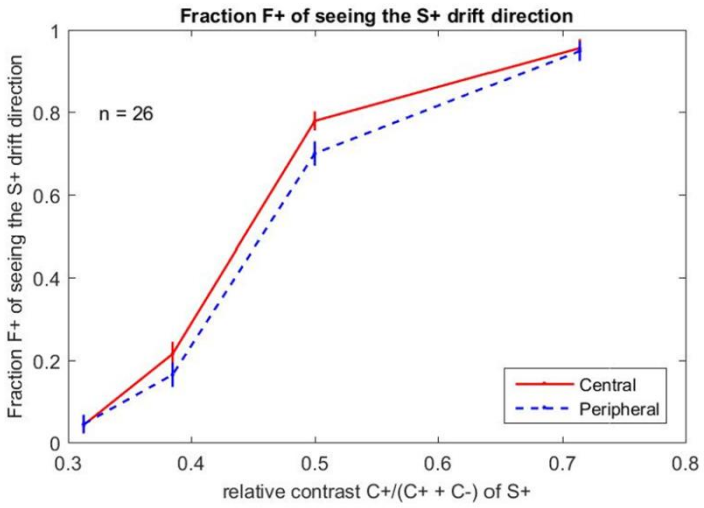


Fig. 3.8: Fractions (F_+) of seeing the S_+ drift direction are shown as psychometric functions of the relative contrast $C_+/(C_+ + C_-)$ at w = 5 Hz temporal frequency. Error bars are the standard errors of the means.

Figure 3.9 shows the fraction (F_+) of seeing the S_+ drift direction at different temporal frequencies. There are significant differences between temporal frequencies 2.5 Hz and 10 Hz (p = 0.001), and also between 5 Hz and 10 Hz (p = 0.000) at the central view condition; whereas there is no significant difference among the 2.5, 5 and 10 Hz at the peripheral view condition. Both results repeat the previous study (Zhaoping, 2017).

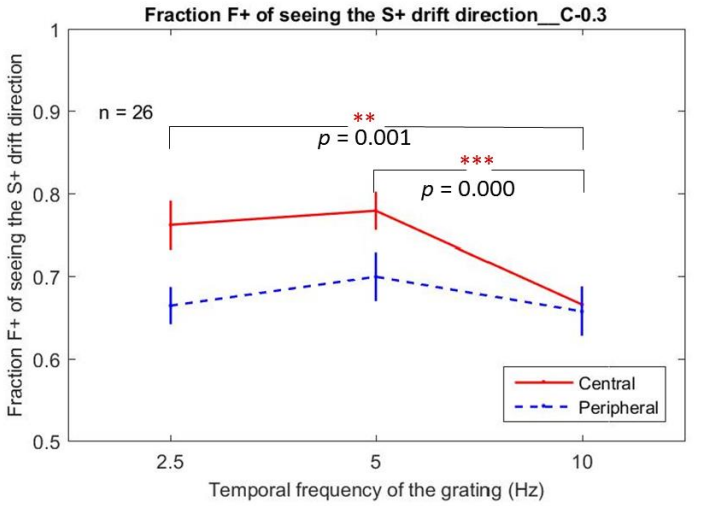


Fig. 3.9: Fractions (F_+) of seeing the S_+ drift direction are shown at different temporal frequencies. Error bars are the standard errors of the means. The red asterisk means there is a significant difference, with the marker as following meaning: $+p < .10$, $*p < .05$, $**p < .01$, $***p < .001$. The p value is calculated by using the matched sample t-test.

In Figure 3.7, as discussed in (Zhaoping, 2017), the bias towards S_+ percept at the central

view condition might be due to the stronger top-down feedback, whereas this feedback is weaker or even absent at the peripheral view condition. This can be explained by the Feedforward – Feedback – Verify – Weight (FFVW) model. When V1 neurons receive the visual inputs from the retina, both the S_+ and S_- channels will produce the feedforward signals to the higher brain areas. Then the higher brain areas give feedback mainly to the S_+ channel based on analysis-by-synthesis computation and the prior knowledge that the binocular inputs are correlated (Zhaoping, 2017).

In Figure 3.8, the psychometric curves and slopes of fraction (F_+) of seeing the S_+ drift direction between both central and peripheral view conditions have obvious differences, suggesting the percept bias towards the binocular summation (S_+) is not caused by the low sensitivity of the peripheral vision. In Figure 3.9, when the temporal frequency is higher (10 Hz), the fraction (F_+) of seeing the S_+ drift direction at central condition is lower. However, this was not the case for the peripheral vision. The fractions (F_+) do not have a significant difference at different temporal frequencies. These different responses at different temporal frequencies between the central and peripheral conditions suggest that the different neural mechanisms are involved in the visual processes. The central vision mainly work for visual decoding. The visual decoding ability might be constrained by the limited brain resources especially when the visual inputs have a relatively higher speed. However, the peripheral vision is mainly for the bottom-up selection, which is not influenced by the temporal frequency.

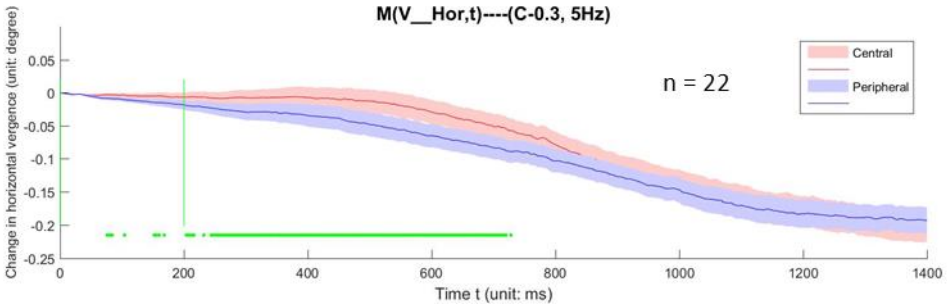
3.2.2 Vergence Eye movements

Eye positions and pupil sizes of 26 participants were recorded by the Eyelink II system. Four particpants' data were excluded because of the tracking data were too noisy (too much blinks or data missing, i.e. more than 30% trials). Those trials with blink or data missing in the time

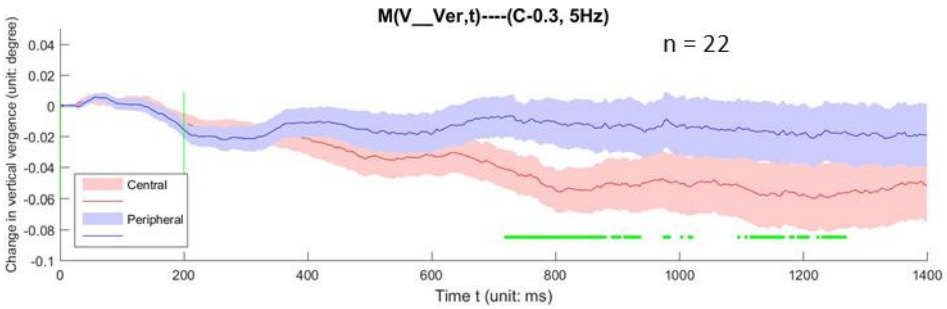
window from the other 22 participants were also excluded (around 9.1% across all participants). We defined t = [0, 1400] as the time window and extracted the horizontal vergence, vertical vergence and pupil size as the method described in data analysis section at each condition.

3.2.2.1 Vergence eye movement at C- = 0.3, w = 5 Hz

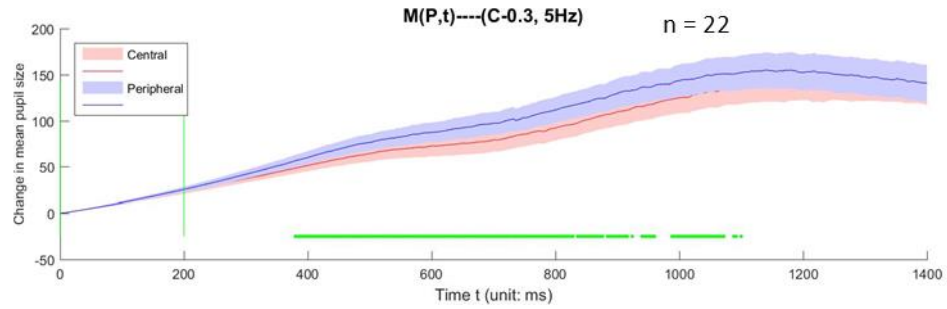
Figure 3.10 (A), (B) and (C) show the change in horizontal vergence, vertical vergence and pupil size respectively with C- = 0.3, w = 5 Hz. The horizontal axes of three figures are the time duration (unit: ms), the vertical axes are the change in horizontal vergence (unit: degree), change in vertical vergence (unit: degree) and the change in pupil size respectively. In all figures, the pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means. The horizontal green bars mean the time duration which have significant difference between the central and peripheral view conditions with t-test ($p < 0.01$). The two vertical green lines mean the dichoptic onset time duration, which have the same meaning in all the figures in this chapter, thus the explanation will be omitted to in the following figures.



(A) Chang in horizontal vergence



(B) Change in vertical vergence



(C) Change in pupil size

Fig. 3.10 the change in horizontal vergence, vertical vergence and pupil size at C- = 0.3, w = 5 Hz. (A), (B) and (C) show the change in horizontal, change in vertical vergence and change in pupil size respectively. The horizontal axes of three figures are the time duration (unit: ms), the vertical axes are the change in horizontal vergence (unit: degree), change in vertical vergence (unit: degree) and change in pupil size respectively. In all figures, the pink and blue curves mean the central and peripheral view conditions respectively. The shade of the curve means the standard errors of the means. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$). The two vertical green lines mean the dichoptic onset time duration.

The horizontal vergence is calculated by subtracting the mean right eye position from the mean left eye position, thus the negative value means divergence.

From Figure 3.10 (A), there are significant differences of change in horizontal vergence (around t = [200, 720] ms) between the central and peripheral view conditions. The amounts of change in horizontal vergence at the former condition are smaller than those at the latter condition. Our explanation is that since the peripheral stimuli have 7.2 degree eccentricity to the fixation, the eyes will diverge to perceive the distant stimuli after the presentation of the dichoptic stimuli, whereas less necessary at the central condition, thus cause the significant difference (see Solé Puig et al., 2013a).

From Figure 3.10 (C), there are significant differences of change in pupil size (around t = [400, 1100] ms) between the central and peripheral view conditions. The amount of change in pupil size at the former condition is smaller than that at the latter condition. One possibility is that the pupils dilate to diverge. Because the divergence is larger at the peripheral condition, it causes larger dilation of the pupil size. Moreover, the significant difference of change in pupil size has around 200 ms latency than that of the vergence change at both view conditions, which also suggests the link between the pupil size and horizontal vergence.

From Figure 3.10 (B), the distribution of change in vertical vergence at the central view condition has negative value, meaning the right-sursumvergence. The absolute amount increases over time. However, at the peripheral view condition, the distribution is quite stable and close to 0. There are significant differences of change in vertical vergence around t >= 700 ms between the two view conditions. Because the stimuli used in this experiment were the summation or differencing of two horizontal gratings, and the two gratings had independent random phases, thus the inputs for each eye were ambiguous gratings. Because of the high acuity of the central vision, to perceive the motion direction of the ambiguous inputs, the visual system need to integrate them into S_+ and S_- channels, which might cause the change in vertical vergence in the visual process. The negative value of change in vertical vergence suggested the right eye had higher position than the left eye. One possibility is that the right eye is the dominant eye for most participants. To integrate the ambiguous gratings, the right eye pay more efforts to realign the gratings. Since the low acuity of the peripheral vision, this integration process is not necessary. Moreover, the longer latency might be because that the FFVW takes time, or even several cycles of FFVW processes are involved for the visual recognition until the decision is made by the participant (Zhaoping, 2017).

In addition, to verify whether the change in vertical vergence is caused by something like artifact, a control experiment was implemented as below.

3.2.2.2 Control experiment

Since the ambiguous stimuli were used in experiment 1, which might involve the top-down feedback for visual recognition. In the control experiment, we used binocular stimuli which were unambiguous for both monocular and binocular percepts. To do so, we set $C_- = 0$; thus the visual inputs to the left and right eyes as equations (3.1) and (3.2) were the same as shown in equations (3.6) and (3.7):

$$S_L = \bar{S}[1 + C_+ S_q/2] \quad (3.6)$$

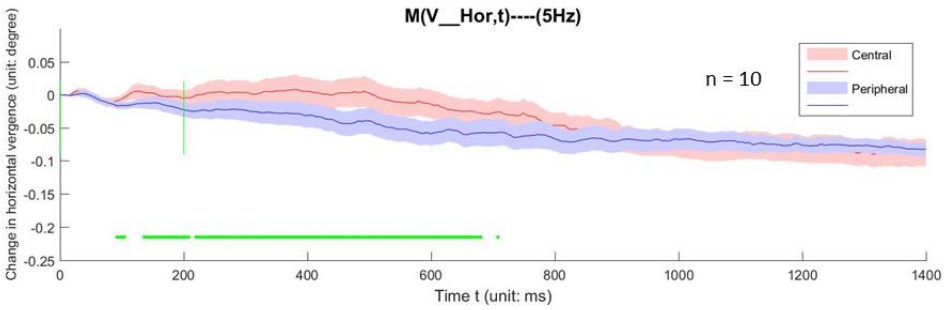
$$S_R = \bar{S}[1 + C_+ S_q/2] \quad (3.7)$$

The experimental setup and procedure were the same as experiment 1. Ten participants (mean age: 24.5 years, 8 males and 2 females) joined in the experiment. Since the motion directions of the left and right eyes were the same, the average correct rate of behavioral data from these ten participants were higher than 99%.

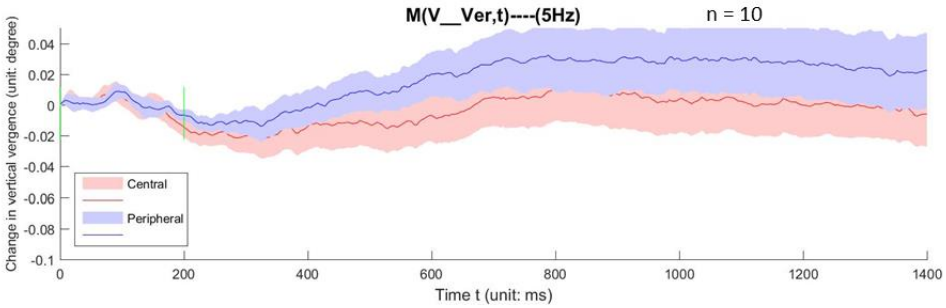
The eye tracking data were analyzed in the same way as experiment 1. Figure 3.11 (A), (B) and (C) show the change in horizontal vergence, vertical vergence and pupil size respectively of the control experiment. The horizontal and vertical axes have the same meanings as those in Figure 3.10. When comparing to the Figure 3.10 (A), there are also significant differences of change in horizontal vergence in the control experiment as shown in Figure 3.11 (A), suggesting the divergence to the distant stimuli as explained before.

When comparing to the change in vertical vergence in Figure 3.10 (B), there is no significant difference of change in vertical vergence in Figure 3.11 (B). Although the change in pupil size at the peripheral condition is larger than that at the central condition, there is no significant difference between them as shown in Figure 3.11 (C). The explanation to Figure 3.11 (B) and (C) is that the motion directions of the stimuli in the control experiment are

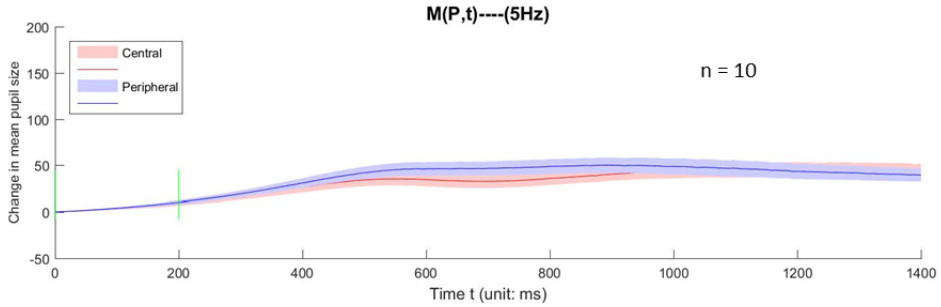
unambiguous, they are very easy to be judged. So the FFVW process is not necessary to be involved or to be finished, thus no vertical vergence eye movement was involved. Since participants feel very easy to do this task, they do not need to pay much attention, which might cause the smaller change in pupil size and also smaller difference between the two view conditions. The comparisons between the experiment 1 and control experiment also indicate that the involvement of the top-down process for the percept of the special designed dichoptic stimuli in experiment 1.



(A) Distribution of change in horizontal vergence



(B) Distribution of change in vertical vergence



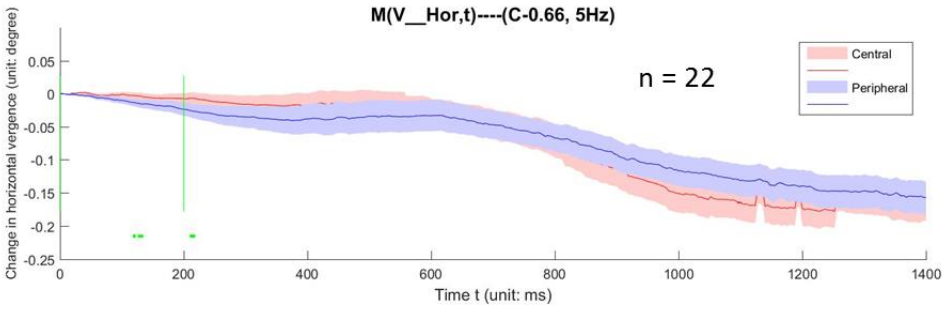
(C) Distribution of change in pupil size

Fig. 3.11: Distributions of change in horizontal vergence, vertical vergence and pupil size at C- = 0, w = 5 Hz in the control experiment. (A), (B) and (C) show the distributions of change in horizontal, vertical vergence and

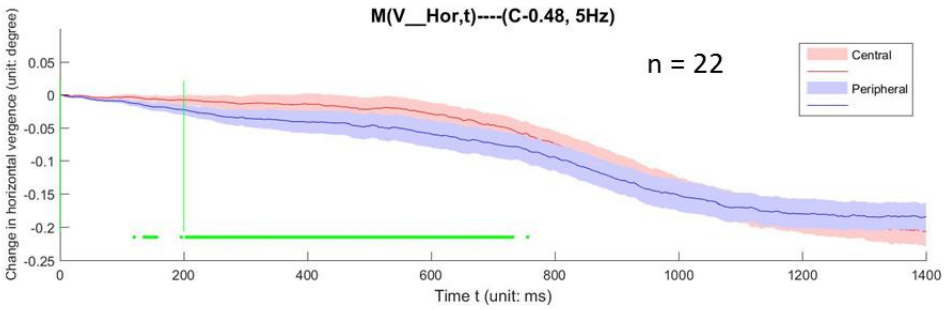
change in pupil size respectively at $C_{-} = 0$, w = 5 Hz. The horizontal axes of the three figures represent the time duration (unit: ms); the vertical axes represent the change in horizontal vergence (unit: degree), vertical vergence (unit: degree) and pupil size respectively. In all figures, the pink and blue curves mean the central and peripheral conditions respectively. The shades of the curve mean the standard errors of the means. The horizontal green bars mean the time duration which have significant differences between the central and peripheral conditions with t-test ($p < 0.01$). The two vertical green lines mean the presentation time duration of the dichoptic stimuli.

3.2.2.3 Vergence eye movement at different relative contrasts

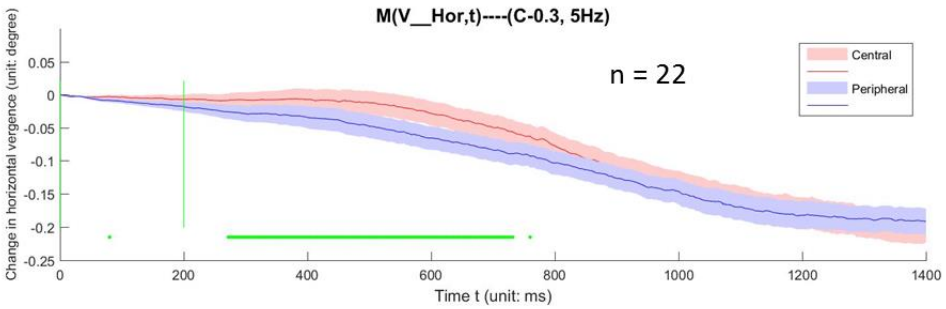
Figure 3.12, 3.13 and 3.14 show the change in horizontal, vertical vergence and pupil size respectively at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in horizontal vergence (unit: degree), vertical vergence (unit: degree) and pupil size respectively. The figures (A), (B), (C) and (D) represent at $C_{-} = 0.66$, 0.48, 0.3, and 0.12 conditions respectively. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$). The pink and blue curves mean the central and peripheral conditions respectively. The shades of the curves means the standard errors of the means.



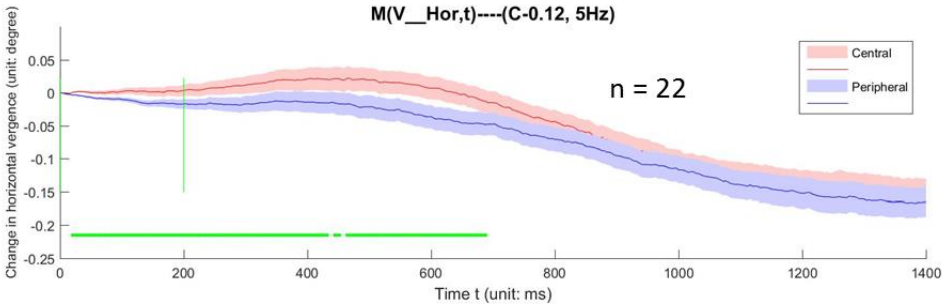
(A) $C_{-} = 0.66$, w = 5 Hz



(B) $C_{-} = 0.48$, w = 5 Hz

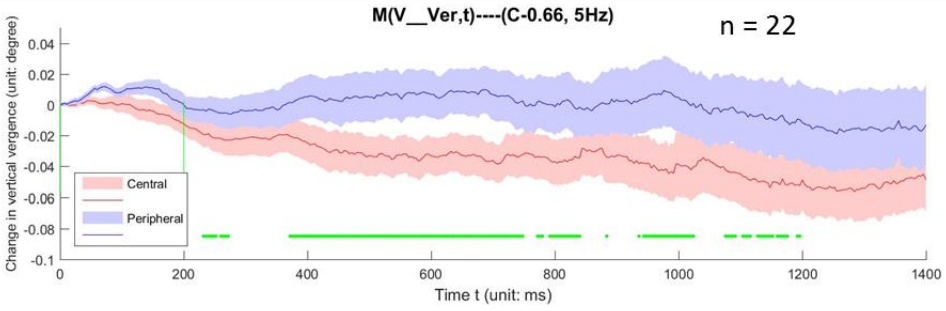


(C) C_{-} = 0.3, w = 5 Hz

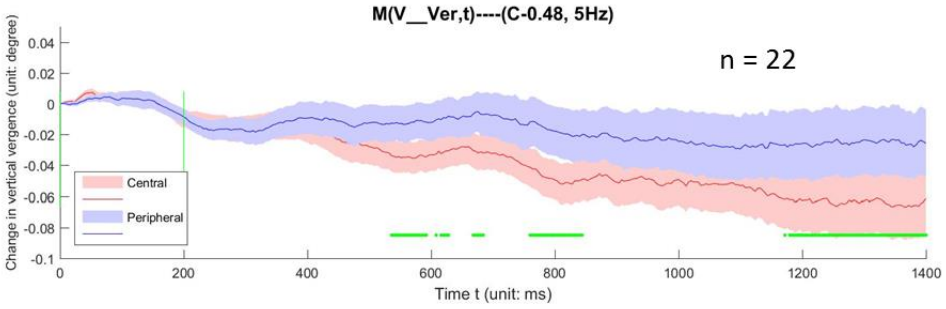


(D) C_{-} = 0.12, w = 5 Hz

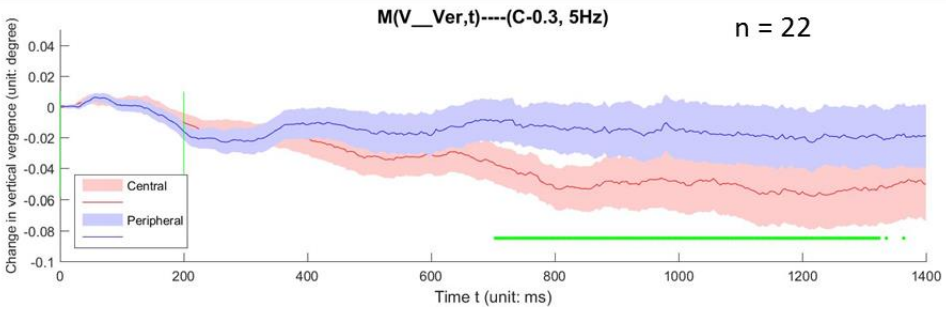
Fig. 3.12: The change in horizontal vergence at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in horizontal vergence (unit: degree). The figures (A), (B), (C) and (D) represent at C- = 0.66, 0.48, 0.3, and 0.12 conditions respectively. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test (p < 0.01). The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means.



