OUKA
Osaka University Knowledge Archive

Title	Binocular cross-correlation computation produces depth perception in a relative frame of reference
Author(s)	青木, 俊太郎
Citation	大阪大学, 2016, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/56097
rights	
Note	

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

Doctoral Thesis

Binocular cross-correlation computation produces depth perception in a relative frame of reference

(両眼相関計算は相対参照枠に奥行き知覚を形成する)

Shuntaro C. Aoki

Laboratory for Cognitive Neuroscience
Graduate School of Frontier Biosciences
Osaka University

Supervisor

Prof. Ichiro Fujita (Osaka Univ.)

Dissertation Committee

Prof. Izumi Ohzawa (Osaka Univ.)
Prof. Hiromichi Sato (Osaka Univ.)
Dr. Kaoru Amano (NICT)

March, 2016

Abstract

Binocular disparity, a cue for depth perception, is encoded based on interocular cross-correlation of visual images and solution of the stereo correspondence problem. The correlation computation produces reversed depth perception in a binocularly anti-correlated random dot stereogram (aRDS) when an adjacent depth reference of a correlated RDS (cRDS) is presented. Removal of the depth reference abolishes reversed depth perception, suggesting that the correlation computation forms depth in a frame relative to the reference plane (a relative frame of reference) but not in a frame centered at a fixation point (an absolute frame of reference). Neural representation of correlation-based disparity signal, however, encodes disparity of a visual feature from the fixation point (absolute disparity) rather than relative disparity between two objects. Neurophysiological evidence thus predicts that correlation computation forms depth in an absolute frame of reference.

To clarify the controversial hypotheses, I psychophysically tested the frame of reference in which reversed depth perception in aRDSs takes place. When viewing an aRDS disk with zero absolute disparity surrounded by a cRDS annulus with non-zero (crossed or uncrossed) disparity, subjects perceived reversed depth. Since the absolute disparity of the center disk was zero, this result cannot be explained by reversal of depth in an absolute frame of reference. When RDSs without disparity (uncorrelated RDSs) were surrounded by cRDSs, subjects did not perceive reversed depth. Hence, relative disparity between aRDSs and cRDSs is essential for the reversed depth perception. The reversed depth perception occurs when aRDSs with zero absolute disparity was presented briefly (94 ms). When disparities of both the center aRDS and the surround cRDS were manipulated, subjects' depth perception was reversed from the depth of the surround but not from the fixation depth. The reversed depth perception also arose for cRDSs surrounded by aRDSs. Combination of correlation-based absolute disparity detectors produced sensitivity to relative disparity that supported reversed depth

perception in a relative frame of reference.

The results indicate that reversed depth perception in aRDSs occurs in a relative, rather than an absolute, frame of reference. I conclude that correlation-based representation produces perception of stereoscopic depth in a relative frame of reference. I suggest that correlation computation generate representation of relative depth in the brain.

Acknowledgments

Firstly, I would like to express my sincere gratitude to Professor Ichiro Fujita. He gave me an opportunity to conduct this research in Osaka University and supervised me with wisdom and kindness throughout my doctoral course. This work could not be completed without his tremendous supports. I would also like to show my particular obligation to Dr. Hiroshi M. Shiozaki. His excellent advice and encouragement are essential to this research. My appreciation goes to the members of the dissertation committee: Professor Izumi Ohzawa, Professor Hiromichi Sato, and Dr. Kaoru Amano for brilliant comments and suggestions. I am also grateful to the members of the Laboratory of Cognitive Neuroscience, Graduate School of Frontier Biosciences, Osaka University. Especially, I acknowledge Dr. Hiroshi Tamura, Dr. Takahiro Doi, Mr. Yoshiya Mori, Dr. Mohammad Abdolrahmani, Mr. Ryosuke Takeuchi, and Mr. Tomofumi Oga, for a lot of helpful discussion, assistance, and fun. Finally, I want to appreciate my family for their kind and patient support during my doctoral course.

