

論文内容の要旨

Three-dimensional (3D) world is full of information. Vision by human two eyes can transform the 3D information into two dimensional (2D) images, and convey them to different brain areas. The binocular visual processing can be divided into three stages: encoding, selection and decoding. However, the hierarchal processes of binocular vision and depth perception still need to be investigated. In this dissertation, two studies were implemented to investigate the hierarchal processes of binocular vision and depth perception. The first study aimed to investigate the circuit of binocular vision by using the dichoptic stimuli, which involved the encoding, selection, decoding with feedforward, top-down feedback, superior colliculus mediated eye movement. The second study was about the depth perception by using the adaptation paradigm, which mainly focused on the decoding process.

In part I, when seeing a scene in the natural world, the left and right eye images are only subtle difference. 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 the redundancy of binocular information. In the efficient coding theory, the correlated left and right eye images were decorrelated into binocular summation (S_+) and binocular suppression (S_-) channels. Based on this theory, Zhaoping (2017) investigated S_+ and S_- percepts at both central and peripheral view conditions, and reported that the bias towards S_+ percept at central vision might be due to the high level top-down feedback. This feedback signals conveyed to early visual area V1, and might play a role to help for the bottom-up selection. However, this feedback was weaker or absent at peripheral vision. In this study, we continued the previous study and investigated the circuit of the visual processing on following aspects: (1) when V1 received the feedback signals, whether it will activate the superior colliculus to mediate the vergence eye movement; (2) whether the bias towards S_+ at central and peripheral view conditions is temporally-modulated; (3) the S_+ percept distribution at different eccentricities.

Thus three experiments were conducted.

In experiment 1, using the dichoptic stimuli presenting ambiguous horizontal gratings as visual inputs for the left and right eyes with 200 ms time duration, the behavioral and eye tracking data were collected at both central and peripheral view conditions with various relative contrasts and temporal frequencies. When the left and right eye inputs were combined into the S₊ and S₋ channels, upwards or downwards motion direction can be perceived. The fractions of S₊ percepts were calculated and the change in horizontal, vertical vergence and pupil size were analyzed. From the behavioral result, the fractions (F₊) of S₊ percept for the central and peripheral conditions were about 78% and 70% respectively. There was a significant difference of fraction (F₊) of S₊ percept between the central and peripheral conditions. This bias towards S₊ percept at central condition might be due to the top-down feedback, whereas the feedback was weaker or absent at peripheral condition (Zhaoping, 2017). The psychometric curves and slopes of fractions (F₊) of S₊ percept against relative contrasts at both central and peripheral conditions had certain value, suggesting the percept bias towards S₊ was not caused by the low sensitivity of peripheral vision (Zhaoping, 2017). When the temporal frequency was higher, the fraction (F₊) of S₊ percept at central condition was lower. However, this was not the case for peripheral vision. The fraction (F₊) did not have a significant difference at different temporal frequencies. These different responses at different temporal frequencies between the central and peripheral conditions suggested the different neural mechanisms were involved for the visual decoding (Zhaoping, 2017). The central vision mainly worked for visual decoding. The visual decoding ability based on the analysis-by-synthesis computation might be constrained by the limited brain resources especially when the visual inputs had higher speed. However, the peripheral vision was for the bottom-up selection, thus was not influenced by the temporal frequency. These behavioral results matched with the previous results as shown in (Zhaoping, 2017).

From the eye information results, the pupil size distribution and horizontal vergence distribution at both central and peripheral condition meet the near-response principle, indicating the bottom-up process. However, when comparing the change in vertical vergence between central and peripheral conditions, the change in vertical vergence at central condition has negative value, with absolute amount increasing with time duration. There are significant

differences of change in vertical vergence between the central and peripheral conditions at $t > 700$ ms. The stimuli used in this experiment were the summation or subtraction of two horizontal gratings which had independent random phases for each eye, to perceive the motion direction, the left and right eyes need to match the ambiguous gratings into S_+ and S_- channels. These might cause the vertical vergence (sursumvergence) change. The negative value of change in vertical vergence suggested the right eye had higher position than the left eye. One possibility is that the central fixation located 3.6 degree left to the center of the CRT display. So the right eye had larger rotation angle than the left eye. When the top-down feedback process was involved, the right eye might need more efforts to keep the larger rotation angle, thus caused the higher position than the left eye. The longer latency might be because that multi-cycles of feedforward-feedback-verify-weight (FFVW) processes are involved, until the decision is made by the participant. Since the decoding process mainly dominates by the central vision, the top-down feedback based on the analysis-by-synthesis is caused at central vision and conveys signals to V1 area. When the V1 neurons receive the top-down feedback, they will convey signals to superior colliculus. Subsequently, the superior colliculus will mediate the V1's vergence eye movements. (The results of chapter 3 and 4 will be summarized here).

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 human vision is more sensitive to high contrast. Another reason is the second-order contrast might not elicit ocular response (Hayashi et al, 2008). From the temporal frequency aspect, when comparing the behavioral results with the eye tracking data, the higher temporal frequency (at $w = 10$ Hz) causes the lower fraction (F_+) of S_+ percept at the central 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.

This experiment used the vergence eye movements as a tool to demonstrate the circuit of

visual processing with S₊ and S. feedforward and feedback, and also the mediation of superior colliculus. Our results suggested the importance of V1 in the central and peripheral visual processes and the effectiveness of efficient coding theory.

Experiment 2 investigated whether the percept bias towards S₊ was temporal modulated, whether the bias of S₊ percept would be enlarged between the central and peripheral conditions against stimuli presenting time duration, and whether and how the participants' gaze and vergence changes along with the percept time duration. Five participants' preliminary data were collected. Experiment 3 examined the eccentricity distribution of S₊ percept (The experiment will be finished by end of December, the results and discussion will be updated then).

In part II, two experiments were conducted to investigate the mechanisms of depth adaptation of the disparity-specified sinusoidal corrugations, which mainly focused on the decoding stage. In experiment 4.1, the retinal dependency and independency of depth aftereffects was examined by manipulating the adaptation stimuli in phase static and randomly-changing conditions. In experiment 4.2, the orientation-independency of depth adaptation was investigated by manipulating the adaptation and test corrugations with different orientations, and the aftereffects were compared with the same orientation pairs. The results show that both the phase-dependency and independency are involved in the depth adaptation. Besides, the orientation independency of depth adaptation is also involved in the depth adaptation. Both experiments suggest the relative early level is involved in depth adaptation.

Experiment 5.1 examined whether the disparity- or the shape-level process was 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. And the aftereffects between the horizontal-horizontal and plaid-horizontal pairs, and also between the horizontal-plaid and plaid-plaid pairs were compared. The results showed that 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 (Cornack 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.

Experiment 5.2 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. The two adaptors were manipulated with the same peak-to-trough amplitudes and the same crossed and uncrossed disparities. The PSE shift was significantly different between the large and small adaptors in both the horizontally-oriented adaptation condition and the noise-shape condition. These results suggest that the disparity-level depth adaptation is more involved than the shape-level adaptation. 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. 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 (like V3A area) of the surface structure process.