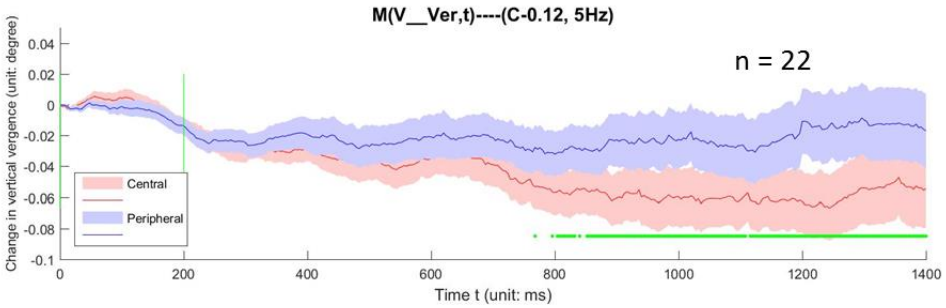
(A) C_{-} = 0.66, w = 5 Hz



(B) C_{-} = 0.48, w = 5 Hz

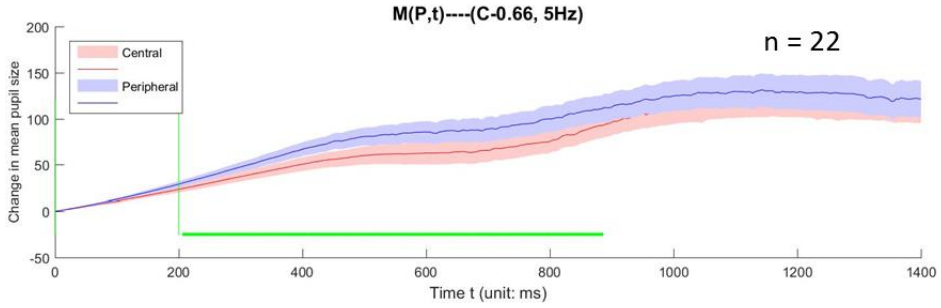


(C) C_{-} = 0.3, w = 5 Hz

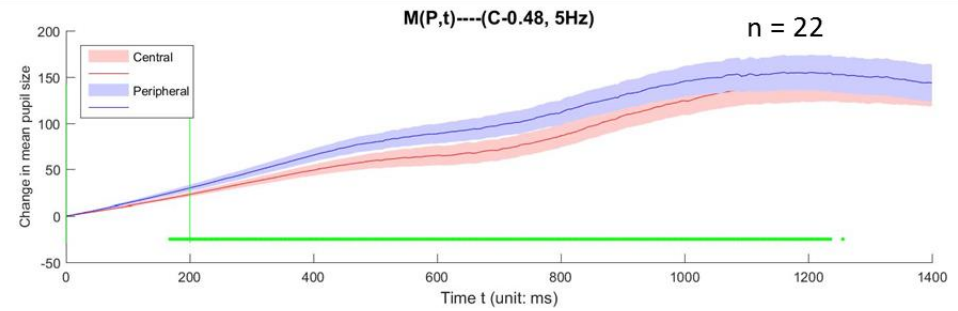


(D) C_{-} = 0.12, w = 5 Hz

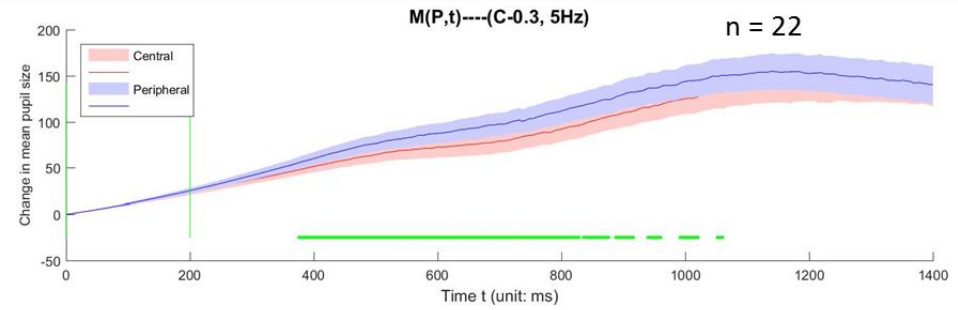
Fig. 3.13: The change in vertical vergence at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in vertical vergence (unit: degree). The figures (A), (B), (C) and (D) represent at C- = 0.66, 0.48, 0.3, and 0.12 conditions respectively. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$). The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means.



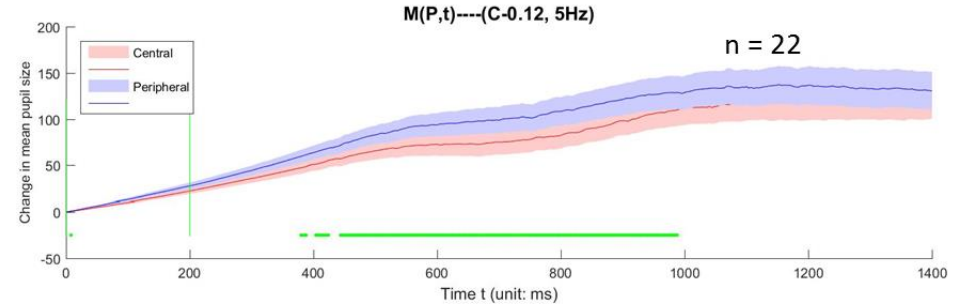
(A) C_{-} = 0.66, w = 5 Hz



(B) $C_{-} = 0.48$, w = 5 Hz



(C) $C_{-} = 0.30$, w = 5 Hz



(D) $C_{-} = 0.12$, w = 5 Hz

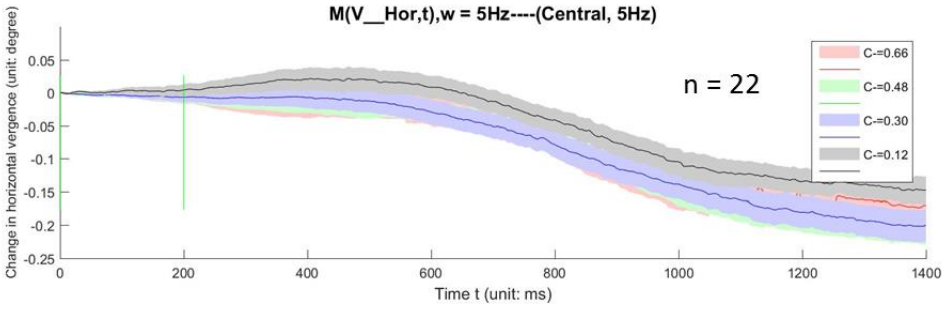
Fig. 3.14: The change in pupil size at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in pupil size. The figures (A), (B), (C) and (D) represent at $C_{-} = 0.66$, 0.48, 0.3, and 0.12 conditions respectively. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$). The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means.

From Figures 3.12, 3.13, and 3.14, there are significant differences of change in horizontal vergence between the central and peripheral view conditions at all relative contrast conditions except $C_{-} = 0.66$. There are significant differences of change in vertical vergence (and also change in pupil size) between central and peripheral conditions at all relative contrast conditions. These results indicate the top-down feedback is involved in the central view

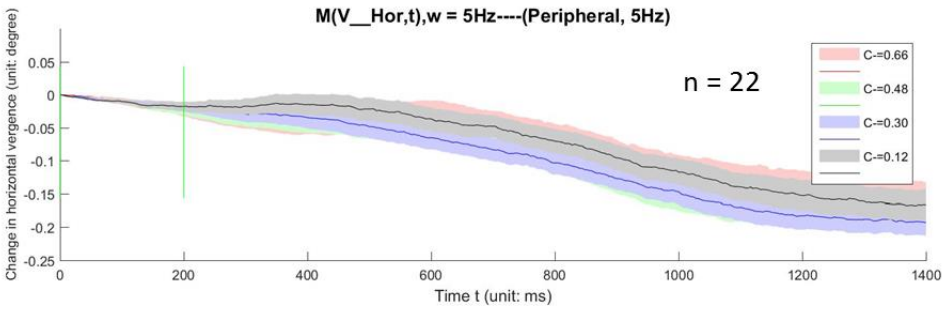
condition as explained in previous section.

However, at C_- = 0.66, we cannot see a significant difference of change in horizontal vergence between central and peripheral conditions. This might be because the ratio of C_+ is relatively low if calculated with relative contrast $C_+/(C_+ + C_-)$. Hence, the percept of binocular suppression S_- might be dominant. The top-down feedback to the limited binocular summation S_+ percept does not cause obvious vergence change at central condition.

Figure 3.15 (also Figure 3.16 and Figure 3.17) separates the central and peripheral conditions as shown in figure (A) and (B) respectively, and compares the change in horizontal vergence (also change in vertical vergence and change in pupil size) at different relative contrasts. The horizontal and vertical axes have the same meaning as in Figure 3.12 (Figure 3.13 and Figure 3.14). The pink, green, blue and gray curves represent C_- = 0.66, 0.48, 0.3, and 0.12 respectively. The shade of the curves means the standard errors of the means.



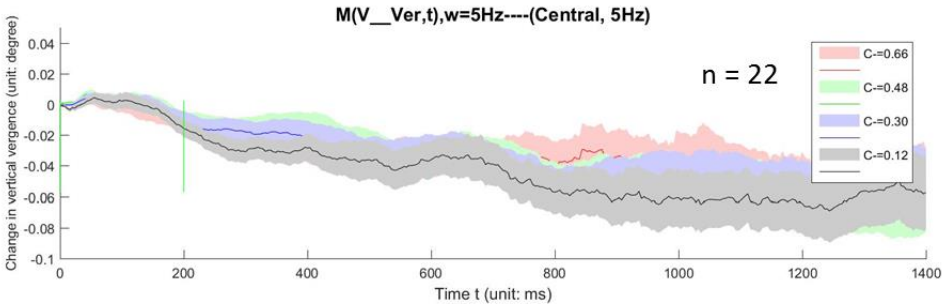
(A) The central view condition, w = 5Hz



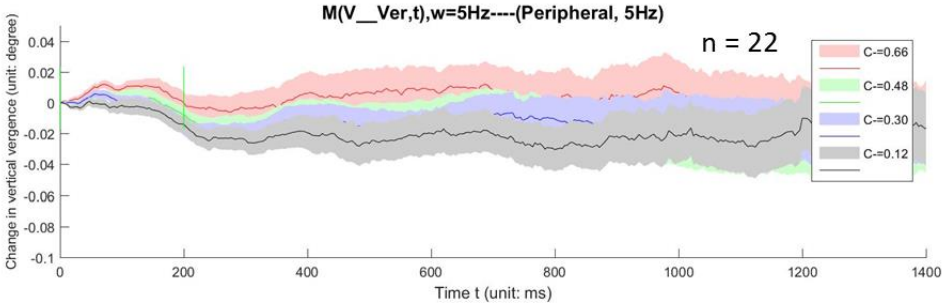
(B) The peripheral view condition, w = 5Hz

Fig. 3.15: Comparison of the change in horizontal vergence at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in horizontal vergence (unit: degree) with w = 5 Hz. The figures (A) and (B) represent the central and peripheral conditions respectively. The pink, green, blue and gray curves represent the C_- = 0.66, 0.3, 0.48, 0.12 respectively. The shade of the

curves meant the standard errors of the means.

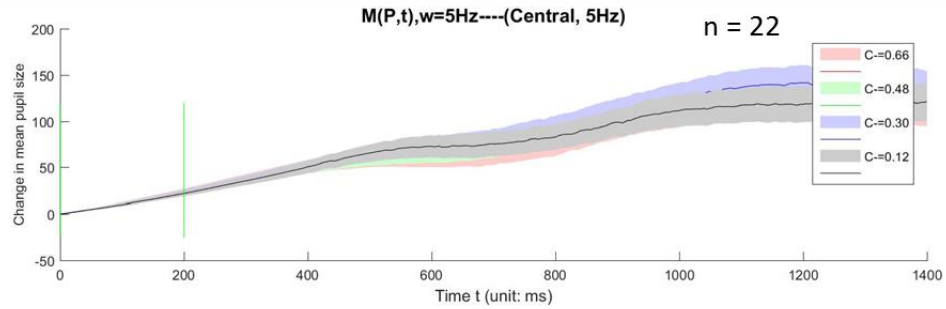


(A) Central condition, w = 5Hz

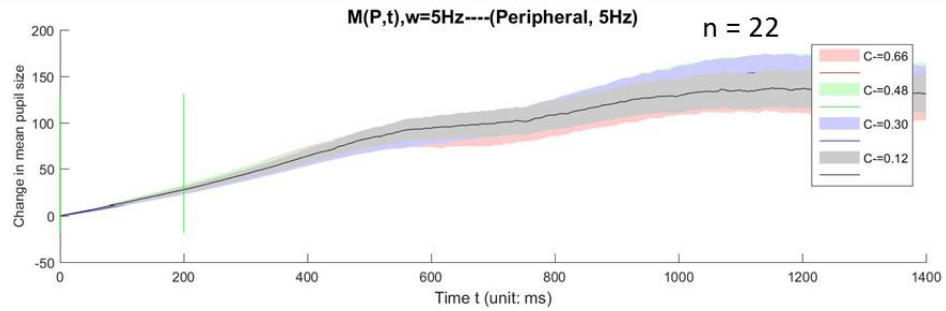


(B) Peripheral condition, w = 5Hz

Fig. 3.16: Comparison of the change in vertical vergence at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in vertical vergence (unit: degree) with w = 5 Hz. The figures (A) and (B) represent the central and peripheral conditions respectively. The pink, green, blue and gray curves represent the C- = 0.66, 0.3, 0.48, 0.12 respectively. The shade of the curves meant the standard errors of the means.



(A) Central condition, w = 5Hz



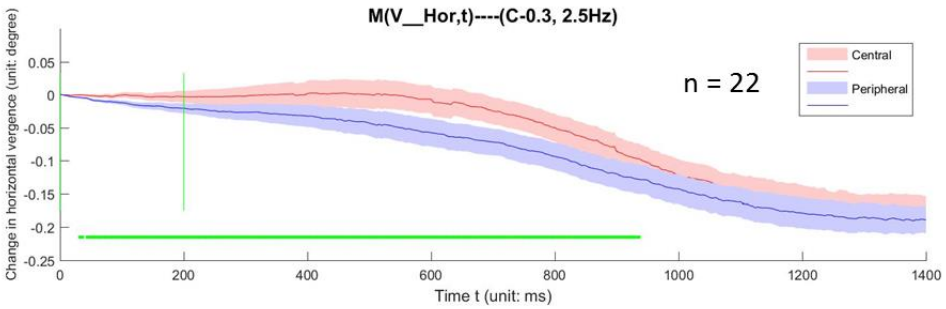
(B) Peripheral condition, w = 5Hz

Fig. 3.17: Comparison of the change in pupil size at different relative contrasts. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in pupil size with w = 5 Hz. The figures (A) and (B) represent the central and peripheral conditions respectively. The pink, green, blue and gray curves represent the C_- = 0.66, 0.3, 0.48, 0.12 respectively. The shade of the curves meant the standard errors of the means.

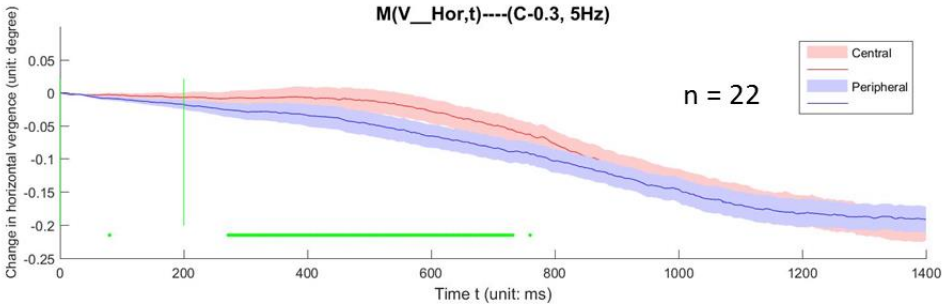
From the results, when comparing the change in horizontal vergence (change in vertical vergence and change in pupil size) at different relative contrasts, there is no obvious difference or clear tendency within the central condition and also within peripheral condition. However, from the behavioral data as shown in Figure 3.8, the fraction (F_+) of S_+ percept at both central and peripheral conditions are the sigmoidal distributions against relative contrasts. The larger relative contrast causes the larger fraction (F_+) of S_+ percept, while does not cause the obvious larger changes in horizontal vergence, vertical vergence or pupil size. One of the possible reason is the bottom-up process is more involved with the increased relative contrast.

3.2.2.4 Vergence eye movement at different temporal frequencies

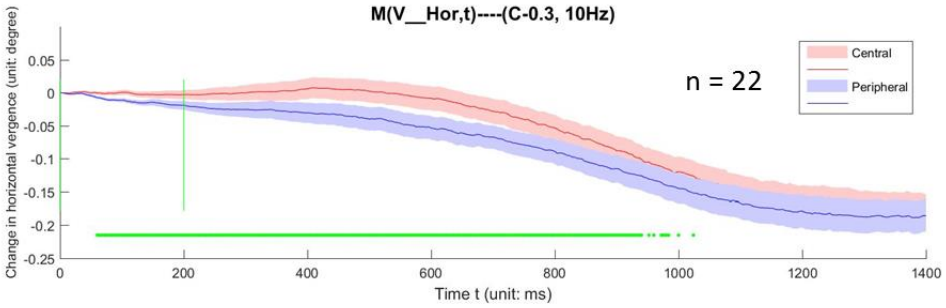
Figure 3.18, Figure 3.19 and Figure 3.20 show the change in horizontal vergence, vertical vergence and pupil size respectively at different temporal frequencies with C_- = 0.3. In all the figures, the horizontal axes are the time duration (unit: ms); the vertical axes are the change in horizontal vergence (unit: degree), change in vertical vergence (unit: degree) and change in pupil size respectively. The figures (A), (B) and (C) represent w = 2.5, 5 and 10 Hz temporal frequencies respectively. The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means. The horizontal green bar means the time duration which has significant difference between the central and peripheral conditions with t-test ($p < 0.01$).



(A) C- = 0.3, temporal frequency w = 2.5 Hz

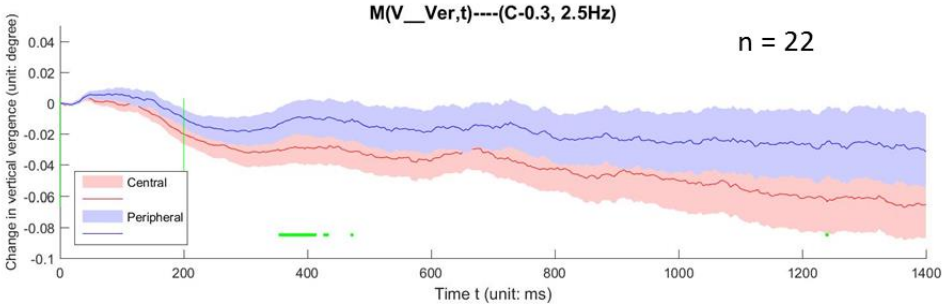


(B) C- = 0.3, temporal frequency w = 5 Hz

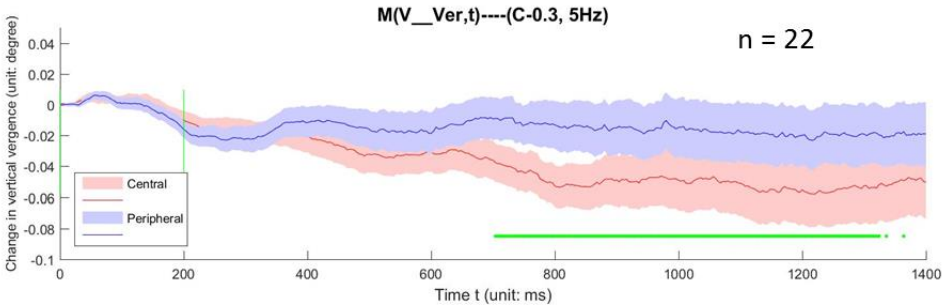


(C) C- = 0.3, temporal frequency w= 10 Hz

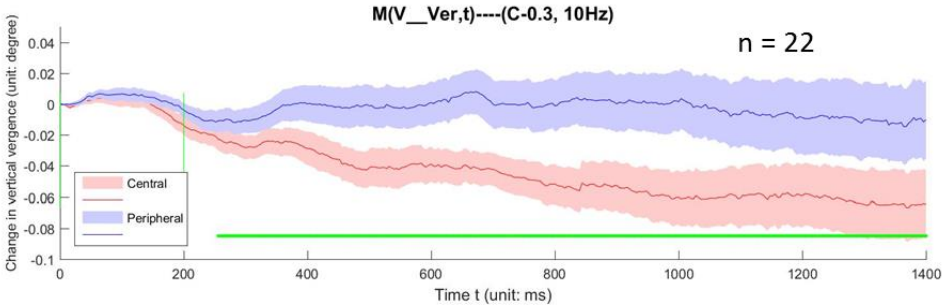
Fig. 3.18: The change in horizontal vergence against time at different temporal frequencies. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in horizontal vergence (unit: degree). The figures (A), (B) and (C) represent at 2.5, 5 and 10 Hz temporal frequencies respectively. The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$).



(A) $C_{-} = 0.3$, temporal frequency = 2.5 Hz

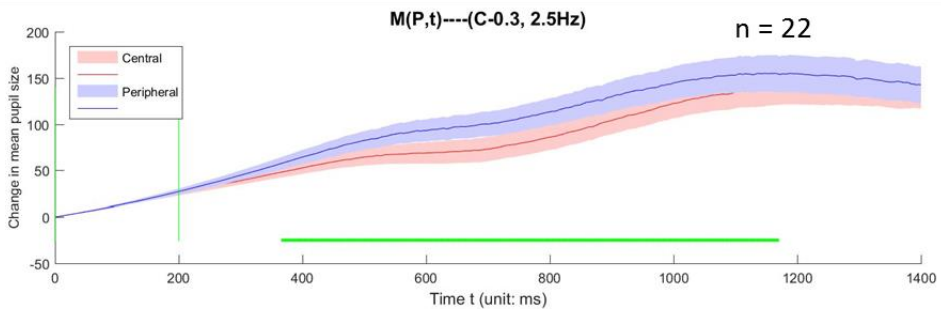


(B) $C_{-} = 0.3$, temporal frequency = 5 Hz

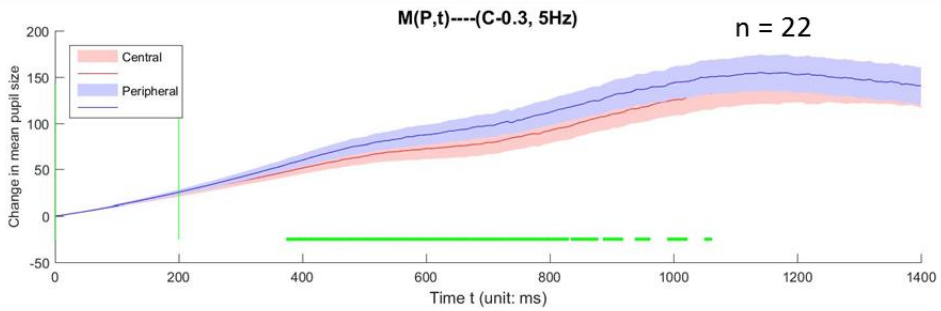


(C) $C_{-} = 0.3$, temporal frequency = 10 Hz

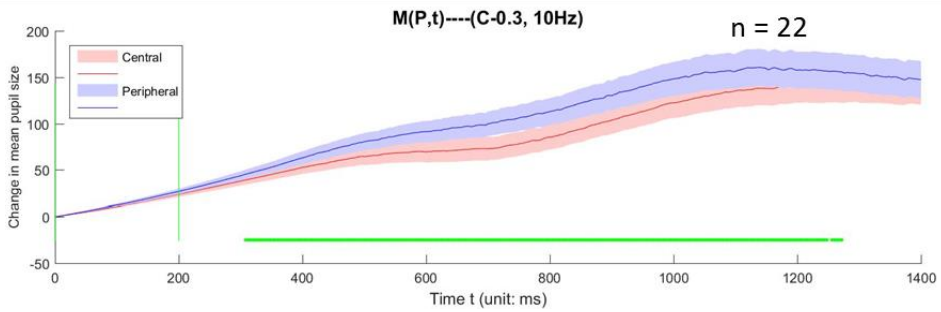
Fig. 3.19: The change in vertical vergence against time at different temporal frequencies. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in vertical vergence (unit: degree). The figures (A), (B) and (C) represent at 2.5, 5 and 10 Hz temporal frequencies respectively. The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$).



(A) $C_{-} = 0.3$, temporal frequency w = 2.5 Hz



(B) C- = 0.3, temporal frequency w = 5 Hz



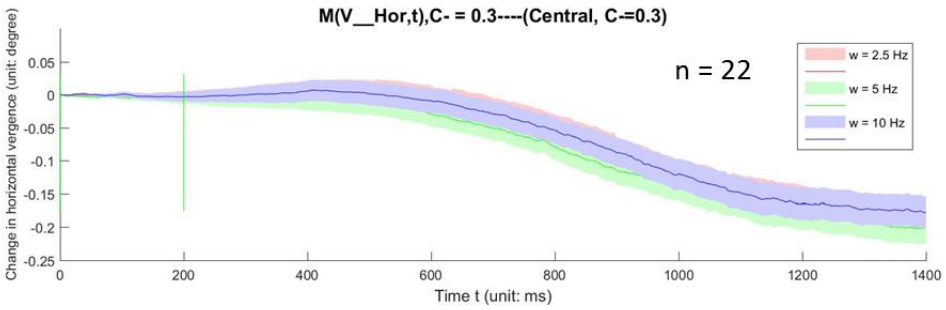
(C) C- = 0.3, temporal frequency w = 10 Hz

Fig. 3.20: The change in pupil size against time at different temporal frequencies. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in pupil size. The figures (A), (B) and (C) represent at 2.5, 5 and 10 Hz temporal frequencies respectively. The pink and blue curves mean the central and peripheral conditions respectively. The shade of the curve means the standard errors of the means. The horizontal green bars mean the time duration which have significant difference between the central and peripheral conditions with t-test ($p < 0.01$).

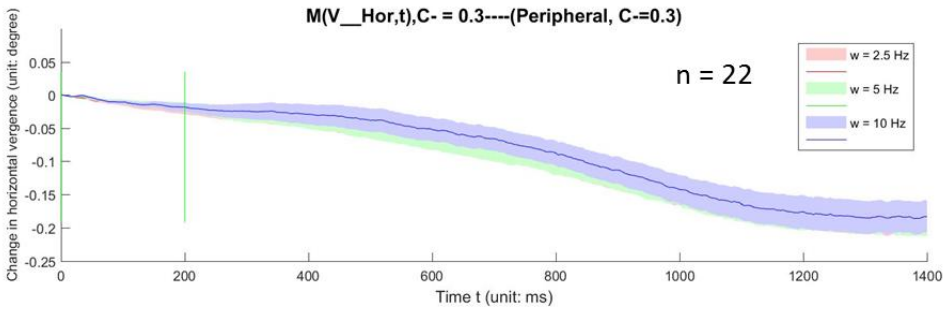
From Figure 3.18 (Figure 3.19 and Figure 3.20), there are significant differences of change in horizontal vergence (also change in vertical vergence and change in pupil size) between central and peripheral conditions at all temporal frequency conditions. These results also indicate the top-down feedback is involved in the central view condition as explained in previous section. However, the change in vertical vergence at w = 2.5 Hz is an exception.