Contents

Chapter 1: General Introduction **9**

Psychophysics and study of perception 9

Binocular disparity and depth perception 10

Neural representation of binocular disparity 16

Contribution of correlation-based representation to depth perception 22

The reference frame of reversed depth perception to aRDSs 24

Aim of this research 25

Chapter 2: Depth perception in aRDSs with zero absolute and non-zero relative disparity **29**

Introduction 29

Methods 33

Results and Discussion 36

Chapter 3: Depth perception in RDSs without binocular correspondence **45**

Introduction 45

Methods 46

Results and Discussion 49

Chapter 4: Depth perception in aRDSs with a short presentation period **55**

Introduction 55

Methods 56

Results and Discussion 59

Chapter 5: Depth perception in aRDSs with various combinations of center and surround disparities 65

Introduction 65

Methods 68

Results and Discussion 71

Chapter 6: Reversed depth perception in cRDSs surrounded by aRDSs 77

Introduction 77

Methods 77

Results and Discussion 80

Chapter 7: Model of neural representation underlying reversed depth perception in aRDSs in a relative frame of reference 85

Chapter 8: General Discussion 93

Summary and conclusion 93

The role of depth reference in depth perception based on correlation-based representation 96

Reversed depth perception for short stimulus duration 97

Reversed depth perception in center-cRDS and surround-aRDS stimuli 99

Cortical visual areas supporting the reversed depth perception in aRDSs 100

References 105

Curriculum vitae 111

Chapter 1: General Introduction

Psychophysics and study of perception

One of the major goals of neuroscience is to reveal how our brains achieve sensation and perception. The quests of perception go back to the ancient ages, and were joined in the line of modern sciences in the 19th century by a group of physicists in Germany. The early works of scientific research of perception mathematically formalized the relation between physical matters and our mind with objective measurements of our sensation caused by physical stimuli. Gustav Theodor Fechner, one of the pioneers in this field, named his project "psychophysics". Psychophysics has revealed successfully the relation between physical stimuli and perception throughout the 20th century, and hold its place in the scientific investigation of the mind in the 21st century, the era of neuroscience (Read, 2014). In modern terms, psychophysics provides input/output analysis of the brains. To understand a brain as a system, it is essential to reveal not only the biological details of the brain but also the computation carried by a brain and the algorithms which the brain uses for the computation (Marr, 1982). Psychophysics formalizes the computational aspects of a brain, and gives insights for the representation and algorithms as well as biological implementation.

Binocular disparity and depth perception

Depth perception is one of the crucial problems that our visual system confronts. We vividly perceive the world with depth; perceptual worlds and objects have three-dimensional structure. The perception of the depth is, however, not a trivial phenomenon but a challenging task for our visual system. In many animals, the visual processing begins at a layer composed of light-sensitive cells in the retina. The input to the visual system is a two-dimensional image projected onto the retina, which lacks depth information of the world. Thus, the visual system needs to reconstruct three-dimensional structure of the world from a two-dimensional image. This reconstruction is a mathematically ill-posed problem, and the solution requires constraints based on the structure of the world. The visual system achieves this reconstruction, or depth perception, by exploiting various cues which arise from the three-dimensional structure of the world. Typical examples of such cues are pictorial information caused by depth structure of the scene: perspective, size of objects, texture gradient, lighting and shading, and occlusion. Another types of cues come from motion of an observer or objects. Those cues are embedded in a single retinal image, or monocular image. In addition to these monocular cues, animals which have two eyes can exploit a different type of cue called ‘binocular disparity’.

Suppose an animal equipped with two eyes in front of their face, such as humans. Since the two eyes are horizontally separated, retinal images in the two eyes contain slight

differences when the subject is viewing something in natural conditions (Figure 1.1). The difference between left and right retinal images is called binocular disparity. Disparities caused by an object nearer and further than the fixation point are called crossed and uncrossed disparities, respectively. Based on the geometrically configuration of the two eyes and objects, the sign and magnitude of binocular disparity reflects the depth between the object and fixation point (Figure 1.2). Thus, the visual system can estimate depth structure of the object and achieve perception of depth by using binocular disparity. The process in which we perceived depth by using binocular disparity is called stereopsis.