Figure 3.21 (Figure 3.22 and Figure 3.23) separate the central and peripheral conditions as shown in figure (A) and (B) respectively, and compare the change in horizontal vergence (also change in vertical vergence and change in pupil size) at different temporal frequencies. The horizontal and vertical axes have the same meanings as in Figure 3.18 (Figure 3.19 and Figure

3.20). The pink, green and blue curves represent the w = 2.5, 5 and 10 Hz respectively. The shade of the curves mean the standard errors of the means.

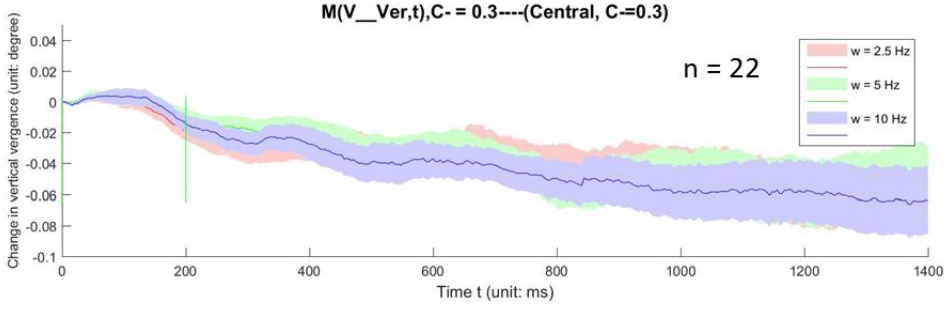


(A) Central view condition

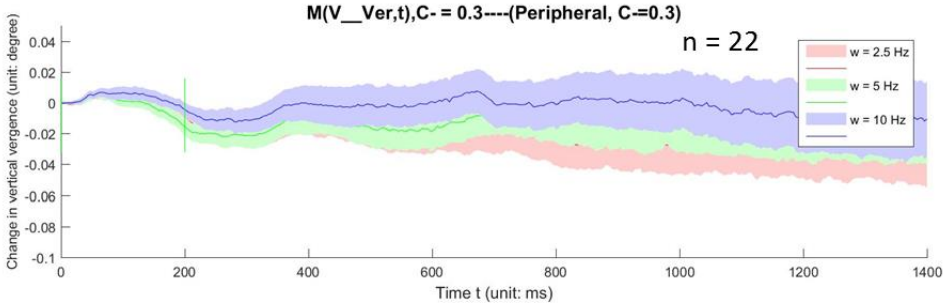


(B) Peripheral view condition

Fig. 3.21: Comparison of the change in horizontal vergence at different temporal frequencies. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in horizontal vergence (unit: degree) with C. = 0.3. The figures (A), and (B) represent the central and peripheral conditions respectively. The pink, green and blue curves represented the w = 2.5, 5, and 10 Hz respectively. The shade of the curves meant the standard errors of the means.

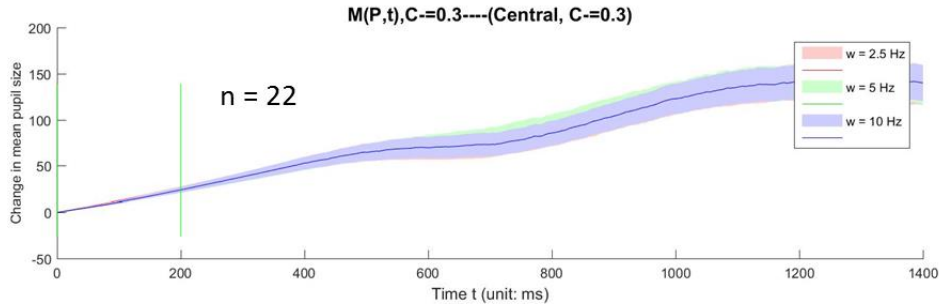


(A) Central view condition

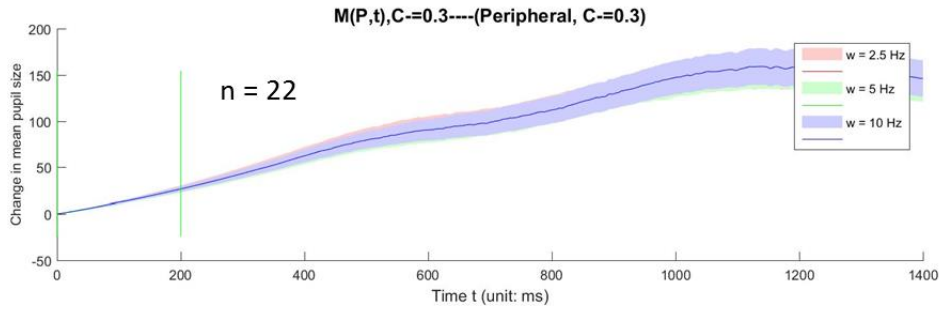


(B) Peripheral view condition

Fig. 3.22: Comparison of the change in vertical vergence at different temporal frequencies. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in vertical vergence (unit: degree) with C- = 0.3. The figures (A), and (B) represent the central and peripheral conditions respectively. The pink, green and blue curves represented the w = 2.5, 5, and 10 Hz respectively. The shade of the curves meant the standard errors of the means.



(A) Central view condition



(B) Peripheral view condition

Fig. 3.23: Comparison of the change in pupil size at different temporal frequencies. In all figures, the horizontal axes are the time duration with unit in ms; the vertical axes are the change in pupil size (unit: degree) with C- = 0.3. The figures (A), and (B) represent the central and peripheral conditions respectively. The pink, green and blue curves represented the w = 2.5, 5, and 10 Hz respectively. The shade of the curves meant the standard errors of the means.

From the results, when comparing the change in horizontal vergence at different temporal frequencies, there is neither significant difference at the central condition, nor at the peripheral

condition. Similarly, when comparing the change in pupil size at different temporal frequencies, no significant difference can be found at both view conditions. There are some shifts of change in vertical vergence, however, there is no significant difference.

From the behavioral data as shown in Figure 3.9, the larger temporal frequency (at w = 10 Hz) causes the lower fraction (F_+) of seeing the S_+ drift direction. As explained, the visual decoding is involved in the central vision. When the stimuli are in high temporal frequency and short time duration, the limited brain resources have bottleneck to decode all the visual inputs, although the top-down feedback is involved.

3.3 Discussion

3.3.1 Vergence eye movements is involved

From Figure 3.10 (A) and (C), the change in pupil size matches the change in horizontal vergence distribution. The significant difference of pupil size has around 200 ms latency than the vergence change, suggesting the vergence change might cause the pupil size change.

From Figure 3.10 (B), there are significant differences of the change in vertical vergence at central condition at t > 700 ms. The change in vertical vergence at central condition has negative value, meaning the sursumvergence during the visual perception. In previous study, Rambold et al. (2010) used two 1-D horizontal sinusoidal gratings with ¼ wavelength difference in phase as dichoptic stimuli (one grating for each eye). The stimuli were produced by modulating the contrast of a high-frequency carrier. They added various amounts of luminance for both eyes, and found the change in vertical vergence (sursumvergence) was dependent on the added luminance, all with latencies less than 150 ms. The similarity between their study and this study, is that we both found the change in vertical vergence (sursumvergence) at central condition with $C_- = 0.5$, w = 5 Hz, since both experiments used the horizontal gratings. The differences are in their study, they just presented the contrast and

luminance modulated gratings and recorded the eye information. Participant did not have task. However, in this study the special designed ambiguous motion gratings were presented for each eye to investigate the perception difference between the central and peripheral conditions. Participant had to judge the motion direction and give their responses. The different experimental design involved different brain mechanisms. In their study, the latency of vertical vergence change was less than 150 ms, whereas in this study, the significant difference shows at least 250 ms after the dichoptic stimuli onset. Based on the previous studies (Theeuwes, 2010; Wolfeetal, 2013), the long latency (> 150 ms) might be caused by the top-down feedback.

Moreover, in this experiment, the eye tracking data and the behavioral data were collected simultaneously. As shown in Figure 3.8, behavioral data showed significant difference of fraction (F_+) of seeing S_+ drift direction between the central and peripheral conditions. Based on Zhaoping (2017), The FFVW model can be used to explain the results. When V1 neurons receive the visual inputs from retina, both the S_+ and S_- channels will produce feedforward signals to the higher brain areas. Then the higher brain areas give feedback mainly to the S_+ channel based on analysis-by-synthesis computation and the prior knowledge that the binocular inputs are correlated. However, this feedback to the S_- channel is weaker or absent (Zhaoping, 2017). Subsequently, the S_+ channel verifies the feedback and may guide the vergence eye movement via the mediation of the superior colliculus (Zhaoping, 2014; Zhaoping, 2017). The circuit of visual processing with S_+ and S_- feedforward and feedback is shown as Figure 3.24. The long latency might be because the multicycles of FFVW were involved (Zhaoping, 2017).

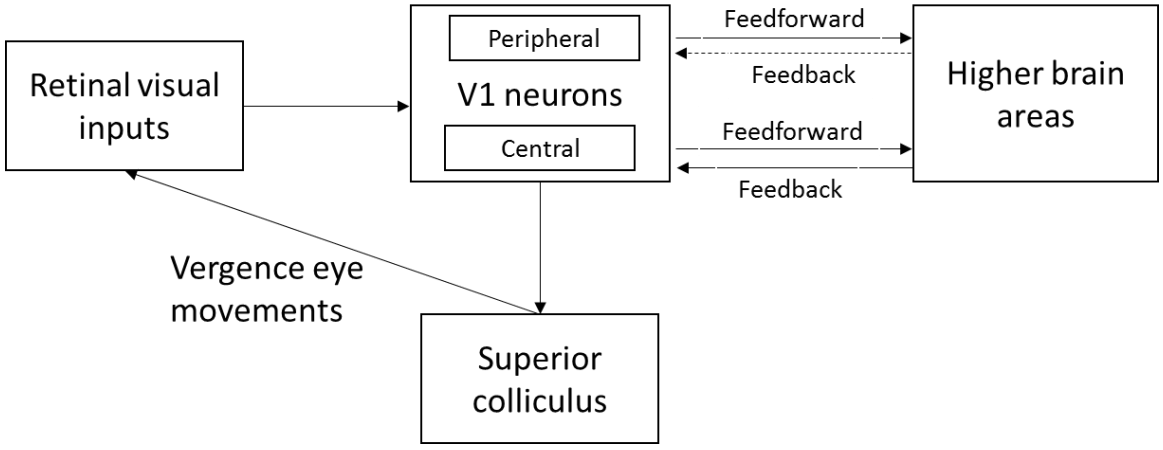


Fig. 3.24: The circuit of visual processing

Moreover, based on the functional differences of the central and peripheral vision, the former mainly takes responsibility for the visual decoding, thus follows the FFVW processes; whereas the latter mainly focuses on the visual selection, and involves the retina-to-V1 process.

3.3.2 The influence of relative contrast and temporal frequency on the vergence eye movements

From the eye tracking data, at different relative contrasts and different temporal frequencies (Figures 3.12 - 3.14 and Figures 3.18 - 3.20), there are significant differences of change in horizontal, vertical and pupil size between central and peripheral conditions, indicating the top-down feedback is involved in the central view condition.

However, at $C_- = 0.66$ w = 5 Hz condition, there is no significant difference of change in horizontal vergence between central and peripheral conditions. This might be because the ratio of C_+ is relatively low if calculated with relative contrast $C_+/(C_+ + C_-)$. Consequently, the top-down feedback to the limited binocular summation S_+ percept does not cause obvious vergence change at central condition, because the binocular suppression S_- percept might be dominant at this condition.

From the relative contrast aspect, when comparing the behavioral results with the eye tracking data, the larger relative contrast causes the larger fraction (F_+) of S_+ percept, while does not cause the obvious larger changes in horizontal vergence, vertical vergence or pupil size. One of the possible reasons is the bottom-up process is more involved along with the increasing of relative contrast $C_+/(C_+ + C_-)$, since the contrast is a relatively early level feature.

From the temporal frequency aspect, when comparing the behavioral results with the eye tracking data, the larger temporal frequency (at w = 10 Hz) causes the lower fraction (F_+) of seeing the S_+ drift direction at the central view condition. Since the visual decoding was involved in the central vision, when the stimuli were presented in high temporal frequency and short time duration, the limited brain resources had bottleneck to decode all the visual inputs. However, there is no obvious change of vergence eye movements across different temporal frequencies, suggesting the top-down feedback is involved in all conditions.

This experiment used the vergence eye movements as a tool to prove the circuit of visual processing with feedforward and feedback for the S_+ and/or S_- percepts, and also the mediation of superior colliculus. The results suggested the importance of V1 in the central and peripheral visual processes; and also the involvement of top-down feedback at the central vision.

Chapter 4 Experiment 2: temporal dynamics for visual recognition involves vergence eye movements and hierarchical processes

When the left and right eyes gaze on a small target, the corresponding images from the two retinas will be combined as an object. However, when the two eyes' inputs are uncorrelated, these dichoptic inputs may cause a perceptual rivalry and work in a "winner-take-all" manner, meaning only one monocular input is the dominant precept each time. Moreover, the dichoptic inputs may also cause the perceptual flip if they are perceived over time, which might be indicated from the blink and (micro)saccade information. For examples, watching a red circular grating and a green radial grating from the left and right eyes accordingly, the percepts altered between the two gratings over time, which were local retinotopic level processes (Chen and He, 2003). Watching a slant surface defined by monocular percept-specified cue or binocular disparity-specified cue separately, the percepts altered between the floor and ceiling slants at the both cue conditions (van Dam and van Ee, 2006a; van Dam and van Ee, 2006b). During the perceptual process, there were reduced probabilities of blink and (micro)saccade at the moment of percept flip; whereas there seemed no relationship between the change in horizontal vergence and the percept flip (van Dam and van Ee, 2006a). Besides, there was also report that the retinal image changes caused by saccade altered the percept (van Dam and van Ee, 2006b).

Zhaoping (2017) increased the stimulus presentation duration from 0.05 s to 0.2 s at the tilt stimulus condition, and found the bias of S_+ percept at central condition increased. Experiment 1 investigated the vergence eye movement with the dichoptic stimuli onset 0.2 s, and found the vergence changes. This experiment aimed to investigate the temporal dynamics

of S_+ and S_- percepts. As in the conventional binocular rivalry, the percepts alter between the left and right eye inputs; whether there is rivalry between the S_+ and S_- percepts over time? To this end, in contrast to the 0.2 s stimulus presentation time duration in previous study (Zhaoping, 2017), we presented the dichoptic stimuli in 300 s, and recorded the real-time behavior response and binocular eye information simultaneously.

4.1 Method

4.1.1 Participants

Totally fifteen people (10 males and 5 females, age from 18 to 34 years, mean age: 23.6 years) from Kochi University of Technology were recruited as participants. All of them had normal or corrected-to-normal vision, and were tested for their stereo acuity and motion acuity with our customized programs. All participants were naïve to the aim of the experiment and were compensated for their time. The authors did not serve as participants. All experiments and procedures were approved by the Research Ethics Committee of Kochi University of Technology and conformed to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to experiments.

4.1.2 Stimuli and apparatus

The dichoptic stimuli were the same as described in experiment 1, except the presentation time was changed from 0.2 s to 300 s. The parameters were $C_+ = C_- = 0.3$, temporal frequency w = 5 Hz. The experimental setup was totally the same as in experiment 1.

4.1.3 Procedure

The experimental procedure was shown as in Figure 4.1.

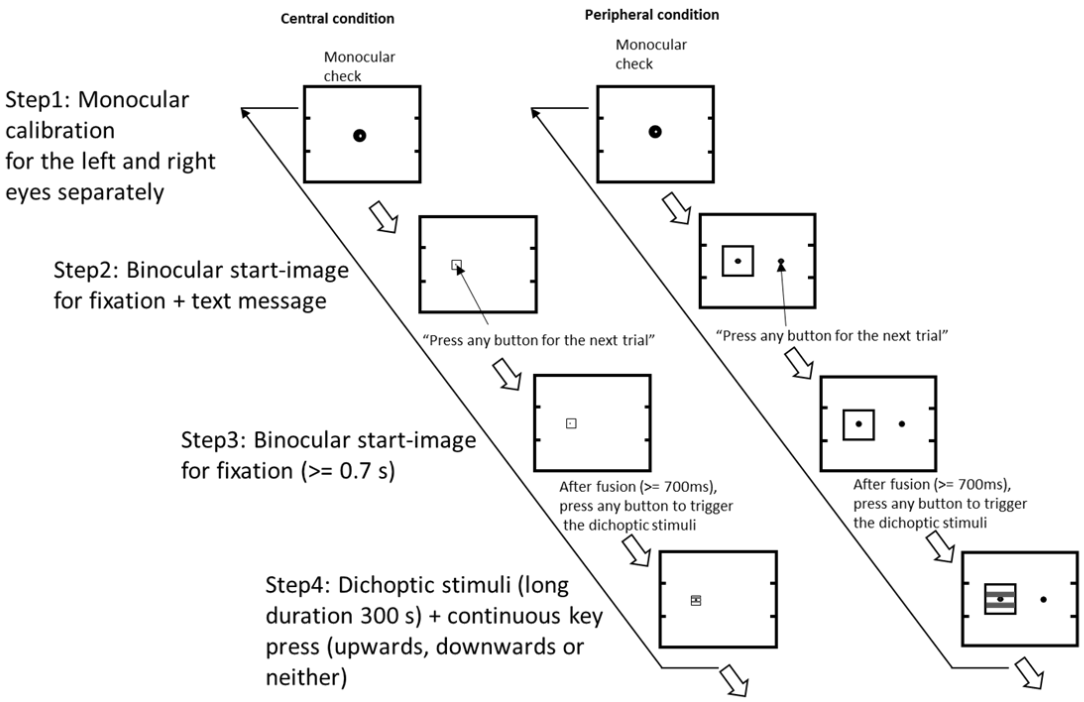


Fig. 4.1: Procedure of experiment 2. The left and right images show the procedures at the central and peripheral view conditions respectively. For both conditions, the procedures contain: the monocular calibration for the left and right eyes separately at beginning of each trial (step 1); binocular start-image for fixation and the text message "press any button for the next trial" to instruct participants to trigger for the next trial (step 2); binocular start-image for fixation, after fusion (>= 700 ms) participants were instructed to press a button to trigger the dichoptic stimuli (step 3); presentation of dichoptic stimuli (long duration for 300 s), participants were asked to judge the motion direction (upwards, downwards, or neither) by continuous key press (step4).

At the beginning of each block, the monocular calibration was implemented for both the left and right eyes separately (step 1). After calibration, the binocular start-image for fixation and the text message "press any button for the next trial" were presented to instruct participants to trigger for the next trial (step 2). The text message was located under the fixation point. Then the binocular start-image for fixation was shown, and participants were instructed to gaze at the fixation point for at least 700 ms (step 3). Meanwhile, the Eyelink II started to record the binocular eye positions and pupil sizes. Only when participants confirmed they focused well on the fixation, they were instructed to press a button to trigger for the dichoptic stimuli. Subsequently, the dichoptic stimuli were presented for 300 seconds (step 4). Participants were asked to judge the motion direction (as upwards, downwards or neither) and give their responses by real-time continuous key press with the specified buttons,

for example continuous upArrow press for upwards percept, continuous downArrow press for the downwards percept, and no button press if the perceived motion direction was neither up nor down, or both up and down. Once the perceived motion direction changed, they were asked to change the key press as timely as possible.

During each trial, participants were asked to keep their heads as steady as possible, and keep the eyes on the fixation. After each trial, participants were asked to take their heads off the chin-rest, remove the Eyelink II and have a short break. This was done to avoid the fatigue to the eyes and necks caused by the high concentration during experiment or the headache caused by the weight of the head-mounted device. After break, participants were asked to redo the calibration before the next trial.

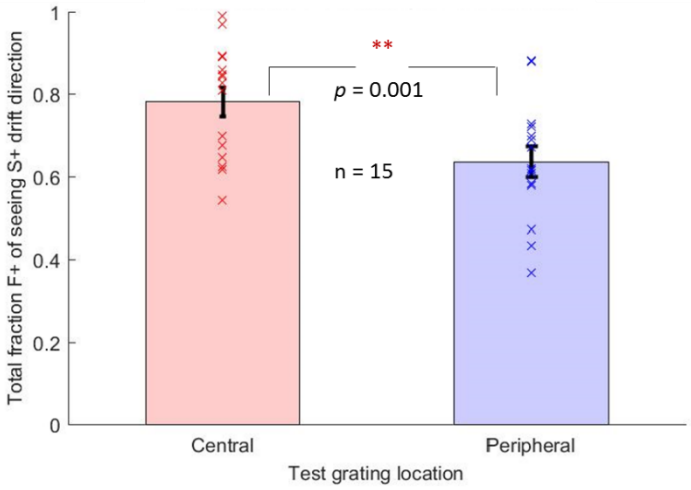
The experiment was divided into 2 sessions and implemented in different days for each participant. In each session, there were 8 trials which contained two view conditions (central and peripheral). Both the view conditions and motion directions (upwards and downwards) were generated in a counterbalanced random order. In total, 16-trial data were collected for each participant.

4.2 Results

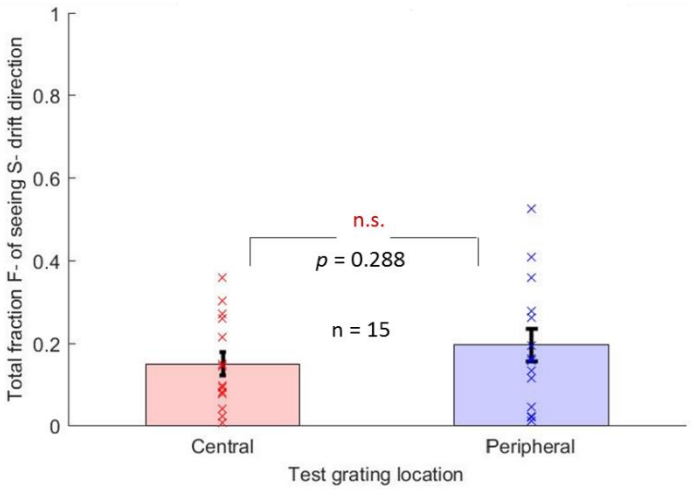
4.2.1 Behavioral data

The total fractions (F) of seeing the three percepts (S_+, S and neither) at both the central and peripheral view conditions were calculated by using 15 participants' data. Figure 4.2 (A), (B) and (C) show the results of F_+, F_- and F_N respectively. The red and blue bars represent the central and peripheral view conditions respectively. The error bars are the standard errors of the means. Matched sample t-test was used to analyze the significance between the two view conditions. From the results, when comparing the central and peripheral view conditions, the

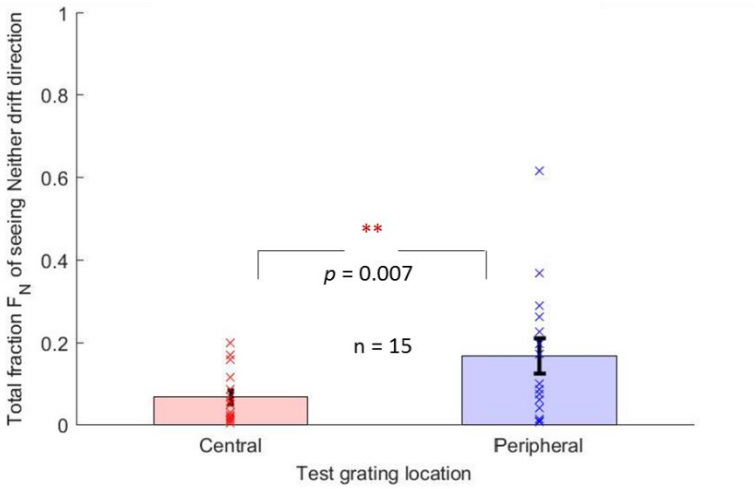
fractions of seeing the S_+, S_- and neither motion directions are 78% vs 63%, 15% vs 20%, and 7% vs 17% respectively. There are significant differences between the central and peripheral view conditions at the S_+ percept condition (p = 0.001, $t(14)$ = 4.06) and at the neither percept condition (p = 0.007, $t(14)$ = 3.18), while no significant difference at the S_- percept condition (p = 0.288, $t(14)$ = 1.12).



(A) Total fraction F_+ of seeing S_+ drift direction



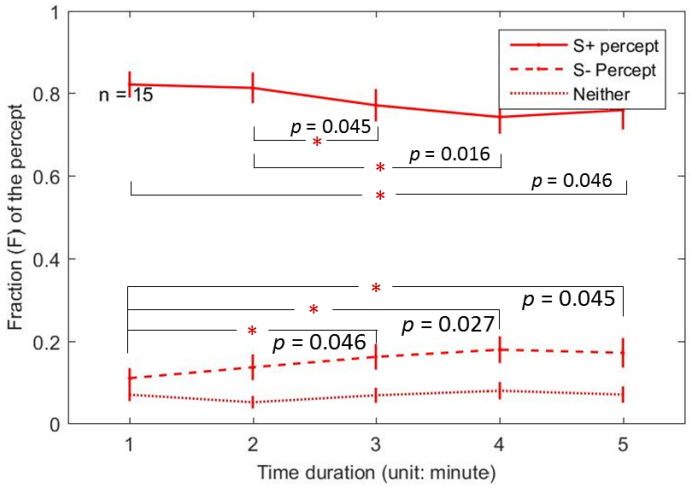
(B) Total fraction F_- of seeing S_- drift direction



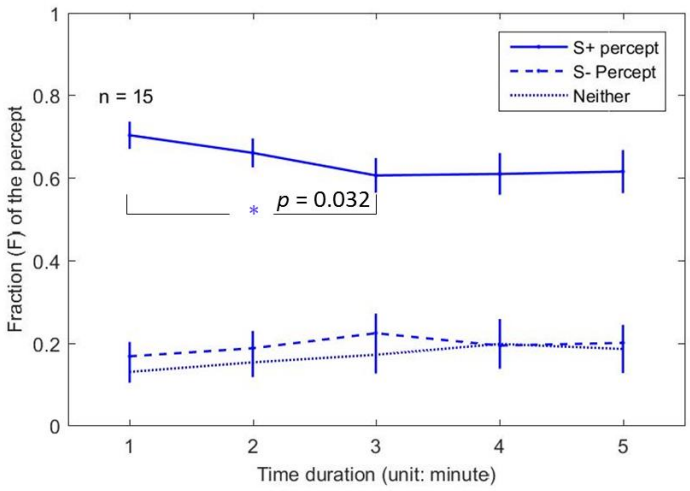
(C) Total fraction F_N of seeing neither drift direction

Fig. 4.2: Total fraction (F) of the three percepts at the central and peripheral view conditions. (A) total fraction F_+ of seeing S_+ drift direction; (B) total fraction F_- of seeing S_- drift direction; (C) total Fraction F_N of seeing neither drift direction. The error bars are the standard errors of the means. The red asterisk means there is a significant difference, with the marker as following meaning: $+p < .10$, $*p < .05$, $**p < .01$, $***p < .001$, n.s. means no significant different. The *p* value is calculated by matched sample t-test.

Then the five-minute time duration were divided into five sections, and marked as td1- td5 respectively. Subsequently, the F_+, F_- and F_N at each section were calculated at both the central and peripheral view conditions. The results were shown in Figure 4.3 (A) and (B). The red and blue lines represent the central and peripheral conditions respectively. The line, dash line and dot line represent the S_+, S_- and neither percepts respectively. The error bars are the standard errors of the means.



(A) Temporal distribution of F at the central view condition



(B) Temporal distribution of F at the peripheral view condition

Fig. 4.3: Temporal distribution of the F_{+}, F_{-} and F_{N} at the central and peripheral view conditions. (A) Temporal distributions of F at the central view condition; (B) Temporal distributions of F at the peripheral view condition. The red and blue lines represent the central and peripheral view conditions respectively. The line, dash line and dot line represent the S_{+}, S_{-} and neither percepts respectively. The error bars are the standard errors of the means. The red and blue asterisks mean there is a significant difference (multi comparison), with the marker as following meaning: +p < .10, *p < .05, **p < .01, ***p < .001.

At the central view condition (Figure 4.3 (A)), 3 (S_{+}, S_{-}, and neither) × 5 (td1 – td5) ANOVA reveals significant main effect of the percept conditions ($F(2, 28) = 133.64$, $p < 0.001$), and significant interaction between the temporal distributions and percept conditions ($F(8, 112) = 2.83$, $p = 0.007$). There are significant differences of three percepts at td1 ($F(2, 28) = 194.99$, $p < 0.001$), at td2 ($F(2, 28) = 133.90$, $p < 0.001$), at td3 ($F(2, 28) = 102.20$, $p < 0.001$), at td4 ($F(2, 28) = 81.21$, $p < 0.001$), and at td5 ($F(2, 28) = 70.85$, $p < 0.001$). Multiple comparison for temporal distribution at S_{+} percept shows significant differences between the td2 and td4 ($t(14) = 2.74$, $p = 0.016$), between the td2 and td3 ($t(14) = 2.20$, $p = 0.044$), and between the td1 and td4 ($t(14) = 2.19$, $p = 0.046$); while no significant difference between other pairs. Multiple comparison for temporal distribution at S_{-} percept shows significant differences between the td1 and td4 ($t(14) = 2.46$, $p = 0.027$), between the td1 and td5 ($t(14) = 2.20$, $p = 0.045$), and between the td1 and td3 ($t(14) = 2.17$, $p = 0.046$); while no significant

difference between other pairs. Multiple comparison for temporal distribution at neither percept shows no significant difference between any pairs.

At the peripheral view condition (Figure 4.3 (B)), 3 (S_+, S_-, and neither) × 5 (td1 – td5) ANOVA reveals significant main effect of the percept conditions ($F(2, 28) = 30.03$, $p < 0.001$), and significant interaction between the temporal distributions and percept conditions ($F(8, 112) = 2.03$, $p = 0.0495$). There are significant differences of three percepts at td1 ($F(2, 28) = 68.68$, $p < 0.001$), at td2 ($F(2, 28) = 37.23$, $p < 0.001$), at td3 ($F(2, 28) = 18.34$, $p < 0.001$), at td4 ($F(2, 28) = 14.24$, $p < 0.001$), and at td5 ($F(2, 28) = 15.48$, $p < 0.001$). There are significant differences of temporal distributions at the S_+ percept ($F(4, 56) = 2.67$, $p = 0.041$); whereas neither significant differences at the S_- percept ($F(4, 56) = 1.32$, $p = 0.27$), nor at the neither percept ($F(4, 56) = 1.56$, $p = 0.20$). Multiple comparison for temporal distribution at the S_+ percept shows a significant difference between the td1 and td3 ($t(14) = 2.38$, $p = 0.032$), while neither significant difference between other pairs at S_+ percept, nor between any pairs at the S_- percept or at the neither percept.

When seeing the temporal distributions of the S_+, S_- and neither percepts, at the central view condition, there is a gradually decreasing trend of the F_+ with significant differences between different time sections. Concomitant with the decreasing of the F_+ is the increasing of the F_- accordingly, which might indicate a competition between the S_+ and S_- percepts; whereas there is almost no change of the F_N. However, at the peripheral view condition, there is a decreasing trend of the F_+ at the first three minutes. Concomitant with the decreasing of the F_+ are the certain amounts of increasing of both the F_N and F_-. But there is no significant difference among different time sections of both the F_N and F_-, meaning there seems no tendency of competition between the S_+ and S_- percepts.

Figure 4.4 shows the temporal distribution of the mean quantities of the percept flips from 15 participant's data. The red and blue lines represent the central and peripheral view

conditions respectively. The error bars are the standard errors of the means. At the central view condition, the mean quantities of the percept flips are increasing during the first four minutes. However, at the peripheral view condition, the mean percept flip quantities are stable during the whole time duration. 2 (view conditions) × 5 (temporal sections) ANOVA reveals there is neither significant difference of the view conditions (p = 0.09) nor of the temporal sections (p = 0.40). The error bars at both view conditions show large variations, indicating the individual differences of the perceptual flip frequencies.

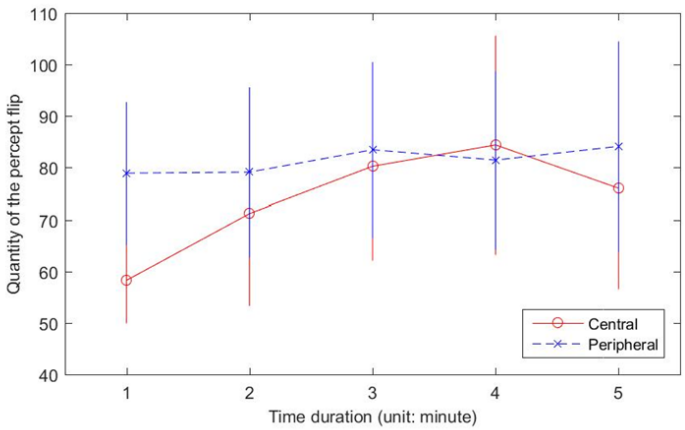
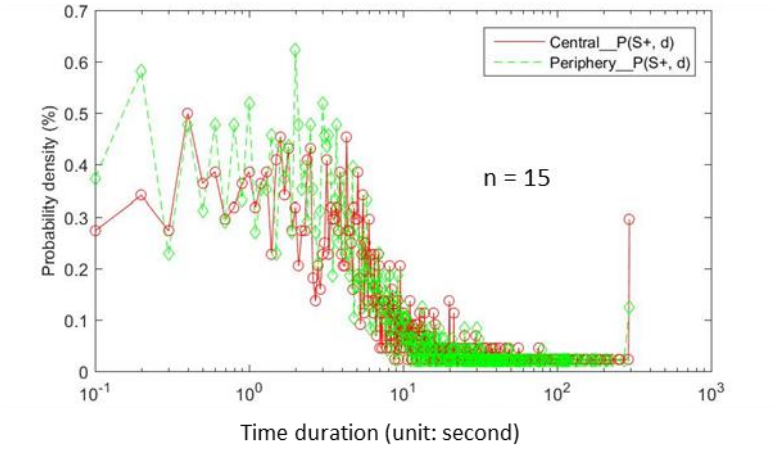


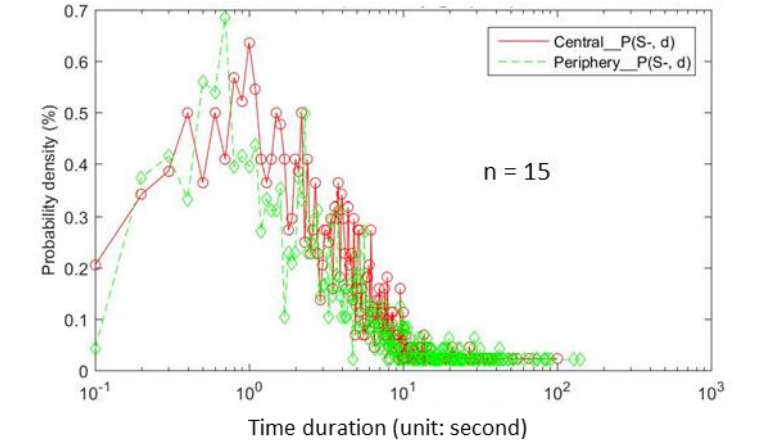
Fig. 4.4: Temporal distribution of the mean quantities of the percept flip at the central and peripheral view conditions. The red and blue lines represent the central and peripheral view conditions respectively. The error bars represent the standard errors of the means. There is no significant difference of the view conditions or of the time sections by using the 2 (view conditions) × 5 (time sections) ANOVA (n = 15).

Figure 4.5 (A) and (B) show the probability distributions (P) of S_+ and S_- percepts respectively using 15 participants' data. The horizontal axes in both figures are the log time duration (unit: second), and the vertical axes are the probabilities of the dominance duration (the dominance duration is calculated in 100 ms). The red and green lines represent the central and peripheral view conditions respectively. From the distributions, we can see the peaks of the probability distributions are around t = 1 s. However, the densities are all lower than 1%. The dominant durations of P_+ and P_- are mainly larger than one second at both the central and peripheral view conditions, some even reach to several-hundred seconds, especially at the S_+

percept.



(A) Probability distribution of S_+ percept



(B) Probability distribution of S_- percept

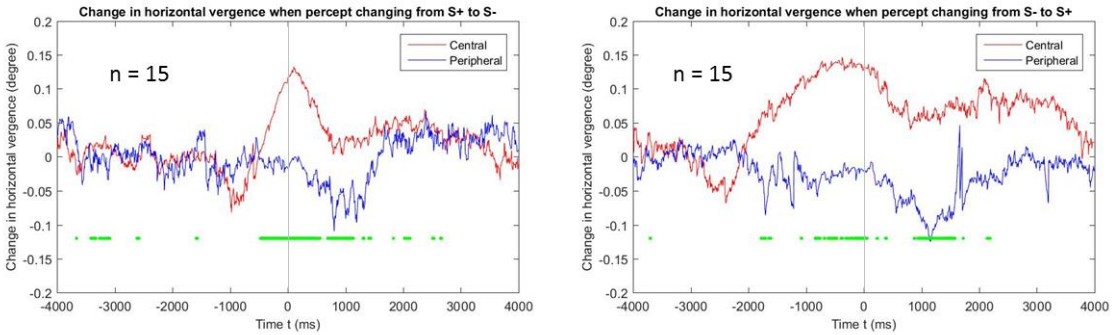
Fig. 4.5: Probability distribution of the S_+ and S_- percepts at the central and peripheral view conditions. (A) Probability distribution at the central view condition; (B) Probability distribution at the peripheral view condition. The horizontal axes in both figures represent the log time duration (unit: second), and the vertical axes represent the probability duration (unit: %) with 15 participants' data. The red and green lines represent the P_+ and P_- respectively.

4.2.2 Vergence eye movements, pupil size, blink and (micro)saccade

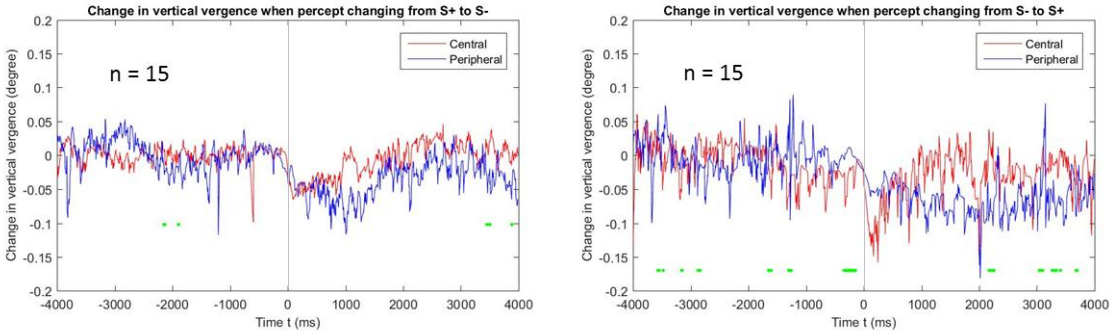
The DataViewer and Matlab software were used to analyze the eye positions and pupil sizes with the fifteen participants' data. Firstly, the eye tracking data were temporally aligned to the behavioral data based on the time stamps recorded by the Eyelink II system. Secondly, the moments of the percept flips based on the key press were set as t0 = 0, t = [-4000, 4000]

ms was defined as the time window, and the x, y positions and the pupil sizes of the left and right eyes within this time window were extracted. Thirdly, the changes in horizontal vergence, vertical vergence and pupil size were calculated by using the method described in chapter 3. Those which contained another flip in the time duration t = [-4000, 0] were excluded to make sure the distribution curves before percept flip were pure. Thirdly, a 9-sample median filter method were used to remove the noise of the raw eye tracking data, and then the mean across all trials and all participants were calculated and plotted into curves.

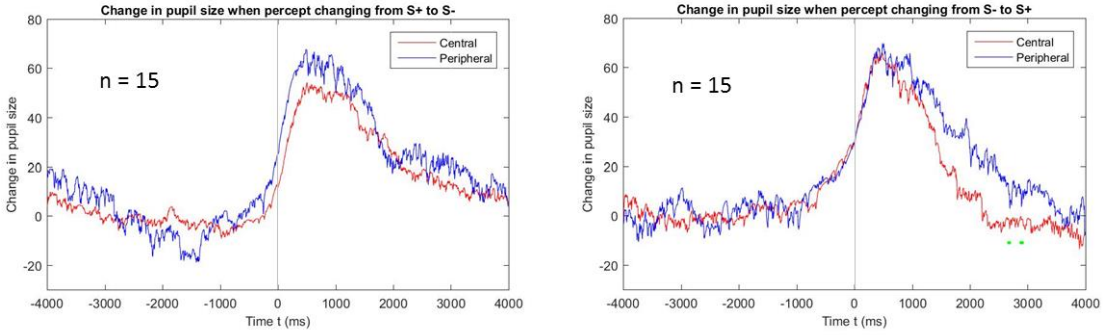
Fourthly, we checked the distribution data in previous step, selected the beginning range of the time window t = [-4000, -3000] ms and calculated the means of changes in horizontal vergence, vertical vergence and pupil size within this range. Then we used these means as the baseline values, shifted the distributions of changes in horizontal vergence, vertical vergence and pupil size by subtracting the baseline means from the distribution data, and calculated the significance between the central and peripheral view conditions. We defined t = [-4000, -3000] ms as baseline, because the distribution data in this range which we got in the third step were quite stable (temporally far enough from the moment of the percept flip) when percept changing from S_+ to S_- and also from S_- to S_+.



(A) Change in horizontal vergence



(B) Change in vertical vergence



(C) Change in pupil size

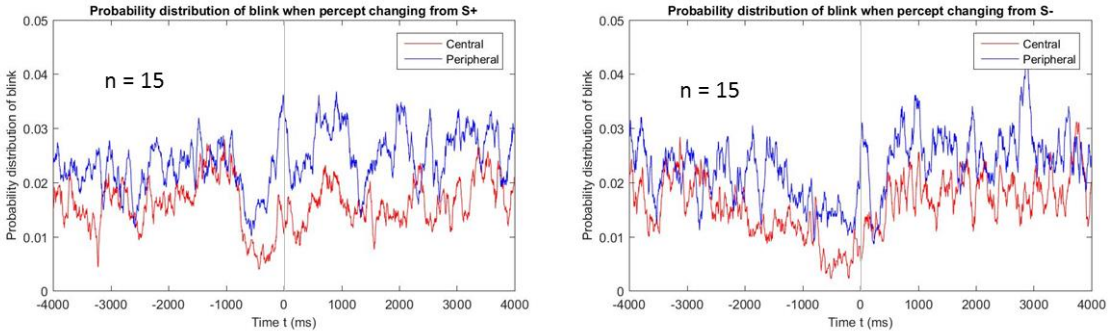
Fig. 4.6: Changes in horizontal vergence, vertical vergence and pupil size when percept changing from S_+ to S_- and from S_- to S_+. The horizontal axes represent the time window t = [-4000, 4000] ms, in which t = 0 means the moment of percept flip based on the key press; and the vertical axes represent the distributions of change in horizontal vergence, vertical vergence, and pupil size respectively. The left and right panels represent the results when percept changing from S_+ to S_- and from S_- to S_+ respectively. The red and blue lines represent the central and peripheral view conditions respectively. The horizontal green bar means the significant difference between the central and peripheral view condition by using the matched t-test ($p < 0.05$, n = 15).

Figure 4.6 shows changes in horizontal vergence, vertical vergence and pupil size when percept changing from S_+ to S_- and also from S_- to S_+. The horizontal axes represent the time window t = [-4000, 4000] ms, in which t = 0 means the moment of the key press when perceiving the flip; and the vertical axes represent the distributions of changes in horizontal vergence, vertical vergence and pupil size respectively. The left and right panels represent the results when percept changing from S_+ to S_- and from S_- to S_+ respectively. The red and blue lines represent the central and peripheral view conditions respectively. The horizontal green bar means the significant difference between the central and peripheral view condition by

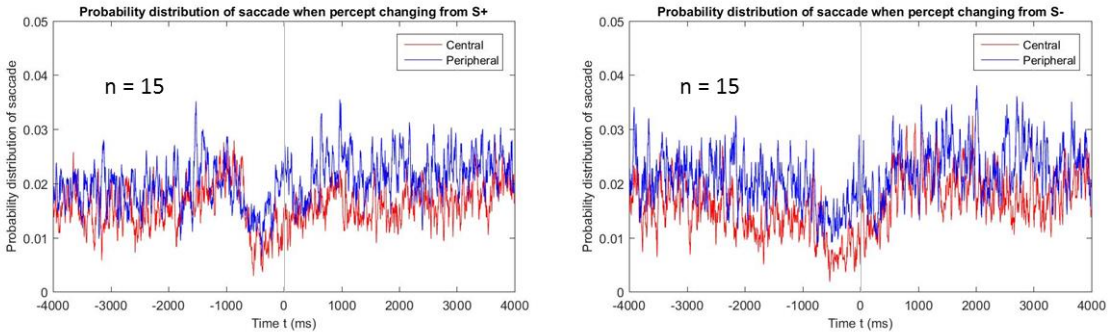
using the matched t-test ($p < 0.05$, n = 15).

In both left and right panels of Figure 4.6 (A), at the central view condition, there are obvious increases (convergence) of change in horizontal vergence before the key press (t = [-1000, 0] ms for the left panel, and t = [-2000, 0] ms for the right panel). However, at the peripheral view condition, there is no obvious vergence change before the key press. The horizontal green bars indicate there are significant differences of changes in horizontal vergence between the central and peripheral view conditions. In both panels of Figure 4.6 (B), there are no obvious difference of change in vertical vergence between the central and peripheral view conditions. In both panels of Figure 4.6 (C), the changes in pupil size keep stable before the key press; whereas there are obvious increases of change in pupil size at both the central and peripheral view conditions after the key press.

Moreover, we also extracted the blink and (micro)saccade data and calculated the possibility distributions within the same time window t = [-4000, 4000] ms. To prevent the relatively low amount of the perceptual flip data in the defined time window, we calculated the possibility distributions when percept changing from S_+ to S_-/neither, and from S_- to S_+/neither separately. Figure 4.7 shows the probability distributions of blink and (micro)saccade when percept changing from S_+ and from S_- separately. The horizontal axes in Figure 4.7 (A) and (B) represent the time window t = [-4000, 4000] ms, in which t = 0 means the moment of percept flip based on key press; and the vertical axes represent the probability distributions of blink and (micro)saccade respectively. The left and right panels in both (A) and (B) represent the results when percept changing from S_+ and from S_- respectively. The red and blue lines represent the central and peripheral view conditions respectively with fifteen participants' data.



(A) Probability distribution of blink



(B) Probability distribution of (micro)saccade

Fig. 4.7: Probability distributions of blink and saccade when percept changing from S_+ and from S_- separately. The horizontal axes represent the time window t = [-4000, 4000] ms, t = 0 means the moment of percept flip based on key press; and the vertical axes represent the probability distributions of blink and saccade respectively. The left and right panels represent the results when percept changing from S_+ and from S_- respectively. The red and blue lines represent the central and peripheral view conditions respectively. The figures are drawn from 15 participants' data.

The time point t = 0 in Figure 4.6 and 4.7 represents the moment of perceptual flip by key press, and may have around 500 ms latency caused by the motor action than the actual moment of percept alters (Liu et al., 1992). If defining t = -500 ms as the moment of perceptual flip, at the central view condition, the changes in horizontal vergence at both the left and right panels are a little bit earlier than this moment (Figure 4.6 (A)). At both the central and peripheral view conditions, the increases of the pupil sizes of the two panels are just at the moment of perceptual flip (Figure 4.6 (C)); there are reduced probabilities of blink and (micro)saccade at the moment of perceptual flip, meaning that the fixation duration is

longer at this moment, which might indicate the onset of the coming percept flip (Figure 4.7 (A) and (B)) (van Dam and Van Ee, 2006b). Moreover, at the peripheral view condition, there are peaks of probabilities of blink and (micro)saccade at the moment of button press (t = 0), which might be caused by the motor action of key press, indicating the involvement of the visual attention or effort (van Dam and Van Ee, 2006b).

4.3 Discussion

4.3.1 Different amounts of the F_+, F_- and F_N suggest different underlying mechanisms between the central and peripheral visions

Figure 4.2 (A) and (C) show there are significant differences of F_+ and F_N respectively between the central and peripheral view conditions, whereas no significant difference of F_- as shown in Figure 4.2 (B). The bias of S_+ percept at the central view condition indicates the stronger top-down feedback as proposed in previous study (Zhaoping, 2017). The F_N is significantly larger at the peripheral view condition than that at the central view condition, which might also indicate different underlying mechanisms between the two view conditions. In the conventional binocular rivalry, the two parallel pathways, as the P (Parvocellular) pathway and M (Magnocellular) pathway, have different roles. The P pathway plays an important role in visual recognition and cannot tolerate for the binocular uncorrelated information, thus will lead to a competition between the binocular uncorrelated inputs (when the contrasts of the two eyes' inputs are not too low (Dayan, 1998; Liu et al., 1992)) and work in a winner-take-all manner. In contrast, the M pathway prefers to combine the binocular information and can tolerate for the binocular uncorrelated inputs since it does not serve for visual recognition (He et al., 2005). Similarly, in this study, based on our participants'

feedbacks, they could perceive the motion direction during almost the whole time duration of the central view condition, while felt hard to judge the motion direction at the peripheral view condition. From their behavioral data, the significantly larger F_N at the peripheral view condition also reflects the difference between the two view conditions. Hence, the different amounts of the F_+, F_- and F_N between the central and peripheral view conditions suggest the functional differences of the central and peripheral visions.

4.3.2 Temporal dynamics of the S_+ and S_- percepts at the central vision might indicate the hierarchy of binocular rivalry

In Figure 4.3, there seems to have a competition between the S_+ and S_- precepts over time at the central view condition; whereas there is no such tendency at the peripheral view condition. These are consistent with the result in the conventional binocular rivalry that the rivalry mainly occurs at the central vision.

To compare the difference between the conventional binocular rivalry and the rivalry in this study, we draw the schematics as shown in Figure 4.8. For the sake of intuition, we use the tilt feature as an example. The left panel shows the schematic of the conventional binocular rivalry. It is the rivalry between the monocular unambiguous inputs, which involves the local retinotopic level process (Chen and He, 2003), and does not cause the vergence change or eye movements (van Dam and van Ee, 2006a; He et al., 2005). The right panel shows the schematic of the rivalry in this study. The stimuli are ambiguous gratings for the monocular percepts, and unambiguous gratings for the S_+ and S_- percepts after the integration in V1 brain area. As proposed in previous study (Zhaoping, 2017), the bias towards the S_+ percept at the central vision might be due to the stronger involvement of the top-down feedback.

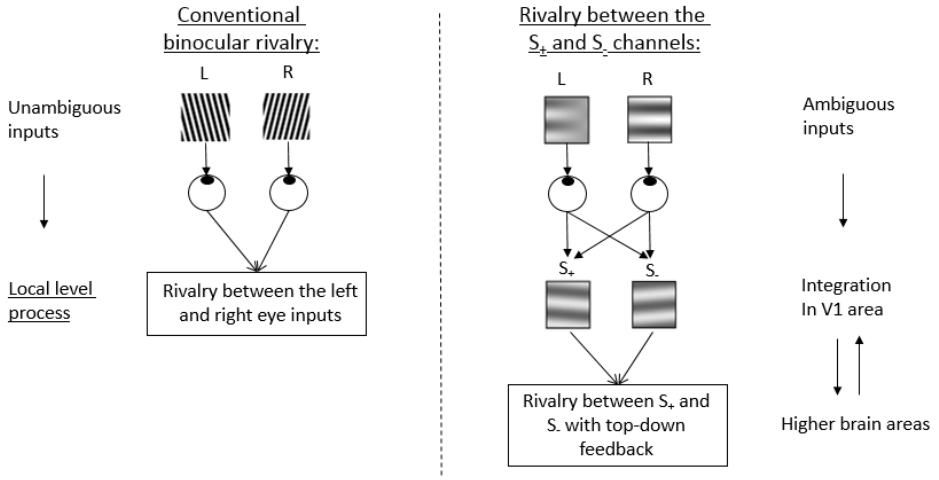


Fig. 4.8: Comparison of the conventional binocular rivalry and the hierarchical rivalry in this study. The left panel shows the schematic of the conventional binocular rivalry, which is the rivalry between the monocular unambiguous inputs, and involves the local retinotopic level process. The right panel shows the schematic of the rivalry in this study, which firstly integrates the monocular ambiguous inputs into S_+ and S_- channels, and involves the rivalry between the binocular neurons in S_+ and S_- channels with top-down feedback (for the sake of intuition, in this schematic we use the tilt feature as an example).