Stereopsis has several advantages as a model system of perception. First, a visual stimulus extracting disparity processing from the other visual functions is available and enables us to examine stereopsis as a stand-alone module. Such stimulus is called a random dot stereogram (RDS; Julesz, 1960; Figure 1.3). An RDS consists of a pair of images. Each image is separately casted on left and right eyes. In each image, a number of dots are randomly distributed. The dots in the left and right images are driven from identical sequence. By shifting the horizontal position of each dot, an experimenter can introduce binocular disparity. The crucial advantage of RDSs is that they do not contain any information of shape or depth in each monocular image. Nevertheless, we clearly see depth in an RDS when dichoptically viewing it. Thus, the visual system can detect binocular disparity without preprocessing of shape or other visual attributes, and RDSs separate stereopsis as an independent module of visual function. Second, the

computational problem of stereopsis is well formalized based on the fact that RDSs cause depth perception. For achieving correct depth perception, the visual system needs to match the images of all features in a scene projected onto one eye to the corresponding images projected onto the other eye (Marr & Poggio, 1979). The problem to find binocularly matching features between left and right images is called “stereo correspondence problem”. Finally, a study of stereopsis has a plenty of literature both in psychophysics and neurophysiology as well as computational studies (for reviews, see Howard & Rogers, 1995; Poggio & Poggio, 1984; Gonzalez & Prez, 1998; Cumming & Parker, 2001; Parker, 2007; Blake & Wilson, 2011; Read, 2014).

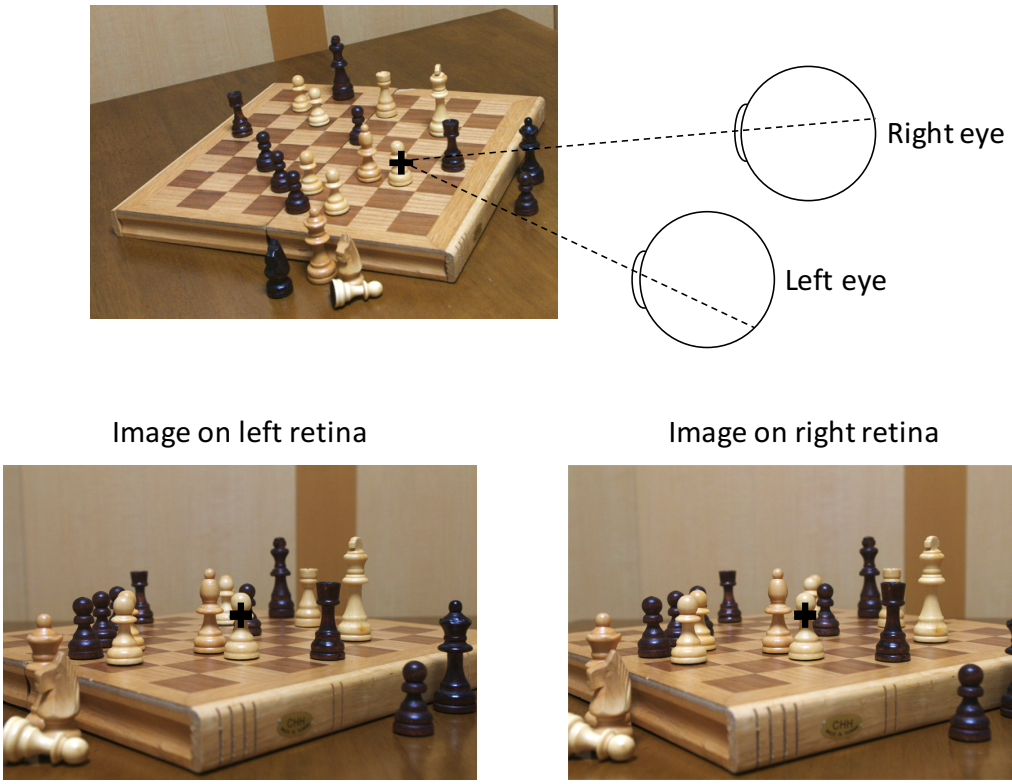

Figure 1.1: Example of stereopsis. An observer looks at chess pieces on a board with two eyes, fixating at the pawn at c2 (indicated by a black cross). Different images are casted on left and right eyes of the observer, depending on the perspective of each eye. When you parallel-fuse the two images, you will experience a clear sensation of depth.