Previous studies proposed the hierarchical models and hierarchical neural mechanisms in the binocular rivalry (Dayan, 1998; Blake and Logothetis, 2002; Scocchia et al., 2014). Based on these studies, the top-down feedback to the visual inputs might have competitions and cause the percept flip, which might involve the early extrastriate areas and posterior inferotemporal cortex (Dayan, 1998; Blake and Logothetis, 2002; Scocchia et al., 2014). In our study, to perceive the motion direction, the bias towards S_+ percept at the central view condition indicates the feedback from higher brain areas are involved in the process (Zhaoping, 2017). To perceive the motion direction over time, the competition related to the higher level feedback may also be involved, which indicates the hierarchy of binocular rivalry, and is different from the direct competition between the visual inputs in the conventional binocular rivalry.

4.3.3 Temporal dynamics and percept flips might indicate the involvement of adaptation

The mechanisms of percept flip of binocular rivalry are quite controversial. Some studies reported that it was the neural adaptation caused the percept flip (Logothetis et al., 2011; Hollins and Hudnell, 1980), whereas others reported that the neural noise triggered the percept flip (Brascamp et al., 2006; Moreno-Bote et al., 2007). Recently, there are reports that both the adaptation and neural noise were involved in the percept flip (Kang and Blake, 2010; Roumani and Moutoussis, 2012; Shpiro et al., 2009).

In the conventional binocular rivalry, the visual input in one eye does not have stronger dominance than the other, and the probability distribution meets the gamma distribution with the peak probability around t = 1 s (Chen and He, 2003; Leopold and Logothetis, 1996). However, in our study, there was a strong dominance of the percept at the S_+ channel. Although the probability distributions in Figure 4.5 show the peaks around t = 1 s, the peak probability densities are less than 1%, which have big difference with the value in the conventional binocular rivalry (Chen and He, 2003). Moreover, the dominant durations of S_+ and S_- percepts in our study distribute from several seconds to even several-hundred seconds (especially at the S_+ percept) at both the central and peripheral view condition, suggesting the higher stabilities of the S_+ and S_- percepts, which might not be driven by the neural noise.

In Figure 4.3, at the first four minutes at the central view condition and the first three minutes at the peripheral view condition, the F_+ decreases gradually; after these time durations, the F_+ keeps stable. These suggest the involvement of adaptation and the saturation of the adaptation respectively. The involvement of the adaptation is similar as previous study (Hollins and Hudnell, 1980). At the central view condition, the top-down feedback was involved and participants needed to concentrate well on the stimuli, which might cause the visual fatigue when perceiving the stimuli over time. Along with time, the neurons at S_+

channel get tired and will give less response than before, thus the percept is alternated to S_- percept. From Figure 4.4, the frequencies of the percept flips increase over time at the central view condition before the saturation of the adaptation, and might correspond to the temporal dynamics as shown in Figure 4.4 (A). After saturation of the adaptation, the F_+ is around 76% at central vision, while only around 60% at the peripheral vision, suggesting the much stronger top-down feedback is involved at the central vision.

4.3.4 Vergence eye movement as a tool to detect the involvement of the higher brain areas

In this study, with the long duration stimuli, the extracted eye data within the time window ([-4000, 4000] ms) show that there are obvious vergence changes when the percepts flip at the central view condition; while almost no vergence change at the peripheral view condition before the percept flip (Figuew 4.7 (A)).

The vergence eye movements suggest the involvement of the binocular disparity neurons. Previous studies (Cumming and Parker, 1997; Cumming and Parker, 2000; Masson et al., 1997; Poggio and Fischer, 1977) reported that the binocular disparity neurons in V1 brain area might have a role to guide vergence eye movements. As discussed in previous study, the bias towards S_+ percept at the central vision suggested the feedback from higher brain areas was involved, whereas this feedback was weaker or even absent at the peripheral vision (Zhaoping, 2017). One possibility is that after receiving the visual inputs from retina, the V1 neurons in both the central and peripheral visual areas provide feedforward signals to the higher brain areas; then the neurons in the central visual area receive feedbacks from the higher brain areas based on the analysis-by-synthesis computation (Zhaoping, 2014; Zhaoping, 2017). Since the prior knowledge is that the binocular inputs are correlated, which is inconsistent with the actual inputs, thus the V1 neurons at central vision will send signals to the superior colliculus for mediating the vergence eye movement (Zhaoping, 2014). The proposal that the superior

colliculus control for the vergence eye movement can be supported by Chaturvedi and van Gisbergen (2000) and Van Horn et al. (2013).

Since the responses of the V1 neurons are the weighted sum of the signals from the S_+ and S_- channels (Zhaoping, 2014; Li and Atick, 1994; May and Zhaoping, 2016), the percept changing from S_+ to S_- might cause different tendencies of V1 responses when comparing with the percept changing from S_- to S_+. However, the changes in horizontal vergence as the left and right panels of Figure 4.7 (A) show the similar tendencies (change to convergence before percept flip). Another possibility is that the higher brain area related to visual attention is involved in the processing. As explained in the introduction, Solé Puig et al. (2013a) used the cue/no-cue paradigm to perceive the peripheral target stimuli, and found the vergence eye movements involved in the cue paradigm. They proposed that the vergence change linked to the covert attention, which might involve the frontal eye field (FEF). Other previous studies also reported about the link of the vergence eye movements and visual attention (Hoffman, 1998; Gamlin and Yoon, 2000; Schafer and Moore, 2007). In this study, the special designed stimuli were used, which produced ambiguous percepts for monocular inputs and unambiguous percepts for the S_+ and S_- channels. To perceive the motion direction, at first the brain needs to integrate the ambiguous monocular inputs into S_+ and S_- channels; to perceive the motion direction over time, the higher brain area for visual attention may be also involved. To explain these results, the schematic is redrawn as shown in Figure 4.9 by referring to Zhaoping (2014). The V1 neurons at both the central and peripheral visual areas provide feedforwards to the higher brain area. However, since there is functional difference between the central and peripheral view conditions that the former mainly works for visual decoding, which involves the higher brain area, and the latter mainly serves for visual selection, which is a bottom-up process, thus the top-down visual attention is more involved at the central vision. Consequently, the visual attention at the central visual area controls the FEF to mediate the

vergence eye movements, indicating the involvement of higher brain areas in the visual processing.

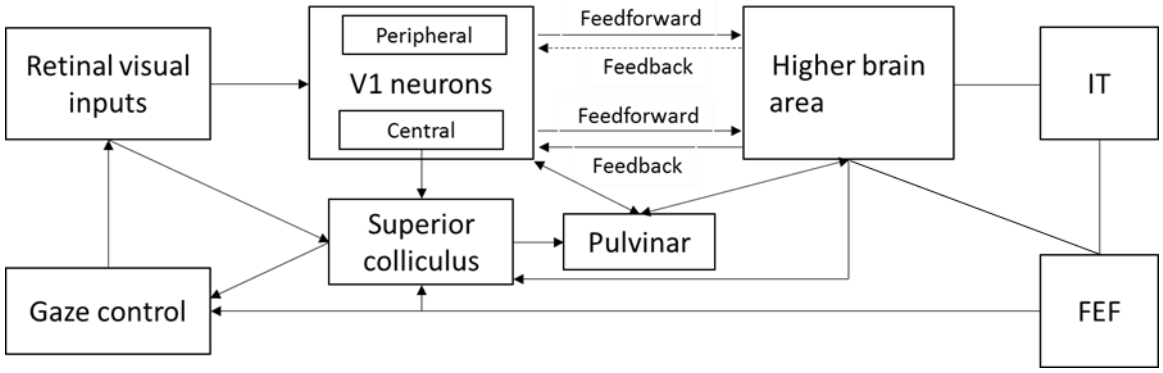


Fig. 4.9: The hierarchical visual processes in the brain.

4.3.5 The change in pupil size might be explained by the existence of the LC-NE complex

When comparing the peaks of changes in pupil size in Figure 4.6 (C) with the peaks of changes in horizontal vergence in Figure 4.6 (A), the former has several-hundred milliseconds latency than the latter. If the change in pupil size is caused by the change in horizontal vergence, the pupil size should be constricted, since for example at the central view condition, the change in horizontal vergence is convergence before the key press. However, the dilations of pupil sizes might suggest that the change in pupil size has no relationship with the change in horizontal vergence.

Figure 4.6 (C) shows the change in pupil size links with the percept flip. After percept flip, the pupil size increases obviously at both the central and peripheral view conditions, which have no relationship with the change in horizontal vergence. We found these results can be supported by the previous study in binocular rivalry. Einhäuser et al. (2008) used four kinds of dichoptic stimuli including plaid, Necker cube, “structure from motion” and auditory to study the change of pupil size in perceptual rivalry. They found that the pupil dilation linked with the percept flip. Their explanation was that during the percept flip, the locus coeruleus

(LC) released a certain amount of norepinephrine (NE), and the NE caused the pupil dilation. In other words, it was the LC-NE complex that were produced at the moment of percept flip and caused the dilation of pupil sizes. In our study, the relationship between the pupil sizes and percept flips are similar as this study (Einhäuser et al., 2008), thus might also be explained by the existence of LC-NE complex when percept flips.

4.4 Conclusion

In conclusion, the special designed dichoptic stimuli were used to produce ambiguous percept for the monocular eye and unambiguous percept for the binocular summation (S_+) channel and binocular differencing (S_-) channel. Based on the previous conclusion that the bias towards S_+ percept at central vision is because top-down feedback is involved for the visual recognition, we investigated the temporal dynamics by elongating the time duration of the dichoptic stimuli. The results show that the temporal dynamics for visual recognition can cause the vergence eye movements, which might involve the even higher brain areas for visual attention; the temporal dynamics of the S_+ and S_- percepts at the central view condition show a tendency of competition between the S_+ and S_-. Since the visual recognition (S_+ or S_- percepts) at the central vision involves the feedback from the higher brain areas, the temporal dynamics of the S_+ and S_- percepts might indicate the hierarchy of the binocular rivalry, which is different from the direct competition between the visual inputs in the conventional binocular rivalry.

Part II: Depth perception study

Chapter 5 Introduction about depth adaptation

5.1 Depth adaptation and previous studies

Adaptation is a process that after prolonged exposure to specific visual stimuli, the neural activities of related brain areas will be changed. Taking this advantage, we can explore the underlying mechanisms of visual perception using different adapting stimuli to selectively manipulate the neuron activities at different brain areas. The adaptation paradigm has been used in numerous studies to reveal the mechanisms of depth perception for decades. For example, Domini et al. (2001) manipulated the viewing distances of adaptation and test stimuli to investigate whether depth adaptation to curved surfaces is a second-order disparity-specified or a shape-level percept-specified process. Based on the principle that disparity is distance-independent and curvature is distance-dependent, they examined the depth aftereffects with four combinations of 20 cm and 80 cm viewing distances as the adaptor-probe pairs. If the adaptation is disparity-specified, there will be no difference of aftereffects between the same and different distances of adaptor-probe pairs. If it is percept-specified, there will be a significant difference. Their results showed a significant difference of aftereffects against distance, suggesting the shape-level adaptation was involved. Similarly, Berends et al. (2005) examined the mechanism of stereo-slant perception based on the similar principle that the percept-specified slant adaptation is viewing distance-dependent.

They fixed the adaptation stimulus at 57 cm distance, and changed the position of the test stimulus at various distances. The results showed that different amounts of aftereffects were induced at different viewing distances of test stimuli, suggesting the percept-specified adaptation. Both studies reported the shape-level adaptation process, which provided good evidence to elucidate the underlying mechanisms of depth adaptation. In the aspect of disparity-specified adaptation, Berends and Erkelens (2001) manipulated the adapting stimuli to be perceived as a fronto-parallel plane by changing the vertical disparity while fixing the horizontal disparities, and test stimuli with only horizontal disparity. Although the perceived adaptation stimuli were fronto-parallel, the test stimuli was significantly different from zero horizontal disparity. They concluded that the disparity per se, not the perceived depth, was adapted by the visual system. Yan and Shigemasu (2015) dynamically changed the location, size, and depth of spherical adaptation stimuli and found that both disparity- and percept-specified processes are involved in stereo-curvature adaptation.

However, the simple stimuli such as a single curved, flat surface and the single objects used in the previous studies might limit the conclusions to a certain range corresponding to simple objects in the real world (Welchman, 2016). To further investigate the mechanism of depth adaptation, we harnessed the characteristics of disparity-specified sinusoidal corrugation that shows the depth structure within a continuous surface. This surface has complex shape information that consists of both crossed and uncrossed disparities, and both positive and negative curvatures (Howard and Rogers, 2012). However, on the other aspect, the sinusoidal corrugation is quite simple in the spatial frequency domain. It contains only one-dimensional information and is widely used as a stimulus in luminance and other domains including binocular disparity. For example, by using the sinusoidal corrugation, investigation about phase-dependent adaptation (Graham and Rogers, 1982; Ryan and Gillam, 1993), phase-independent adaptation (Schumer and Ganz, 1979), and both adaptations (Graham and

Rogers, 1983) was reported.

Moreover, the combination of two orthogonal corrugations as plaid stimulus, which contains two dimensional information in frequency domain, is used as stimulus in both the luminance and disparity domains to examine whether there is relatively higher-level mechanisms (Bowd et al., 2000; Georgeson and Shackleton, 1994; Hibbard and Langley, 1998). In previous studies, some researchers focused on the investigations of the relationship between the one-dimensional sinusoidal structure and two-dimensional plaids. If there had some relationship between them, we could use the basis in one-dimensional structure to predict the mechanisms of the two-dimensional structure or objects, and vice versa. For example, in luminance domain, Georgeson and Shackleton (1994) examined the perceived contrast of sine-wave gratings and plaids by manipulating different spatial frequencies and orientations. The results showed that the perceived contrast of the plaids was lower than the gratings when both of them had the same physical contrast. Cherniawsky and Mullen (2016) reported a similar study which was extended in both luminance and chromatic domains. They investigated the perceived contrasts by employing chromatic and achromatic sine-wave gratings and plaids separately, and found in contrast to the component gratings, the summation relationship of the perceived contrast of chromatic plaids was lower than that for the achromatic plaids, suggesting the greater suppression in chromatic domain. In disparity domain, Hibbard and Langley (1998) examined the thresholds of slant and inclination of disparity-defined sinusoidal corrugation and plaid surface at different orientations. The stimuli were manipulated in rotation, horizontal shear and horizontal compression three binocular transformations. They found the slant and inclination thresholds of the sinusoidal corrugation could be used to predict the thresholds of the plaid surface. Bowd et al. (2000) investigated the perceived coherence of moving plaids by adapting to moving plaids or moving components separately. The adapting plaids or components and test plaids had eight

combinations cross the luminance and disparity domains. And they found the cross-domain adaptation, meaning that adapting to moving plaids or gratings in luminance domain could cause negative aftereffects of test plaids in disparity domain, and vice versa.

However, in the studies of the phase dependency and independency depth adaptation (Graham & Rogers, 1982; Ryan & Gillam, 1993; Schumer & Ganz, 1979; Taya et al., 2005; Graham & Rogers, 1983), neither the vertical-orientated corrugation nor the relationship between the plaids and components was investigated. In the studies of summation and suppression relationship between plaids and components in different domains, the aftereffects of depth adaptation were not examined (Georgeson & Shackleton, 1994; Cherniawsky and Mullen, 2016; Bowd et al., 2000). Although Hibbard and Langley (1998) mentioned the common mechanism was involved in the processing for the disparity-defined grating and plaid stimuli, they mainly focused on the orientation effects (slant and inclination) of the two stimuli.

5.2 The purpose of this study

To have a further investigation on the depth adaptation, we conducted experiment 3 and experiment 4, which might involve the early and middle level of brain areas, like V1, V2, V3 (might V4) as shown in Figure 5.1.

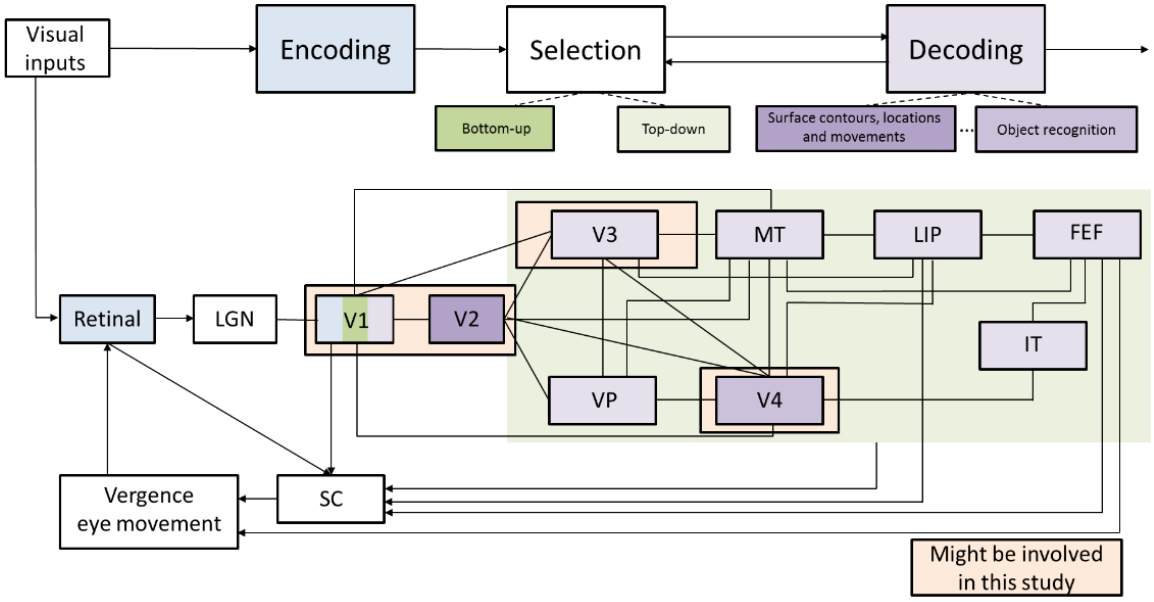


Figure 5.1 schematic of brain areas might be involved in the depth adaptation study (Zhaoping, 2014)

In experiment 3, we investigated the retinal dependency and/or independency of depth aftereffects (experiment 3.1) and orientation-independency of depth adaptation (experiment 3.2). Three kinds of disparity-defined adaptor-probe pairs (horizontal-horizontal, vertical-vertical corrugations and plaids-plaids) were used as stimuli in experiment 3.1. Four combinations of disparity-defined adaptor-probe pairs between horizontal and vertical corrugations were used as stimuli in experiment 3.2. By manipulating the phases of the corrugations and plaids in static or randomly changing conditions, we separated the local retinal-level depth adaptation. By adapting to the same- and different-oriented corrugations, we verified whether any relative higher level of depth adaptation was involved.

In experiment 4, we investigated whether disparity- or shape-level process was related to the adaptation effects of the corrugated surface, and also whether any different amounts of depth aftereffects were induced between the adaptors with and without certain surfaces. Different from experiment 3, in this experiment, the adaptor and probe were different shapes.

Chapter 6 Experiment 3: Phase- and orientation-independency of depth adaptation

In experiment 3.1, we investigated the retinal dependency and independency of depth aftereffects by manipulating the adaptation stimuli in phase static and randomly-changing conditions. There were three kinds of disparity-defined adaptor-probe pairs as horizontal-horizontal, vertical-vertical corrugations and plaids-plaids stimuli. The adaptation and test stimuli were always the same-shape and same-orientation pairs in experiment 3.1.

However, whether the cross-orientation of the adaptor-probe pairs could cause different amounts of depth adaptation was still unclear. To have a further investigation on this issue, in experiment 3.2, we examined the orientation-independency of depth adaptation by manipulating the adaptation and test corrugations with different orientations, and compared the aftereffects with the same orientation pairs. The experiment was implemented at phase randomly-changing condition, so the local retinal level depth adaptation was cancelled out.

6.1 Experiment 3.1: retinal-dependency and independency of depth adaptation

To investigate whether both local retinal level and higher level depth adaptation were involved, we examined the retinal-dependency and independency of depth aftereffects by using three kinds of disparity-defined stimuli, which were horizontal-, vertical-orientated corrugations and the combination (plaids) of these two corrugations. For each stimulus, we

manipulated the phases in static and randomly-changing conditions. In phase static condition, the adaptation and test stimuli had the same phase within-trial, but randomly changed phases between trials. In phase randomly-changing condition, the phases were evenly distributed in all the position, the average disparity at each position was close to zero.

If the depth adaptation only involves the retinal dependent processing, there will be no aftereffects in phase randomly changing condition. If it only involves the retinal independent stage, the aftereffects in phase static condition will be equal to the amount of phase randomly changing condition. If it involves both the retinal-dependent and independent processing, the aftereffects in both conditions will not be zero, and the amount of aftereffects in phase static condition should be larger than the amount in phase randomly-changing condition.

6.1.1 Method

6.1.1.1 Participants

Ten students aged 20-35 years old (3 males, mean age: 26.5 years) from Kochi University of Technology were recruited as participants. All of them had normal or corrected-to-normal vision, and passed the stereo perception and stereo acuity test (less than 1 arcmin) with our own program. Participants were naïve to the experiment aims and paid for their time. The authors did not serve as participants. All experiments and procedures were approved by the Research Ethics Committee of Kochi University of Technology and conformed to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to experiments.

6.1.1.2 Apparatus

Stimuli were presented on a 22-inch CRT color display (RDF223H, Mitsubishi) with 1024 × 768 resolution and 120 Hz frame refresh rate. The luminance of the display was measured by CS-100A colorimeter measurement (Minolta, Japan) and linearized using a look-up table

method. We programmed to present the experimental stimuli by using Matlab (Mathworks, Natick, MA) with PsychToolbox Version 3 (Brainard, 1997; Pelli, 1997). During experiments, participants sat in a dark room with frontal parallel to the surface of the display and watched the stimuli via a pair of stereoscopic wireless LCD glasses (NuVision 60GX, MacNaughton, Inc., OR, USA). The refresh rate of the LCD glasses was 120 Hz. Thus for each eye the frame rate was 60 Hz. No flicker was reported by participants. A chin rest was used to prevent head movement.

6.1.1.3 Stimuli

Anti-aliased pseudo random white dots (29.7 cd/m^2) were presented on a gray background (9.9 cd/m^2). The random dots were presented in dynamic pattern with 5 Hz frequency to prevent afterimages. Their density was 30.6 dots/deg^2. At the center of the display, a nonius fixation with lower part T- and upper part reversed T-shape was shown to the left and right eye separately. To guarantee the eye vergence, participants were asked to maintain vertical lines of the two T parts collinear and horizontal lines overlapped during the whole experimental procedure. After eye vergence, the nonius was perceived as a cross. The lengths of both the horizontal and vertical lines were 1.17 arcdeg.

6.1.1.4 Procedure

In both phase static and randomly-changing conditions, the spatial frequency of the sinusoidal corrugations was 0.25 cpd. In phase randomly-changing condition, the phase of adaptation stimuli were presented in random order with 5 Hz temporary frequency. Method of constant stimuli was used in both conditions.

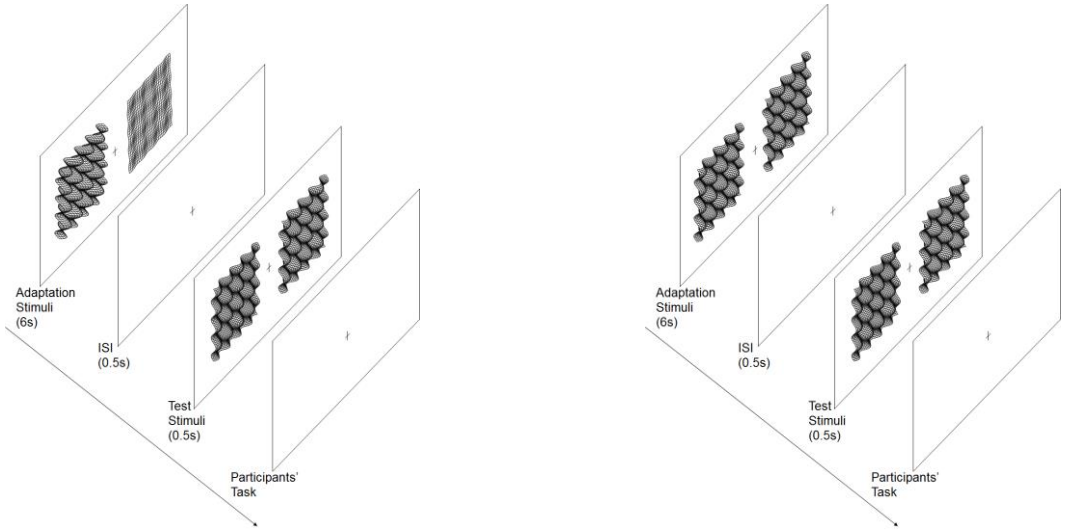
For each stimulus type and phase condition, we used two adaptors simultaneously positioned at the left and right sides of the CTR display with large-small or middle-middle adaptation amplitudes respectively. Figure 6.1 showed the schematic of the test procedure by using plaid stimuli as an example. The two adaptors were presented for 6 s. The size of the

stimuli at each side was 14.0 × 14.0 arcdeg. After 0.5 s time interval of blank gray background, the test stimuli with one side comparison and the other side test stimulus were presented simultaneously in 0.5 s time duration. The comparison stimulus had a fixed amplitude, while the test stimulus had nine levels of amplitudes and was presented in random order in each trial. The positions of the test and comparison stimulus were presented at the left and right sides of the display in a counterbalanced randomized order. Participants' task was to judge which side had the larger amplitude based on their subjective perception, and report their choice by a button press with the two-alternative forced-choice (2AFC) method. No feedback of correctness was given. After participants made their choice, the next trial was triggered automatically.

In large-small amplitude adapting condition, one side of the adaptors had large amplitude, and the other side had small amplitude. The sides for the large and small adaptors were in a counterbalanced random order. In middle-middle amplitude adapting condition, both adaptors had the same amplitude. For plaid stimuli, the disparities of the adaptors and test stimuli were doubled in contrast to the horizontal- and vertical-orientated corrugations. The parameters were shown in table 6.1.

Table 6.1: Parameters in large-small and middle-middle amplitude adapting conditions

Stimulus Type	**Depth Amplitude of Sinusoidal corrugation**	**Depth Amplitude of Plaid stimulus**
Adaptation stimulus		
-- Large	20.2 arcmin	40.4 arcmin
-- Middle	12.1 arcmin	24.2 arcmin
-- Small	4.1 arcmin	8.2 arcmin
Test stimulus	9.1-15.2 arcmin(9 levels)	18.2-30.4 arcmin(9 levels)
Comparison stimulus	12.1 arcmin	24.2 arcmin



(a) Large-small depth amplitudes condition (b) Middle-middle depth amplitudes condition

Fig 6.1. Schematic of the test procedure by using plaid stimuli as an example. (a) The large-small amplitude adapting condition, (b) the middle-middle amplitude adapting condition. The horizontal- and vertical-orientated stimuli were shown in the same way. The two adaptors were presented for 6 s. After 0.5 s time interval of blank gray background, the test stimuli with one side comparison and the other side test stimulus were presented simultaneously in 0.5 s time duration. The comparison stimulus had a fixed amplitude, while the test stimulus had nine levels of amplitudes and was presented in random order in each trial. The positions of the test and comparison stimulus were presented at the left and right sides of the display in a counterbalanced randomized order. Participants' task was to judge which side had the larger amplitude based on their subjective perception, and report their choice by a button press with the two-alternative forced-choice (2AFC) method.

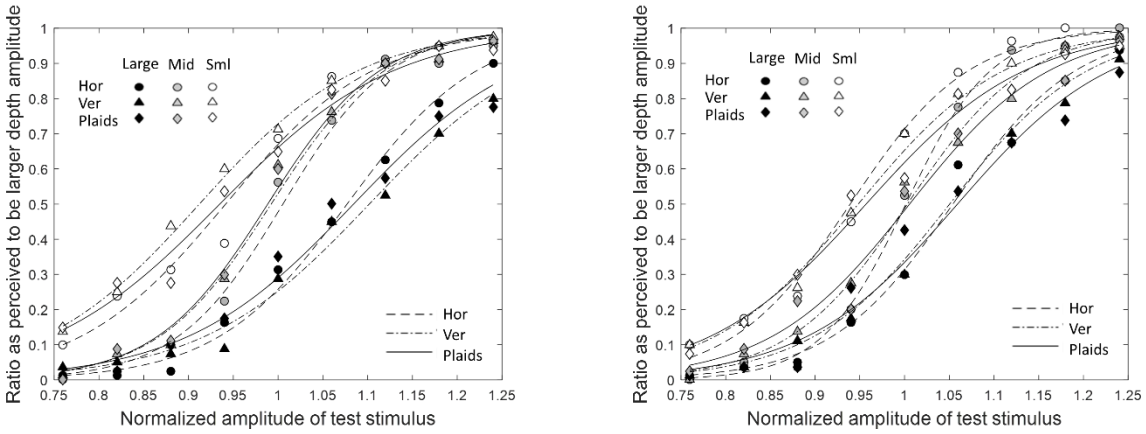
Before the experiment, participants were trained by our own practice program. In the practice trials, the adaptation process was eliminated to avoid any potential influence on the experiment results, and only the test stimuli were shown. Feedback of the correctness was given to make participants have a basic concept on their perception. There were 36 practice trials for each type of stimuli.

During experiment, the stimuli were presented with combinations of two phase types (phase static and randomly-changing conditions), three stimulus types (Horizontal-, vertical-orientated and plaids conditions), and three adaptation-amplitude types (large, middle and small conditions). For each condition, there were 216 trials to produce 8 repeats at each test point. Different conditions were block designed, and divided into six sessions which were implemented at different days. In each session, blocks were presented subsequently with two

minutes break. All the blocks were counterbalanced within- and between-subjects.

6.1.2 Results

In experiment 3.1, the ratio as perceived to be larger depth amplitude was calculated. Figure 6.2 showed the psychometric sigmoidal curves fitted with ten participants' average data by using logistic function method (Kingdom and Prins, 2010). Figure 6.2 (a) represented the phase static condition and (b) the phase randomly-changing condition. In both figures, the horizontal axis showed the normalized amplitude of test stimulus, and the vertical axis showed the ratio as perceived to be larger depth amplitude. When comparing with these curves, we could see the sigmoidal curves showed obvious shifts caused by the adapting amplitude type, while little shifts related to the stimulus type.



(a) Phase static condition (b) Phase randomly-changing condition

Fig 6.2. Fitted psychometric sigmoidal curves in phase static and phase randomly-changing conditions with ten participants' average data. (a) Phase static condition; (b) Phase randomly-changing condition. In both panels, the horizontal, vertical and plaid stimuli were shown as circles, triangles and diamonds respectively. The average values of the "ratio as perceived to be larger depth amplitude" after adapting to large-, middle- and small-amplitude adaptors were shown in solid, gray and open symbols respectively. The fitted curves of horizontal-, vertical-orientated and plaid stimuli were drawn as dash-dot, dash and solid lines respectively.

The point of subjective equality (PSE) were calculated with each participant's data by using the generalized linear model method (Kingdom and Prins, 2010). A three-factor repeated measures analysis of variance (ANOVA) (2 x 3 x 3) was used to analyze PSE data with phase

type, stimulus type and adaptation amplitude type as independent variables. The results showed a significant difference of adaptation amplitude type ($F(2, 18) = 20.35$, $p = 0.00$, generalized $\eta^2 = 0.57$), while no significant difference of phase type ($F(1, 9) = 1.34$, $p = 0.28$, generalized $\eta^2 = 0.00$) or stimulus type ($F(2, 18) = 0.08$, $p = 0.93$, generalized $\eta^2 = 0.00$). There were significant interaction between phase type and adaptation amplitude type ($F(2, 18) = 6.02$, $p = 0.01$, generalized $\eta^2 = 0.07$). The post hoc analysis showed there were a significant difference of phase type at large amplitude adaptation condition ($F(1, 9) = 8.48$, $p = 0.02$, generalized $\eta^2 = 0.11$), while no significant difference of phase type at middle ($F(1, 9) = 1.77$, $p = 0.22$, generalized $\eta^2 = 0.04$) and small (($F(1, 9) = 2.98$, $p = 0.12$, generalized $\eta^2 = 0.04$)) amplitude adaptation condition. To show the PSE value in a more readable way, we calculated the PSE shifts by subtracting 1 from the PSE values. As a result, the PSE shifts in large, middle and small amplitude adapting conditions were positive, close to zero and negative respectively (Figure 6.3).

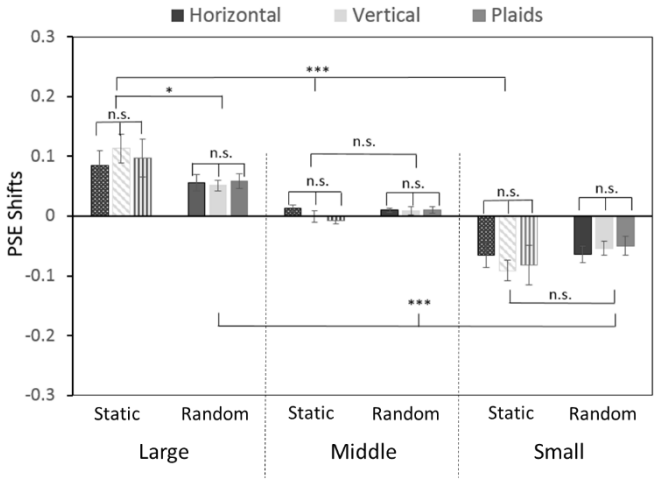


Fig 6.3. Results of experiment 3.1 with ten participants' data. There was a significant difference of adaptation stimulus amplitude (large, middle and small amplitude conditions), a significant difference of phase type (phase static and randomly changing) at large amplitude adapting condition, while no significant difference of stimulus type (horizontal-, vertical-orientated and plaids). The standard error bars of the means were shown in all conditions.

6.1.3 Discussion

One sample t-test showed the amount of aftereffects at phase randomly changing condition were not zero, suggesting the phase-independent of depth adaptation. In addition, when adapting to the large amplitude adaptors, the amount of the aftereffects at phase randomly-changing condition were significantly smaller than the value at phase static condition, which suggested the phase dependent of depth adaptation. Since the randomly-changed phases caused the random retinal position of the adaptation stimuli, thus it could be said that both retinal-dependent and independent of depth adaptation were involved in this experiment.

Compared with previous related studies about the phase-dependency and independency, this experiment had some differences. In the reports from Graham and Rogers (1982) and Ryan and Gillam (1993), participants moved their gaze to track or scan on the stimuli. In contrast, in this experiment participants focused on the central fixation through the whole trial, which avoided the eye movement. In the report from Graham and Rogers (1983), they manipulated the phase of the adaptation stimulus reversed (180°) every few seconds, probed with test stimulus in phase 0°, 90°, 180°, 270° conditions, and got the distribution curves of depth aftereffects. Since the adapting phase were reversed with several seconds time interval, the adaptation might be involved in both phase dependent and independent processing. However, in our experiment, the phase of the adaptors changed every 0.5 s and was presented evenly at all locations in the phase randomly-changing condition, so the phase dependent processing was excluded or in very limited range. Taya et al. (2005) demonstrated the retinal independent depth aftereffects by presenting the adaptation and test stimuli at overlapped and separated retinal-position conditions, and found the existence of aftereffects at the separated-position condition. In our experiment, the adaptation and test stimuli were always presented at the same retinal position, which cancelled out the influence of the different retinal

position of adaptation stimuli.

Moreover, in this experiment, not only the horizontal-orientated corrugations, but also the vertical ones and plaids were used. By using the vertical-orientated pairs, we verified whether any anisotropy of depth adaptation was involved in both phase dependency and independency conditions. By using the plaid pairs, we examined whether the combination of two corrugations could cause different level of depth adaptation, since the combining processing of visual perception might involve higher stage processing. The results showed no significant difference between the horizontal and vertical orientated corrugations, suggesting the isotropy of depth adaptation. The results also showed the plaid stimuli had the similar level of phase-dependency and independency as their two components in the normalized amplitude condition. Since the plaid stimuli had doubled peak-to-trough amplitude as that of the horizontal and vertical-orientated corrugations, if we showed the PSE shift results against the absolute amplitude of test stimulus, the PSE shifts of plaids would be doubled. The doubled PSE shifts might be due to the doubled interval between the adaptation and test stimuli. However, whether the depth adaptation of plaid stimuli caused higher stages of visual processing was still not clear in the experiment 3.1.

In this experiment, in both phase static and randomly-changing conditions, there were significant differences caused by the adapting amplitude in three stimulus types, suggesting the amplitude of the adaptor was an important factor for depth adaptation. However, the adaptor-probe pairs were the same shape, whether any cross-orientation or cross-shape of depth adaptation were involved in the relatively higher stages than the retinal level adaptation were still unclear. To have further investigations on above issues, in experiment 3.2 we investigated the orientation-independency of depth adaptation by manipulating the adaptation and test corrugations with same and different orientations separately.

6.2 Experiment 3.2: Orientation-independency of depth adaptation

In experiment 3.1, we found no significant difference of depth adaptation among the three types of stimuli pairs in both phase static and randomly-changing conditions. However, our interests were to investigate the adaptation mechanism at relative higher stages than the retinal level processing. Whether any orientation independency of depth adaptation was involved in the relative higher stages was still unclear. Therefore, in experiment 3.2, we examined whether cross-orientation of adaptor-probe pairs could cause different level of depth adaptation. To do so, we manipulated different orientated adaptor-probe pairs, and compared the aftereffects with that of the same orientated pairs. To exclude the local retinal level adaptation, we implemented the adaptation stimuli in phase randomly-changing condition.

6.2.1 Method

6.2.1.1 Participants and apparatus

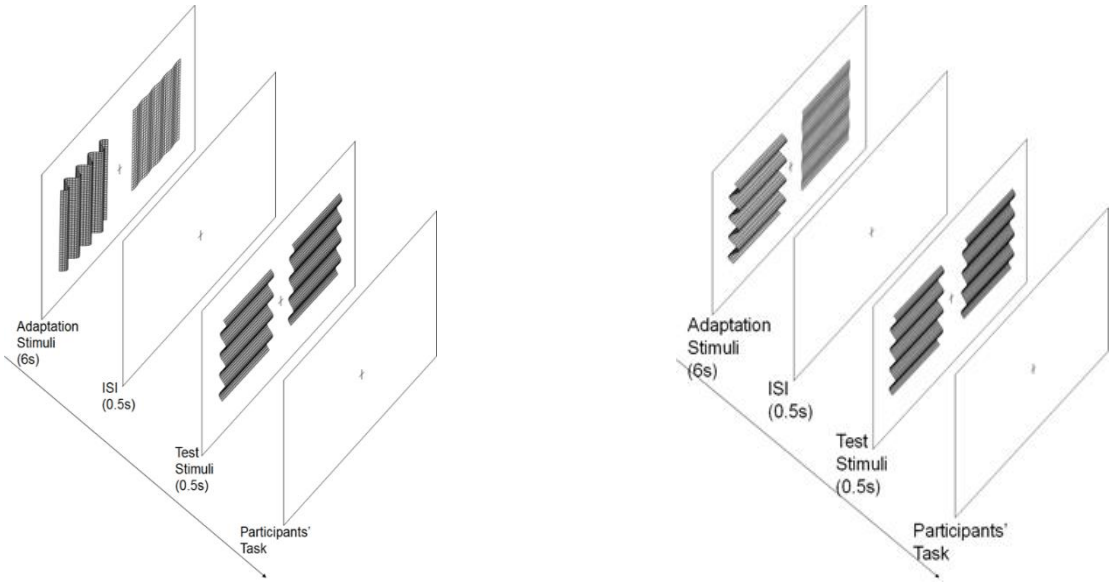
Nine of the ten participants in experiment 3.1 (3 males, mean age: 26.3 years) joined in this experiment. The one did not join in this experiment because of time conflict. The apparatus was the same as in experiment 3.1.

6.2.1.2 Stimuli and procedure

By presenting the adaptation stimuli in phase randomly-changing condition as experiment 3.1, we examined the aftereffects that were adapted to vertical-orientated corrugations but probed by horizontal-oriented corrugations, and compared the value of the horizontal-horizontal pair. We also compared the depth aftereffects between the horizontal-vertical and vertical-vertical corrugations of adaptor-probe pairs (Figure 6.4). The parameters for both horizontal and

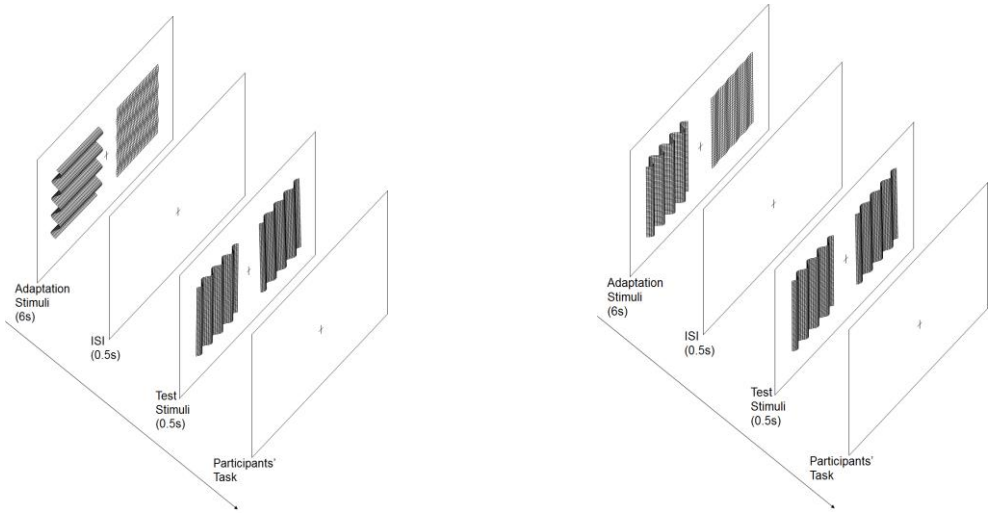
vertical-orientated corrugations were the same as the sinusoidal depth amplitude as shown in table 6.1.

In each condition, there were 216 trials to produce 8 repeats at each test point. The time duration and procedure were the same as experiment 3.1. The experiment was block designed and each session was implemented at different day.



(i) Vertical -> horizontal (ii) Horizontal -> horizontal

(a) Horizontal-orientated test stimulus condition



(i) Horizontal -> vertical (ii) Vertical -> vertical

(b) Vertical-orientated test stimulus condition

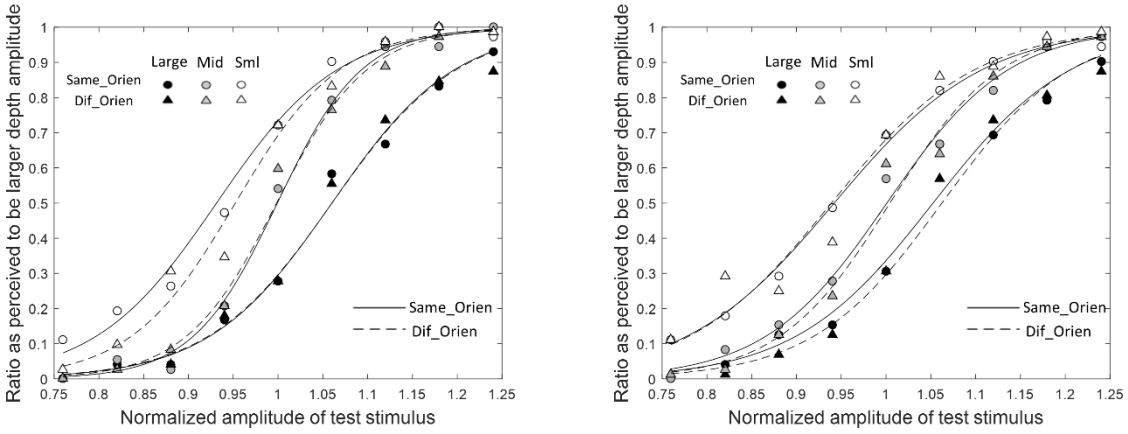
Fig 6.4. Schematic procedure of orientation-independent depth adaptation experiment. (a) The schematic of horizontal-orientated test stimulus condition, which included the vertical-horizontal and horizontal-horizontal

adaptor-probe pairs. (b) The schematic of vertical-orientated test stimulus condition, which included horizontal-vertical and vertical-vertical adaptor-probe pairs.

6.2.2 Results

In experiment 3.2, stimuli were presented with four combinations of adaptor-probe pairs (horizontal-horizontal and vertical-horizontal conditions, vertical-vertical and horizontal-vertical conditions) and three adaptation-amplitude types (large, middle and small adapting conditions). We also fitted the psychometric sigmoidal curves (Figure 6.5) with nine participants' average data and calculated the PSE shift by using the PSE value subtracting 1 for each participant as in experiment 3.1.

Figure 6.5 (a) represented the horizontal-orientated test stimulus condition and (b) the vertical-orientated test stimulus condition. In both figures, the horizontal axis showed the normalized amplitude of test stimulus, and the vertical axis showed the ratio as perceived to be larger depth amplitude. In contrast with these curves, we could see the sigmoidal curves showed obvious shifts caused by the adapting amplitude type, while little shifts related to the same and different-orientated of adaptor-probe pairs.



(a) Horizontal-oriented test stimulus condition (b) Vertical-oriented test stimulus condition

Fig 6.5. Psychometric sigmoidal curves in horizontal- and vertical-oriented test stimulus conditions with nine participants' average data. (a) Horizontal-orientated test stimulus condition; (b) Vertical-orientated test stimulus condition. In both panels, the average values of the "ratio as perceived to be larger depth amplitude" after adapting to large-, middle- and small-amplitude adaptors were drawn as solid, gray and open symbols

respectively. The same and different orientated adaptor-probe pairs were shown in circles and triangles respectively. The fitted curves of the same- and different-orientated adaptor-probe pairs were drawn as solid and dash lines respectively. In panel (a), the same-orientated and different-orientated pairs were horizontal-horizontal and vertical-horizontal pairs respectively. In panel (b), the same-orientated and different-orientated pairs were vertical-vertical and horizontal-vertical pairs respectively.

A two-factor repeated measures ANOVA (4 x 3) was used to analyze the PSE shifts of the four combinations with each participant's data. The results showed a significant difference of the adaptation amplitude type ($F(2, 16) = 25.63$, $p = 0.00$, generalized $\eta^2 = 0.68$), while no significant difference of four combined pairs ($F(3, 24) = 0.42$, $p = 0.74$, generalized $\eta^2 = 0.01$). There were no significant interaction between adaptor-test stimulus pairs and adaptation amplitude types ($F(6, 48) = 0.68$, $p = 0.66$, generalized $\eta^2 = 0.02$) (Figure 6.6).

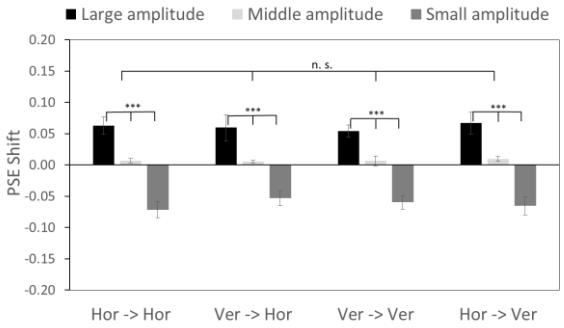


Fig. 6.6. Results of experiment 3.2 with nine participants' data. There was a significant difference of adaptation stimulus amplitude (large, middle and small amplitude adapting conditions), while no significant difference of four adaptor-probe pairs (horizontal- horizontal, vertical-horizontal, vertical- vertical, and horizontal- vertical orientation). The standard error bars were shown in all conditions.

6.2.3 Discussion

Results showed that different orientated adaptor-probe pairs caused similar level of depth aftereffects as the same orientated pairs, suggesting that the depth adaptation was orientation independent.

In previous studies, Tyler (1975) used the disparity-specified sinusoidal corrugations to study about the tilt and size aftereffects. He mainly focused on whether adapting to different orientated corrugations could cause the perceived orientation of the test stimulus changed,

while not the depth changes. However, in our experiment, we examined the perceived depth changes after adapting to different orientated corrugations, which had different purpose as Tyler's study (1975). From the spatial frequency aspect, Serrano-Pedraza et al. (2013) demonstrated that both horizontal and vertical-oriented corrugations had multi-channel mechanisms, while the sensitivity of vertical corrugation was a little lower than the horizontal one, which caused the anisotropy. Witz et al. (2014) examined the disparity thresholds for horizontal and vertical corrugations against spatial frequencies, and got the similar optimal sensitivities for both oriented corrugations. They made a conclusion that similar multi-channel mechanism was involved in both horizontal and vertical corrugations. Moreover, they also claimed that to process the stimuli, both the low-level local disparity detectors and higher-level global detectors were involved, which might have some interactions with each other. In our study, we manipulated the amplitudes of the horizontal and vertical test stimuli at around 12.1 arcmin, which was much higher than the discrimination thresholds of both horizontal and vertical-orientated corrugations. Thus the small anisotropy of discrimination sensitivities revealed between the horizontal and vertical corrugations did not have influence on the depth perception in our experiment.

In our experiment, we manipulated the adaptation stimuli in phase randomly-changing condition, so the local retinal-level depth adaptation was excluded. At the very beginning, the hypothesis was that the similar level of aftereffects in this experiment might be due to the higher stage of global level depth adaptation, which meant that the corrugations were perceived as a whole "object", no matter what the presented orientation was. And the orientation independency might be due to the isotropy of receptive fields of the related detectors. In this case, the amount of the aftereffects was determined by the peak-to-trough amplitudes of adaptation stimuli, while not by the orientation of corrugations. However, the low level disparity-specified mechanism might also explain this result. During the adaptation

period, the randomly changed phases were evenly distributed in all locations of the adaptation stimuli, thus produced similar amounts of disparity, no matter what the orientation was. This similar amounts of disparity induced by the same and different-oriented adaptors caused the similar level of depth aftereffects.

Chapter 7 Experiment 4: Evaluation of disparity- and shape-level depth adaptation

In this chapter, two experiments were conducted. In experiment 4.1, we examined whether the disparity- or the shape-level process is related to the adaptation effects of the corrugated surface. The combinations of disparity-defined horizontal corrugation and plaid surfaces as adaptor-probe pairs were used. We compared the aftereffects between the horizontal-horizontal and plaid-horizontal pairs, and also between the horizontal-plaid and plaid-plaid pairs. The adaptors in the four pairs had the same peak-to-trough amplitudes. In experiment 4.2, we verified whether any different amounts of depth aftereffects were induced between the adaptors with and without certain surfaces by using horizontally oriented corrugation and noise-shape as adaptors, while manipulating the same peak-to-trough amplitudes and the same crossed and uncrossed disparities of the two adaptors. In both experiments, we dynamically changed the phase of the stimuli to prevent local retinal-level depth.

7.1 Experiment 4.1: disparity- or shape-level depth adaptation

7.1.1 Methods

7.1.1.1 Participants

Ten students aged 20-35 years (5 male, mean age: 21.7) from Kochi University of Technology were recruited as participants. All of them had normal or corrected-to-normal vision, and passed the stereo perception and stereo acuity test (less than 1 arcmin) with our own program. Participants were naïve to the aims of the experiments and were compensated for their time. The authors did not serve as participants. All experiments and procedures were approved by

the Research Ethics Committee of Kochi University of Technology and conformed to the tenets of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to experiments.

7.1.1.2 Apparatus

Stimuli were presented on a 22-inch CRT color display (RDF223H; Mitsubishi, Tokyo, Japan) with a 1024 × 768 resolution and 120 Hz frame refresh rate. The luminance of the display was measured using a CS-100A colorimeter (Minolta, Japan) and linearized using a look-up table method. We created a program using Matlab (Mathworks, Natick, MA, USA) with PsychToolbox Version 3 to present the experimental stimuli (Brainard, 1997; Pelli, 1997). During experiments, participants sat in a dark room fronto-parallel to the surface of the display and observed the stimuli via a pair of stereoscopic wireless LCD glasses (NuVision 60GX; MacNaughton, Inc., OR, USA). The refresh rate of the LCD glasses was 120 Hz, so the frame rate was 60 Hz for each eye. No flicker was reported. A chin rest was used to prevent head movement.

7.1.1.3 Stimuli

Random dot stereograms with horizontal disparity were used for the stimuli. Anti-aliased pseudo-random white dots (29.7 cd/m^2) were presented on a gray background (9.9 cd/m^2). The phase and dot patterns of the adaptation stimuli were randomly changed every 200 ms. The density of the dot pattern was 30.6 dots/deg^2.

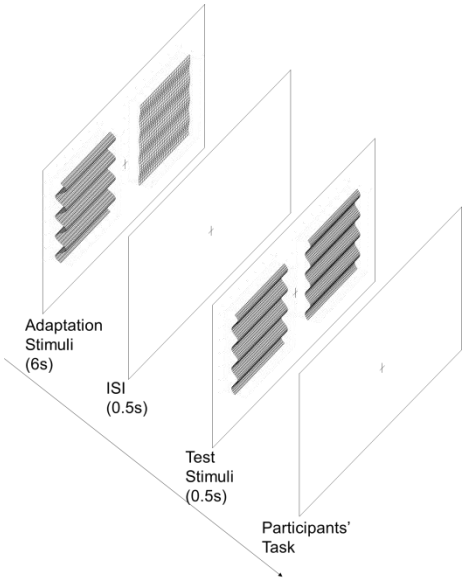
At the center of the display, a nonius fixation with lower part T- and upper part reversed T-shape was shown to the left and right eyes separately. To ensure eye vergence, participants were asked to maintain the vertical lines of the two T parts collinearly and the horizontal lines overlapped during the entire experimental procedure. With the correct vergence, the nonius was perceived as a cross. The lengths of both the horizontal and vertical lines were 1.17 arcdeg.

7.1.1.4 Procedure

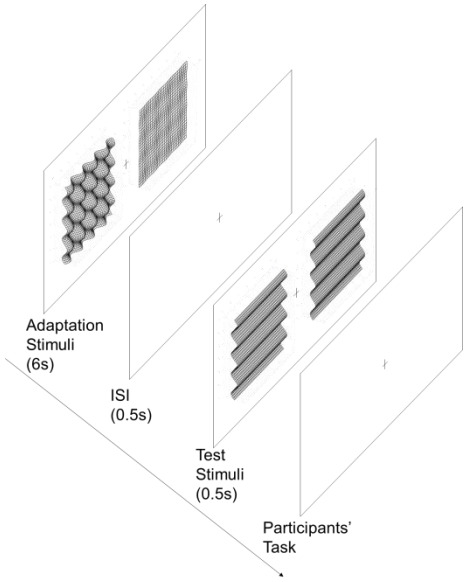
In this experiment, we aimed to investigate whether disparity or shape-level depth adaptation was involved. We used the combinations of horizontally oriented corrugation and plaid as adaptor-probe pairs. Thus, in this experiment, we compared the depth aftereffects between the horizontal - horizontal and plaid - horizontal pairs as the horizontal corrugation test stimulus condition, and between the plaid - plaid and horizontal - plaid pairs as the plaid test stimulus condition.

Both the horizontally oriented corrugation and plaid adaptors had the same peak-to-trough amplitude of disparity, but different distributions of crossed and uncrossed disparity. If the depth aftereffects are more related to shape-level adaptation, the same-shape adaptor-probe pairs would show larger aftereffects than the different-shape pairs. Otherwise, the same-shape adaptor-probe pairs would not always cause larger aftereffects than the different-shape pairs.

To control the peak-to-trough amplitude of the plaid adaptor to be the same as that of the horizontally oriented corrugation adaptor, we defined the amplitudes of the horizontal and vertical corrugated components of the plaids to be half, as the value of the amplitude will be doubled when the two components are linearly added up. The spatial frequency of the sinusoidal corrugations and plaid components was 0.25 cpd. The size of the stimuli at each side was 14.0 × 14.0 arcdeg.

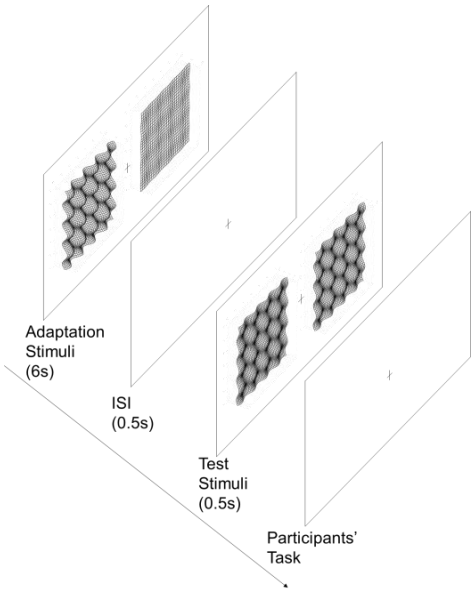


(i) Horizontal -> horizontal

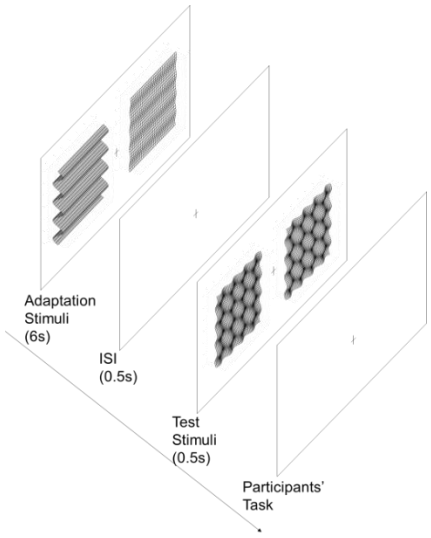


(ii) Plaid -> horizontal

(a) Horizontal-orientated test stimulus condition



(i) Plaid -> plaid



(ii) Horizontal -> plaid

(b) Plaid test stimulus condition

Fig 7.1. Schematic of test procedure in experiment 5.1. (a) Horizontal-orientated test stimulus condition with (i) horizontal-horizontal and (ii) plaid-horizontal adaptor-probe pairs; and (b) plaid test stimulus condition with (i) plaid-plaid and (ii) horizontal-plaid adaptor-probe pairs. In all conditions, both adaptors were presented for 6 s. After a 0.5-s time interval of blank gray background, a test stimulus was presented on one side and a comparison stimulus on the other side for 0.5 s simultaneously. The comparison stimulus had a fixed amplitude (12.1 arcmin), whereas the test stimulus had nine levels of amplitudes (9.1-15.2 arcmin with constant intervals); these stimuli were presented in a random order. The participants' task was to judge which side had the larger amplitude based on their subjective perception, and report their choice by a button press with a two-alternative forced-choice

method. After the participants made their choice, the next trial was triggered automatically.

For each adaptor-probe experimental procedure, we used two adaptors simultaneously positioned at the left and right sides of the CRT display with large-small or middle-middle adaptation amplitude pairs. The left panel of Figure 7.1(a) showed the same shape of horizontal-horizontal as the adaptor-probe pair, with the stimulus amplitudes in the large-small mode. The right panel showed the different shape of plaid-horizontal as the adaptor-probe pair. The left and right panels of Figure 7.1(b) were plaid-plaid and horizontal-plaid pairs, respectively. The sides for the large and small adaptors were presented in a counterbalanced random order. In a middle-middle amplitude-adapting condition, both adaptors had the same amplitude. This was used as a control condition. In all conditions, both adaptors were presented for 6 s. After a 0.5-s time interval of blank gray background, a test stimulus was presented on one side and a comparison stimulus on the other side for 0.5 s simultaneously. The comparison stimulus had a fixed amplitude (12.1 arcmin), whereas the test stimulus had nine levels of amplitudes (9.1-15.2 arcmin with constant intervals); these stimuli were presented in a random order. The positions of the test and comparison stimuli were presented on the left and right sides of the display in a counterbalanced random order. The participants' task was to judge which side had the larger amplitude based on their subjective perception, and report their choice by a button press with a two-alternative forced-choice method. No feedback of correctness was given. After the participants made their choice, the next trial was triggered automatically. The parameters of the horizontally oriented corrugation and plaid stimuli are shown in Table 7.1.

In the large-small adaptor condition, if an adaptation effect is caused by the two adaptors, the one with large amplitude causes the amplitude of the probe to appear smaller than the actual value, whereas the other with small amplitude causes the amplitude of the probe to appear larger. Thus, the points of the subjective equality in large and small adaptor conditions

shift to opposite directions.

Table 7.1: Parameters of the horizontal oriented corrugation and plaid stimuli

Stimulus Type	Depth Amplitude
Adaptation stimulus	
-- Large	20.2 arcmin
-- Middle	12.1 arcmin
-- Small	4.1 arcmin
Test stimulus	9.1-15.2 arcmin (9 levels)
Comparison stimulus	12.1 arcmin

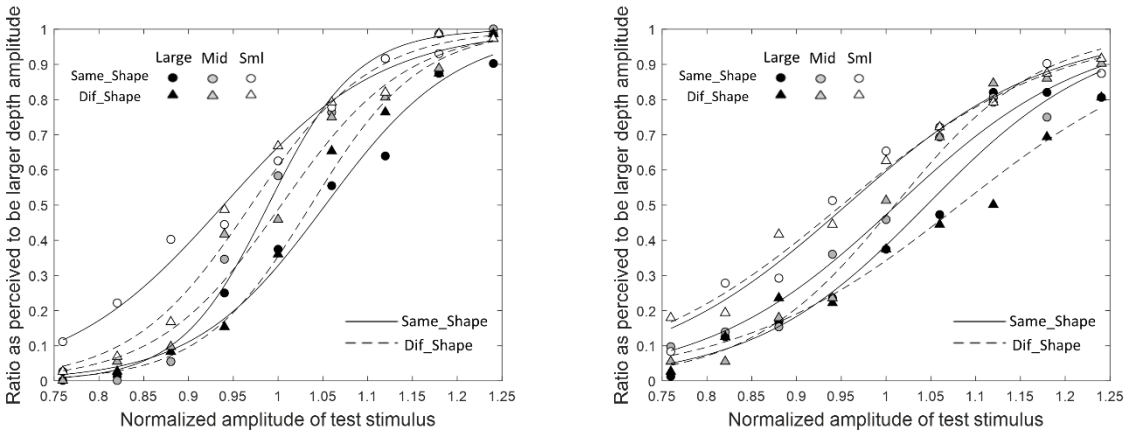
Before the experiment, participants were trained using our own practice program. In the practice trials, the adaptation process was eliminated to avoid any potential influence on the experimental results, and only the test stimuli were shown. Feedback of the correctness was given to participants to enable them to have a basic concept on their perception. There were 36 practice trials for each stimuli type.

During the experiment, stimuli were presented with two combinations of adaptor-probe pairs in each test stimulus condition (i.e., horizontal-horizontal and plaid-horizontal pairs as the horizontally corrugated test stimulus condition, and plaid-plaid and horizontal- plaid pairs as the plaid test stimulus condition) and three adaptation-amplitude types (large, middle, and small amplitude-adapting conditions). For each condition, there were 216 trials to produce 8 repeats at each test point. Different conditions were block designed and divided into six sessions that were implemented at different days. In each session, blocks were presented subsequently with a 2-min break. All the blocks were counterbalanced within- and between-subjects.

7.1.2 Results

Stimuli were presented with four combinations of adaptor-probe pairs and three

adaptation-amplitude types (large, middle, and small amplitude-adapting conditions). The ratio as perceived to be larger depth amplitude was calculated. Figure 7.2 shows the sigmoidal curves as a psychometric function fitted with nine participants' average data by using the generalized linear fitting method (Kingdom and Prins, 2010) with Matlab program (Mathworks, Natick, MA, USA). Data of one male participant were excluded from the analysis owing to his relatively low score of the visual acuity test. Figure 7.2(a) shows the result of horizontally corrugated test stimulus condition, in which the horizontal-horizontal adaptor-probe pair is the same-shape surfaces, and the plaid-horizontal pair is the different-shape surfaces. Figure 7.2(b) shows the plaid test stimulus condition, in which the plaid-plaid adaptor-probe pair as the same-shape surfaces, and the horizontal-plaid pair as the different-shape surfaces. In both figures, the horizontal axis shows the normalized amplitude of test stimulus, and the vertical axis shows the ratio as perceived to be larger depth amplitude. Comparison of these curves showed that shifts are noted among the large, middle, and small adaptor conditions. There are also shifts between the same and different-shape conditions.



(a) Horizontally corrugated test stimulus condition (b) Plaid test stimulus condition

Fig. 7.2: Fitted psychometric sigmoidal curves in horizontally corrugated and plaid test stimulus conditions. (a) Horizontal test stimulus condition, and (b) plaid test stimulus condition. In both panels, the average values of the ratio as perceived to be larger depth amplitude after adapting to large-, middle- and small-amplitude adaptors are shown by solid, gray, and open symbols, respectively. The same and different shapes of adaptor-probe pairs are shown in circles and triangles, respectively. The fitted curves of the same and different shape of adaptor-probe pairs are indicated by solid and dashed lines, respectively. In panel (a), the same and different types of

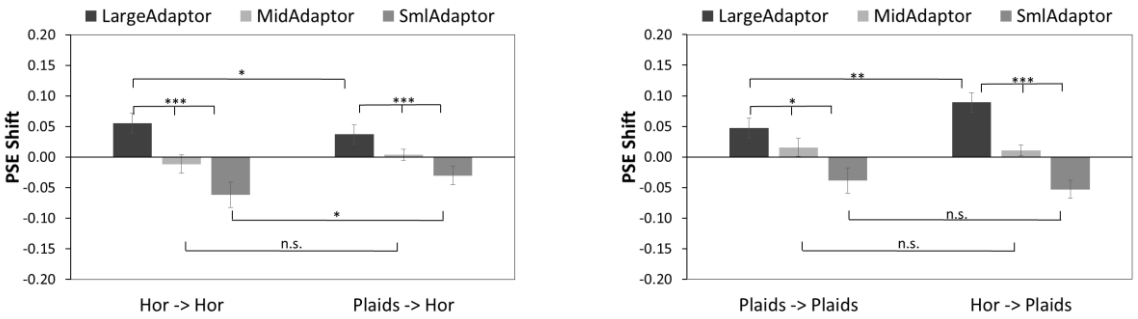
adaptor-probe pairs are horizontal-horizontal and plaid-horizontal, respectively. In panel (b), the same and different types of adaptor-probe pairs are plaid-plaid and horizontal-plaid, respectively.

The 50% point on the fitted psychometric function was calculated for each participant as a point of subjective equality (PSE). The results of PSE in each condition are plotted in Figure 3. A two-factor repeated measures ANOVA (2 x 3) was used to analyze the PSE of the horizontally corrugated and plaid test stimulus conditions separately.

In the horizontally corrugated test stimulus condition (Figure 7.3(a)), ANOVA revealed a significant main effect of adaptor-probe pairs ($F(1,8) = 4.74$, $p = 0.06$, generalized $\eta^2 = 0.03$), amplitude type ($F(2,16) = 19.88$, $p = 0.00$, generalized $\eta^2 = 0.62$), and a significant interaction ($F(2,16) = 5.06$, $p = 0.02$, generalized $\eta^2 = 0.11$). Significant simple main effects of adaptor-probe pairs were noted in large ($F(1,8) = 5.48$, $p = 0.05$, generalized $\eta^2 = 0.11$) and small ($F(1,8) = 6.05$, $p = 0.04$, generalized $\eta^2 = 0.17$) amplitude-adapting conditions. Thus, the absolute values of PSE shift are significantly larger in horizontally corrugated adaptor than in plaid adaptor condition. No significant difference in middle amplitude-adapting condition ($F(1,8) = 2.86$, $p = 0.13$, generalized $\eta^2 = 0.09$). Multiple comparison tests showed significant differences between every pair of amplitude-adapting conditions in horizontal- horizontal condition ($p = 0.00$ between the large and small amplitude-adapting conditions, $p = 0.00$ between the large and middle amplitude-adapting conditions, and $p = 0.02$ between the middle and small amplitude-adapting conditions) and plaid-horizontal condition ($p = 0.00$ between the large and small amplitude-adapting conditions, $p = 0.02$ between the large and middle amplitude-adapting conditions, and $p = 0.04$ between the middle and small amplitude-adapting conditions).

In the plaid test stimulus condition (Figure 7.3(b)), ANOVA revealed a significant main effect of amplitude type ($F(2,16) = 13.56$, $p = 0.00$, generalized $\eta^2 = 0.52$) and a significant interaction ($F(2, 16) = 3.96$, $p = 0.04$, generalized $\eta^2 = 0.07$), whereas no significant

difference in adaptor-probe pairs ($F(1,8) = 1.31$, $p = 0.29$, generalized $\eta^2 = 0.01$). A significant simple main effect of adaptor-probe pairs was found in large condition ($F(1, 8) = 14.48$, $p = 0.01$, generalized $\eta^2 = 0.19$), whereas no significant difference in middle ($F(1, 8) = 0.10$, $p = 0.75$, generalized $\eta^2 = 0.003$) and small amplitude-adapting conditions ($F(1, 8) = 0.52$, $p = 0.49$, generalized $\eta^2 = 0.02$). Thus, the absolute value of PSE shift is significantly larger in horizontally corrugated than in plaid adaptor stimuli. Multiple comparison tests showed significant differences in plaid-plaid condition ($p = 0.02$ between the middle and small amplitude conditions, $p = 0.04$ between the large and small amplitude conditions) and horizontal-plaid condition ($p = 0.00$ between each two amplitude-adapting conditions among the large, middle, and small amplitude adaptors).



(a) Horizontally oriented test stimulus condition (b) Plaid test stimulus condition

Fig. 7.3: PSE shift in experiment 5.1. (a) Horizontal-orientated test stimulus condition. (b) Plaid test stimulus condition. A significant difference of amplitude-adapting (large, middle, and small amplitude-adapting conditions) was found at both adaptation conditions. The standard error bars are shown in all conditions.

7.1.3 Discussion

The results showed significantly different PSE shifts based on the amplitudes of the adaptors at each adaptor-probe pair condition. These results indicate that the disparity-level depth adaptation is involved in both the same- and different-shape adaptor-probe pairs. Based on the model of multi-channel disparity detectors (Cormack et al., 1993; Stevenson et al., 1992) and adaptation mechanism (Carandini, 2000; Georgeson, 2004), the amplitudes of the large and

small adaptors used in this experiment are within the disparity-tuning range; thus, significant difference can be found.

In the horizontally oriented test stimulus condition, the same-shape pair (horizontal-horizontal) caused larger amount of PSE shifts than the different-shape pair (plaid-horizontal). However, in the plaid test stimulus condition, it was the opposite. The same-shape pair (plaid-plaid) caused smaller amount of PSE shifts than the different-shape pair. In our hypothesis, if the depth adaptation is more related to the shape-level adaptation, the same-shape pair will always cause larger amount of aftereffects. The contrary results indicate that other factor(s) but shape-level adaptation is involved. Although the horizontally oriented corrugation and plaid adaptors had the same peak-to-trough amplitude, they had different disparity distributions. The former had about 30% disparity information distributed at the peak area of the 90%-100% amplitude range, whereas the latter only had 5% disparity information. We also separated the crossed and uncrossed disparities and calculated the disparity distributions of the two adaptation stimuli separately. We found that the horizontally oriented corrugations had larger average disparities than the plaid (for both the crossed and uncrossed average disparities, the horizontal adaptation stimulus is about 1.56 times as the plaid one). In the large adaptor condition, the larger average disparities of horizontal corrugation adaptor might induce a larger amount of PSE shifts than the plaid. In addition, as the phase of stimuli was randomly changed, the adaptation effect was not induced by the local process.

Thus, these results suggest that the adaptation in corrugated surfaces was mainly caused by the disparity-level process, which might involve the disparity integration process through some time window since the phase was dynamically changed.

However, whether the disparity-level process is the only factor for depth adaptation or the shape-related process, such as continuous surface, can also enhance the effect of depth

adaptation remains unclear. To further investigate this issue, we used two adaptors in experiment 4.2. Both adaptors had the same peak-to-trough amplitude and the same average disparity distribution. However, one adaptor had a continuous shape structure, and the other did not.

7.2 Experiment 4.2 depth adaptation by using surface and non-surface adaptors

7.2.1 Methods

In experiment 4.1, we found that the depth adaptation of horizontal corrugation might be related to the disparity-level process, because we found that the different shape adaptor-probe pair showed even larger adaptation effects in the plaid test stimulus condition. However, whether the shape-related process, such as continuous surface, is necessary to induce the aftereffect is unknown. In this experiment, we used a new noise-shape adaptor. This adaptor does not have a continuous surface. Thus, it will not involve the higher-order and shape-level process when compared with the horizontal corrugation adaptor. In both conditions, the two adaptors had the same peak-to-trough amplitude and the same amount of disparity distribution.

7.2.1.1 Participants and apparatus

The ten participants who joined in experiment 4.1 took part in this experiment. The apparatus was the same as in experiment 4.1.

7.2.1.2 Stimuli

The noise-shape adaptor had the same disparity information as the adaptor of horizontally oriented corrugation, but distributed in random positions without a continuous surface. The horizontal corrugation had the same parameters as that used in previous experiment. The schematic of the stimuli is shown in Figure 7.4.

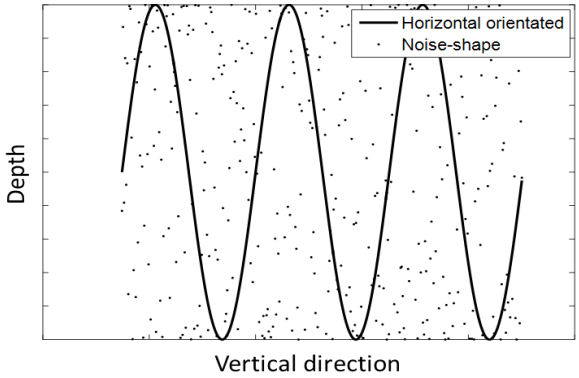
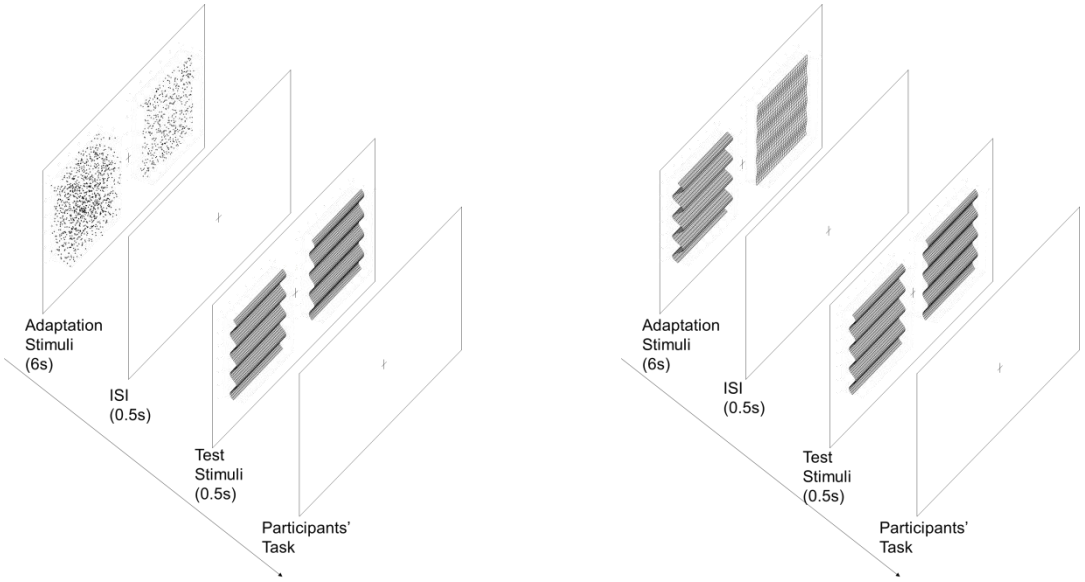


Fig. 7.4: Schematic of the noise-shape and horizontally oriented corrugation adaptors. The horizontally oriented corrugation adaptor was shown in solid line and the noise-shape adaptor was shown in dots.

7.2.1.3 Procedure

The schematic of procedure is shown in Figure 7.5. In each condition, there were 216 trials to produce 8 repeats at each test stimuli. The time duration and the procedure were the same as in experiment 4.1. The participants' task was to judge which side had the larger amplitude based on their subjective perception. The experiment was block designed, and each session was conducted on a different day.



(a) noise-shape-adapting condition (b) horizontal corrugation-adapting condition

Fig. 7.5: Schematic of procedure of experiment 4.2. (a) Noise-shape-adapting condition and (b) horizontal corrugation-adapting condition.

7.2.2 Results

In experiment 4.2, the stimuli were presented with two combinations of adaptor-probe pairs (horizontal-horizontal and noise-shape -horizontal) and three adaptation-amplitude types (large, middle, and small amplitude-adapting conditions).

The psychometric sigmoidal curves were fitted, and PSE values were calculated from the data of the nine participants. We also excluded the tenth participant as described in experiment 4.1. Figure 7.6 shows the fitted psychometric sigmoidal curves from the average data of participants. The horizontal and vertical axes represent the normalized amplitude of test stimulus and the ratio as perceived to be larger depth amplitude, respectively. The sigmoidal curves show shifts caused by the amplitude-adapting types and also shifts between the horizontally oriented condition and noise shape condition.

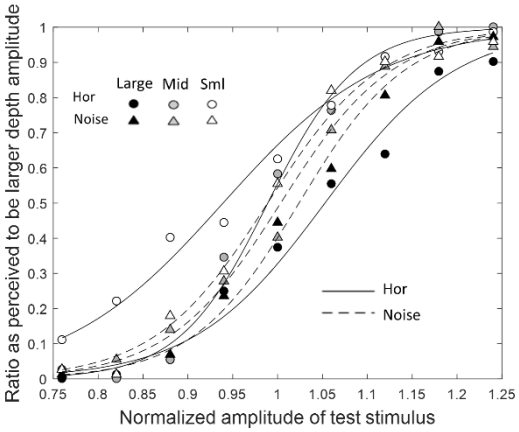


Fig. 7.6: Fitted psychometric sigmoidal curves in horizontally oriented corrugation and noise-shape adaptor conditions. The average values of the ratio as perceived to be larger depth amplitude after adapting to large-, middle-, and small-amplitude adaptors are shown by solid, gray, and open symbols, respectively. The horizontally oriented and noise-shape adaptation conditions were shown in circles and triangles respectively. The fitted curves of the horizontally oriented and noise-shape adaptation conditions are indicated by solid and dashed lines.

A two-factor repeated measures ANOVA (2 x 3) was used to analyze the PSE of the horizontally oriented and noise-shape adaptation conditions. ANOVA revealed a significant main effect of adaptor-probe pairs ($F(1,8) = 5.37$, $p = 0.05$, generalized $\eta^2 = 0.05$), amplitude type ($F(2,16) = 15.62$, $p = 0.00$, generalized $\eta^2 = 0.53$), and a significant interaction ($F(2,16) = 9.26$, $p = 0.00$, generalized $\eta^2 = 0.25$). Significant simple main effects of adaptation amplitude type were observed at both adaptation conditions (horizontal adaptor: $F(2,16) = 17.60$, $p = 0.00$, generalized $\eta^2 = 0.68$; noise-shape adaptor: $F(2,16) = 3.81$, $p = 0.04$, generalized $\eta^2 = 0.26$). Significant simple main effects of depth adaptation were noted at large ($F(1,8) = 11.44$, $p = 0.01$, generalized $\eta^2 = 0.28$) and small ($F(1,8) = 9.55$, $p = 0.01$, generalized $\eta^2 = 0.32$) amplitude-adapting conditions, whereas no significant difference at middle ($F(1,8) = 3.34$, $p = 0.10$, generalized $\eta^2 = 0.14$) amplitude-adapting condition (Figure 7.7).

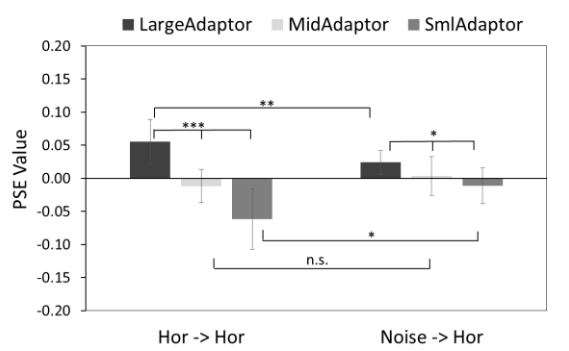


Fig. 7.7: Results of experiment 4.2. Significant differences of adaptation amplitude (large, middle, and small amplitude-adapting conditions) were found at both adaptation conditions. Significant differences of depth adaptation were also noted between the two adaptation conditions in both large and small amplitude-adapting conditions, but not in middle amplitude-adapting condition.

7.2.3 Discussion

The PSE shift was significantly different between the large and small adaptors in both the horizontally oriented adaptation condition and the noise-shape condition, which indicates the disparity-level depth adaptation. Moreover, the difference in PSE shifts between these two conditions was also significant. Since the peak-to-trough amplitudes and the crossed and uncrossed average disparity were the same in two conditions, the significant result suggests the importance of surface structure in depth adaptation. If the adaptor does not have a specific continuous surface, the adaptation effect might be limited.

7.3 General discussion

We conducted two experiments to investigate the mechanisms of depth adaptation of the disparity-specified sinusoidal corrugations. We aimed to examine whether the disparity or the shape-level process was involved in depth adaptation. In both experiments, we used different types of adaptors with the same peak-to-trough amplitude. We obtained different amounts of aftereffects, which revealed the process of depth adaptation of corrugated surfaces as follows.

In experiment 4.1, we examined the depth aftereffects of the same- and different-shape adaptor-probe pairs at horizontally oriented and plaid test stimulus conditions separately and found significant differences in both conditions. In the horizontally oriented test stimulus condition, the same-shape pair (horizontal-horizontal) caused larger amount of aftereffects than the different-shape pair (plaid-horizontal). However, the results were contrary in the plaid test stimulus condition. The different-shape pair (horizontal-plaid) caused larger amount of aftereffects than the same-shape pair at the large amplitude-adapting condition. Although we controlled the same peak-to-trough amplitudes of the adaptors in the same- and different-shape pairs, their disparity distributions were different. The horizontal corrugated adaptor had larger amount of average disparity, thus caused larger amount of aftereffects. Moreover, if the shape-level depth adaptation was involved, the different-shape pair should cause smaller aftereffects than the same-shape pair, which would be in an opposite direction. However, we still found the larger aftereffects, suggesting that the disparity-level depth adaptation is more involved.

In experiment 4.2, we used the randomly distributed noise dots as the adaptor and compared the amount of aftereffects with that caused by the horizontal corrugated adaptor. To eliminate the influence caused by the difference of the average disparities as in experiment 5.1, we manipulated the noise-shape adaptor to have both the same peak-to-trough amplitude and the same disparity distribution as the horizontal corrugated stimulus, while no continuous surface. The results showed that the non-structured noise-shape adaptor caused smaller aftereffects, suggesting the importance of surface structure in the depth adaptation process. In this process, a certain surface structure at the peak area might provide more related disparity information when compared with the disparity information at the same area of the noise-shape adaptor. Consequently, this more related disparity information might cause a significantly larger adaptation effect in the corrugated adaptor than in the noise-shape adaptor. In the brain

imaging study, Tsao et al. (2003) reported that the brain areas of V3A, V7, and V4d topology and the caudal parietal disparity region (CPDR) are strongly activated in the near/far judgment when compared with zero disparity. They found that these brain areas are responsible for the 3D structure. Ban and Welchman (2015) reported that V3A is responsible for capturing the 3D structure of the surrounding surfaces. Based on these previous studies, the depth adaptation in our study might also involve these brain areas, which indicates the relatively early level of the disparity perception and middle level (like V3A area) of the surface structure process.

Although we found that both disparity and surface structure were important for depth adaptation, the influence was not analyzed quantitatively in this study. Moreover, we separated the adaptor-probe pairs as the same and different shapes, but the horizontal corrugation was the component of the plaid. Thus, the shape-level depth adaptation might also be involved in the process. These factors will be examined in the future.

In summary, with regard to the disparity-defined corrugated stimuli, the larger average disparity distribution caused the larger amount of depth aftereffects. This finding suggests that the disparity-level depth adaptation is more involved than the shape-level adaptation. Meanwhile, the continuous surface of horizontal corrugation adaptor caused the larger amount of depth adaptation than the noise-shape condition, suggesting the importance of surface structure in the depth adaptation.

Chapter 8 General discussion

In this dissertation, two studies were implemented by using the dichoptic stimuli and adaptation paradigm to investigate the different stages of human 3D visual processing. The conclusion can be summarized as follows.

8.1 Dichoptic study

Based on the efficient coding theory (Li and Atick, 1994), Zhaoping (2017) investigated binocular summation (S_+) and binocular suppression (S_-) percepts at central and peripheral view conditions, and reported that the bias towards binocular summation (S_+) percept at central vision might be due to the high level top-down feedback, whereas this feedback was weaker or absent at peripheral vision.

In this study, we continued the previous study (Zhaoping, 2017) and had further investigations on following questions: (1) whether the vergence eye movement is involved in the process? (2) Since the two eyes' inputs are dichoptic stimuli which may cause the rivalry, how are the temporal dynamics of S_+ and S_- percepts? To this end, two experiments were implemented.

Firstly, the vergence eye movements were investigated by using the short duration dichoptic stimuli. The results showed that there was obvious change in vertical vergence at the central view condition, whereas no such tendency at the peripheral view condition. Matched sample t-test showed there was a significant difference of change in vertical vergence between the two view conditions at around $t > 700$ ms after the dichoptic stimuli onset. The stimuli used in this experiment were the summation or differencing of two horizontal gratings, which had independent random phases of each grating. To perceive the motion direction, the left and

right eyes needed to integrate the ambiguous gratings into S_+ and S_- channels based on the top-down feedback. The integration might cause the change in vertical vergence. The significant difference had longer latency after the dichoptic stimuli onset, which might be because that multi-cycles of feedforward-feedback-verify-weight (FFVW) processes are involved before participants made their choices.

Secondly, temporal dynamics of S_+ and S_- percepts were investigated by using the long duration dichoptic stimuli. The results show that: (1) the change in horizontal vergence is involved at the central vision when percept changing from S_+ to S_- and also from S_- to S_+, which might indicate the involvement of the high level visual attention; (2) the temporal dynamics of the S_+ and S_- percepts at the central view condition show a tendency of competition between the S_+ and S_-. Since the visual recognition (S_+ or S_- percepts) at the central vision involves the feedback from the higher brain areas, the temporal dynamics of the S_+ and S_- percepts might indicate the hierarchy of the binocular rivalry, which is different from the direct competition between the visual inputs in the conventional binocular rivalry; (3) the neither percept at the peripheral view condition is significantly larger than that at the central view condition, suggesting the different underlying mechanisms, in which the former could tolerate for the binocular conflicts, whereas the latter could not tolerate for the binocular uncorrelated information and works in “winner-take-all” manner; (4) the temporal dynamics of the S_+, S_- and neither percepts and the quantities of percept flips might indicate the involvement of adaptation; (5) the change in pupil size can be explained by the existence of the LC-NE complex, which might indicate the involvement of the visual attention when percept flips.

8.2 Depth adaptation study

In this study, we conducted two experiments to investigate the mechanisms of depth

adaptation of the disparity-specified sinusoidal corrugations. We aimed to examine whether the disparity or the shape-level process was involved in depth adaptation. In both experiments, we used different types of adaptors with the same peak-to-trough amplitude and obtained different amounts of aftereffects. The results revealed the process of depth adaptation of corrugated surfaces as follows.

Firstly, regarding to the disparity-defined corrugated stimuli, both the phase-dependency and independency were involved in the depth adaptation. By using the same- and different-oriented adaptor-probe pairs in phase randomly changing condition, we found the orientation independency of depth adaptation.

Secondly, regarding to the disparity-defined corrugated stimuli, the larger average disparity distribution caused the larger amount of depth aftereffects. This finding suggests that the disparity-level depth adaptation is more involved than the shape-level adaptation. Meanwhile, the continuous surface of horizontal corrugation adaptor caused the larger amount of depth adaptation than the noise-shape condition, suggesting the importance of surface structure in the depth adaptation. In this process, a certain surface structure at the peak area might provide more related disparity information when compared with the disparity information at the same area of the noise-shape adaptor. Consequently, this more related disparity information might cause a significantly larger adaptation effect in the corrugated adaptor than in the noise-shape adaptor. Physiologically, previous studies (Tsao et al., 2003; Ban and Welchman, 2015) reported that the V3A, V4d were responsible for capturing the 3D structure of the surrounding surfaces, or were strongly activated in the near/far judgment when compared with zero disparity. Based on these previous studies, the depth adaptation in our study might also involve these brain areas, which indicates the relatively early level of the disparity perception and middle level of the surface structure process.

Thirdly, the limitations of this study were, although we found that both disparity and

surface structure were important for depth adaptation, the influence was not quantitatively analyzed in this study. Moreover, we separated the adaptor-probe pairs as the same and different shapes, but the horizontal corrugation was a component of the plaid. Thus, the shape-level depth adaptation might also be involved in the process. These factors will be examined in the future.

References

Akao, T., Mustari, M. J., Fukushima, J., Kurkin, S., & Fukushima, K. (2005). Discharge characteristics of pursuit neurons in mst during vergence eye movements. *Journal of Neurophysiology, 93*(5), 2415- 2434.

Baker, R., & Highstein, S. M. (1975). Physiological identification of interneurons and motoneurons in the abducens nucleus. *Brain Research, 91*(2), 292-298.

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision, 9*(12), 13.1-18.

Ban, H., & Welchman, A. E. (2015). Fmri analysis-by-synthesis reveals a dorsal hierarchy that extracts surface slant. *Journal of Neuroscience the Official Journal of the Society for Neuroscience, 35*(27), 9823-9835.

Bando, T., Hara, N., Takagi, M., Yamamoto, K., & Toda, H. (1996). Roles of the lateral suprasylvian cortex in convergence eye movement in cats. *Progress in Brain Research, 112(08),* 143-156.

Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. Physiological laboratory, *Cambridge Universitiy.* 217-234.

Berends, E. M., & Erkelens, C. J. (2001). Adaptation to disparity but not to perceived depth. *Vision Research, 41*(7), 883-892.

Berends, E. M., Liu, B., & Schor, C. M. (2005). Stereo-slant adaptation is high level and does not involve disparity coding. *Journal of Vision, 5*, 71–80.

Bowd, C., Donnelly, M., Shorter, S., & Patterson, R. (2000). Cross-domain adaptation reveals that a common mechanism computes stereoscopic (cyclopean) and luminance plaid motion. *Vision Research, 40*, 331–339.

Blake, R. and Logothetis, N. (2002). Visual competition. *Nature Reviews Neuroscience*, 3(1),

pp.13-21.

Bradshaw, M. F., & Rogers, B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Research, 39,* 3049-3056.

Brascamp, J., van Ee, R., Noest, A., Jacobs, R. and van den Berg, A. (2006). The time course of binocular rivalry reveals a fundamental role of noise. *Journal of Vision*, 6(11), pp.1244-1256.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Carandini, M. (2000). Visual cortex: fatigue and adaptation. *Current Biology Cb, 10*(16), 1-3.

Chaturvedi, V. and Van Gisbergen, J. (2000). Stimulation in the rostral pole of monkey superior colliculus: effects on vergence eye movements. *Experimental Brain Research*, 132(1), pp.72-78.

Chen, X. and He, S. (2003). Temporal characteristics of binocular rivalry: visual field asymmetries. *Vision Research*, 43(21), pp.2207-2212.

Chen, Y. C., Maurer, D., Lewis, T. L., Spence, C., & Shore, D. I. (2017). Central-peripheral differences in audiovisual and visuotactile event perception. *Attention Perception & Psychophysics*(4), 79:2552–2563.

Cherniawsky, A. S., & Mullen, K. T. (2016). The whole is other than the sum: perceived contrast summation within color and luminance plaids. *i-perception, September-October*, 1-14.

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1993). Disparity-tuned channels of the human visual system. *Visual Neuroscience, 10,* 585–596.

Cumming, B. G., & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature,389*(6648), 280-283.

Cumming, B. and Parker, A. (2000). Local Disparity Not Perceived Depth Is Signaled by Binocular Neurons in Cortical Area V1 of the Macaque. *The Journal of Neuroscience,*

20(12), pp.4758–4767.

Daniel PM, Whitteridge D (1961) The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology, 159,* 203-21.

Dayan, P. (1998). A Hierarchical Model of Binocular Rivalry. *Neural Computation*, 10(5), pp.1119-1135.

Domini, F., Adams, W., & Banks, M. S. (2001). 3D after-effects are due to shape and not disparity adaptation. *Vision Research, 41*, 2733–2739.

Duke, P. A., & Wilcox, L. M. (2003). Adaptation to vertical disparity induced-depth: implications for disparity processing. *Vision Research, 43*, 135-147.

Einhäuser, W., Stout, J., Koch, C. and Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences*, 105(5), pp.1704-1709.

Fukushima, K., Yamanobe, T., Shinmei, Y., Fukushima, J., Kurkin, S., & Peterson, B. W. (2002). Coding of smooth eye movements in three-dimensional space by frontal cortex. *Nature, 419*(6903), 157-162.

Gamlin, P. D. R. (2002). Neural mechanisms for the control of vergence eye movements. *Annals of the New York Academy of Sciences*, *956*(1), 264-272.

Gamlin, P. D., & Clarke, R. J. (1995). Single-unit activity in the primate nucleus reticularis tegmenti pontis related to vergence and ocular accommodation. *Journal of Neurophysiology, 73*(5), 2115-2119.

Gamlin, P. D., & Yoon, K. (2000). An area for vergence eye movement in primate frontal cortex. *Nature, 407*(6807), 1003-1007.

Gamlin, P. D., Yoon, K., & Zhang, H. (1996). The role of cerebro-ponto-cerebellar pathways in the control of vergence eye movements. *Eye, 10 (Pt 2)*(2), 167-171.

Georgeson, M. A., & Shackleton, T. M. (1994). Perceived contrast of gratings and plaids: non-linear summation across oriented filters. *Vision Research, 34*(8), 1061-1075.

Georgeson, M. (2004). Visual aftereffects: cortical neurons change their tune. *Current Biology, 14*(18), R751-R753.

Gnadt, J. W., & Mays, L. E. (1995). Neurons in monkey parietal area lip are tuned for eye-movement parameters in three-dimensional space. *Journal of Neurophysiology, 73*(1), 280-297.

Graham, M., & Rogers, B. (1982). Simultaneous and successive contrast effects in the perception of depth from motion-parallax and stereoscopic information. *Perception, 11*(3), 247-262.

Graham, M. E., & Rogers, B. J. (1983). Phase-dependent and phase-independent depth aftereffects. *Perception, 12*(Abs), A16.

Hayashi, R., Miura, K., Tabata, H., & Kawano, K. (2008). Eye movements in response to dichoptic motion: evidence for a parallel-hierarchical structure of visual motion processing in primates. *Journal of Neurophysiology, 99*(5), 2329-2346.

He, S., Carlson, T. and Chen, X. (2005). Parallel pathways and temporal dynamics in binocular rivalry. *Binocular rivalry*, the MIT press, pp.81-97.

Hegdé, J., & Van Essen, D. C. (2005). Role of primate visual area v4 in the processing of 3-d shape characteristics defined by disparity. *Journal of Neurophysiology, 94*(4), 2856-2866.

Hibbard, P. B., & Langley, K. (1998). Plaid slant and inclination thresholds can be predicted from components. *Vision Research, 38*, 1073–1084.

Hoffman, J. (1998). Visual attention and eye movements. *In H. Pashler (Ed.), Attention* (pp. 119-153). Psychology Press.

Hollins, M. and Hudnell, K. (1980). Adaptation of the binocular rivalry mechanism. *Invest*

Ophthalmol Vis Sci, 19 (9), pp.1117-1120.

Howard, I. P., & Rogers, B. J. (2012). Perceiving in depth, Volume 2: stereoscopic vision. New York, NY: Oxford University Press.

Judge, S. J., & Cumming, B. G. (1986). Neurons in the monkey midbrain with activity related to vergence eye movement and accommodation. *Journal of Neurophysiology, 55*(5), 915-930.

Kang, M. and Blake, R. (2010). What causes alternations in dominance during binocular rivalry?. *Attention, Perception, & Psychophysics*, 72(1), pp.179-186.

Kingdom, F. A. A., & Prins, N. (2010). Psychophysics: a practical introduction. London, UK: Academic Press.

Leopold, D. and Logothetis, N. (1996). Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature*, 379(6565), pp.549-553.

Levi, D. M. (2008). Crowding - an essential bottleneck for object recognition: a mini-review. *Vision Research, 48*(5), 635-654.

Li, Z., & Atick, J. J. (1994). Efficient stereo coding in the multiscale representation. *Network Computation in Neural Systems, 5*(2), 157-174.

Liu, L., Tyler, C. and Schor, C. (1992). Failure of rivalry at low contrast: Evidence of a suprathreshold binocular summation process. *Vision Research*, 32(8), pp.1471-1479.

Logothetis, N., Theodoni, P., Panagiotaropoulos, T. I., Kapoor, V., Logothetis, N. K., & Deco, G. (2011). Cortical microcircuit dynamics mediating binocular rivalry: The role of adaptation in inhibition. *Frontiers in Human Neuroscience*. DOI: 10.3389/fnhum.2011.00145.

Masson, G., Busettini, C. and Miles, F. (1997). Vergence eye movements in response to binocular disparity without depth perception. *Nature*, 389(6648), pp.283-286.

Mays, L. E. (1984). Neural control of vergence eye movements: convergence and divergence

neurons in midbrain. *Journal of Neurophysiology, 51*(5), 1091-1108.

Mays, L. E., Porter, J. D., Gamlin, P. D., & Tello, C. A. (1986). Neural control of vergence eye movements: neurons encoding vergence velocity. *Journal of Neurophysiology, 56*(4), 1007-1021.

May, K. A., & Zhaoping, L. (2016). Efficient coding theory predicts a tilt aftereffect from viewing untilted patterns. *Current Biology, 26*(12), 1571-1576.

May, Keith , Zhaoping, Li, Hibbard, & Paul . (2012). Perceived direction of motion determined by adaptation to static binocular images. *Current Biology, 22*(1), 28-32.

Minini, L., & Parker, A. J. (2010). Neural modulation by binocular disparity greatest in human dorsal visual stream. *Journal of Neurophysiology,104*(1), 169-178.

Moreno-Bote, R., Rinzel, J. and Rubin, N. (2007). Noise-Induced Alternations in an Attractor Network Model of Perceptual Bistability. *Journal of Neurophysiology,* 98(3), pp.1125-1139.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Poggio, G. E. (1995). Mechanisms of stereopsis in monkey visual cortex. *Cerebral Cortex, 5*(3), 193-204.

Poggio, G. and Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *Journal of Neurophysiology,* 40(6), pp.1392-1405.

Rambold, H. A., Sheliga, B. M., & Miles, F. A. (2010). Evidence from vergence eye movements that disparities defined by luminance and contrast are sensed by independent mechanisms. *Journal of Vision, 10*(14), 1-34.

Rogers, B., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22,* 261-270.

Rogers, B., & Graham, M. (1985). *Motion Parallax and the Perception of Three-Dimensional Surfaces*. *Brain Mechanisms and Spatial Vision*. 95-111.

Rosenholtz, R. (2016). Capabilities and limitations of peripheral vision. *Annu Rev Vis Sci, 2*(1), 437-457.

Roumani, D. and Moutoussis, K. (2012). Binocular rivalry alternations and their relation to visual adaptation. *Frontiers in Human Neuroscience,* 6.

Ryan, C., & Gillam, B. (1993). A proximity-contingent stereoscopic depth aftereffect: evidence for adaptation to disparity gradients. *Perception, 22*(4), 403-418.

Schafer, R. and Moore, T. (2007). Attention Governs Action in the Primate Frontal Eye Field. *Neuron,* 56(3), pp.541-551.

Schumer, R., & Ganz, L. (1979). Independent stereoscopic channels for different extents of spatial pooling. *Vision Research, 19*, 1203–1314.

Scocchia, L., Valsecchi, M. and Triesch, J. (2014). Top-down influences on ambiguous perception: the role of stable and transient states of the observer. *Frontiers in Human Neuroscience,* 8.

Serrano-Pedraza, I., Brash, C. & Read, J. C. A. (2013). Testing the horizontal-vertical stereo anisotropy with the critical-band masking paradigm. *Journal of Vision, 13(11):15,* 1-15.

Shpiro, A., Moreno-Bote, R., Rubin, N. and Rinzel, J. (2009). Balance between noise and adaptation in competition models of perceptual bistability. *Journal of Computational Neuroscience,* 27(1), pp.37-54.

Snowden, R. J., Thompson, P., & Troscianko, T. (2012). *Basic vision: an introduction to visual perception*. Oxford University Press.

Solé, P. M., Pérez, Z. L., Aznar-Casanova, J. A., & Supèr, H. (2013a). A role of eye vergence in covert attention. *Plos One, 8*(1), e52955.

Solé, P. M., Puigcerver, L., Aznar-Casanova, J. A., & Supèr, H. (2013b). Difference in visual

processing assessed by eye vergence movements. *Plos One, 8*(9), e72041.

Solé, P. M., Romeo, A., Cañete, C. J., & Supèr, H. (2017). Eye vergence responses during a visual memory task. *Neuroreport, 28*(3), 123-127.

Stevenson, S. B., Cormack, L. K., Schor, C. M., & Tyler, C. W. (1992). Disparity tuning in mechanisms of human stereopsis. *Vision Research, 32*(9), 1685-1694.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: a review. *Journal of Vision, 11*(5), 13, 1–82.

Sylvestre, P. A., Choi, J. T., & Cullen, K. E. (2003). Discharge dynamics of oculomotor neural integrator neurons during conjugate and disjunctive saccades and fixation. *Journal of Neurophysiology*, *90*(2), 739-754.

Taya, S., Sato, M. & Nakamizo, S. (2005). Stereoscopic depth aftereffects without retinal position correspondence between adaptation and test stimuli. *Vision Research, 45,* 1857–1866.

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in v2. *Nature Neuroscience, 5*(5), 472-478.

Toda, H., Tanimoto, N., Takagi, M., Abe, H., & Bando, T. (2006). Visual cortical contribution to open-loop and feed-back control of convergence eye movements in the cat. *Neuroscience Research, 54*(4), 302-312.

Toda, H., Yoshizawa, T., Takagi, M., & Bando, T. (2001). The properties of convergence eye movements evoked from the rostral and caudal lateral suprasylvian cortex in the cat. *Neuroscience Research, 39*(3), 359-367.

Tsao, D. Y., Vanduffel, W. Y., Fize, D., Knutsen, T. A., Mandeville, J. B., & Wald, L. L., et al. (2003). Stereopsis activates v3a and caudal intraparietal areas in macaques and humans. *Neuron*, *39*(3), 555-568.

Tyler, C. W. (1975). Stereoscopic tilt and size aftereffects. *Perception, 4*, 187-192.

van Dam, L. and van Ee, R. (2006a). The role of saccades in exerting voluntary control in perceptual and binocular rivalry. *Vision Research*, 46(6-7), pp.787-799.

van Dam, L. and van Ee, R. (2006b). Retinal image shifts, but not eye movements per se, cause alternations in awareness during binocular rivalry. *Journal of Vision*, 6(11), pp.1172-1179.

van Horn, M., Waitzman, D. and Cullen, K. (2013). Vergence Neurons Identified in the Rostral Superior Colliculus Code Smooth Eye Movements in 3D Space. *Journal of Neuroscience,* 33(17), pp.7274-7284

Welchman, A. E. (2016). The human brain in depth: how we see in 3D. *Annual Review of Vision Science, 2*, 345–376.

Witz, N., Zhou, J. & Hess, R. F. (2014). Similar mechanisms underlie the detection of horizontal and vertical disparity corrugations. *PLOS one, 9.*

Yan, P. & Shigemasu, H. (2015). Stereo-Curvature Aftereffect Is Due to More Than Shape Curvature Adaptation. *Perception, 44*, 790-813.

Zhang H, Gamlin PD. (1998). Neurons in the posterior interposed nucleus of the cerebellum related to vergence and accommodation. I. Steady-state characteristics. *Journal of Neurophysiology*. *79(3):* 1255-1269.

Zhang, Y., Gamlin, P. D. R., & Mays, L. E. (1991). Antidromic identification of midbrain near response cells projecting to the oculomotor nucleus. *Experimental Brain Research, 84*(3), 525-528.

Zhaoping, L. (2014). *Understanding vision: theory, models, and data*. Oxford University Press.

Zhaoping, L. (2017). Feedback from higher to lower visual areas for visual recognition may be weaker in the periphery: Glimpses from the perception of brief dichoptic stimuli. *Vision Research (136),* 32-49.