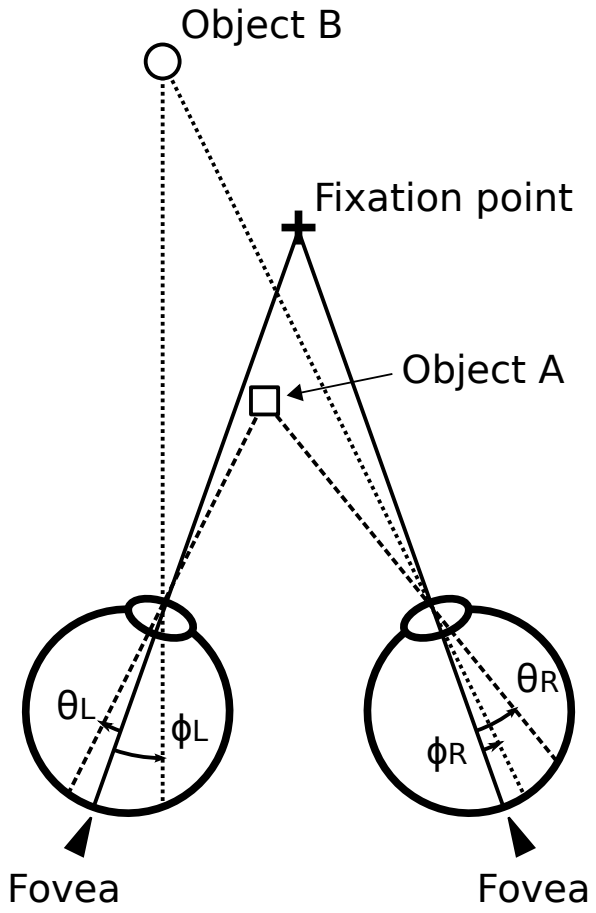

Figure 1.2: Geometrical configuration of binocular disparity. Here, an observer is fixating at the fixation point (cross) and viewing object A (square) and B (circle). The fixation point casts on the fovea in each retina (black triangles). Taking a coordinate centered at the fovea, each object casts on different retinal position between left and right eyes (object A: θ_L and θ_R; object B: ϕ_L and ϕ_R). When we define binocular disparity as left retinal position subtracted by right retinal position, the nearer object (A) has a negative disparity ($\theta_L - \theta_R$). By contrast, the farther object (B) has a positive disparity ($\phi_L - \phi_R$). Generally, the retinal coordinate is taken to make near and far objects produce negative (crossed) and positive (uncrossed) disparity, respectively.

Figure 1.3: Example of a random dot stereogram. The pair of images composes a binary RDS, the original RDS proposed by B. Julesz. If you parallel-fuse the stereogram (i.e., see the left image with your left eye and the right image with your right eye), you will observe that the central area pops out from the surrounding region.

Neural representation of binocular disparity

Over the past half century, a plenty of studies examined neural representation of binocular disparity (Gonzalez & Perez, 1998; Cumming & Parker, 2001; Parker, 2007). Disparity is first encoded in the primary visual cortex, V1, in which the inputs from two eyes firstly converge into a single neuron (Barlow et al., 1967; Nikara et al., 1968). Many visual areas including both dorsal and ventral visual pathways contain disparity sensitive neurons. The literature shows that neural representations of disparity develop toward a solution of the stereo correspondence problem along the cortical hierarchy (Nieder, 2003; Janssen, Vogels, Liu, & Orban, 2003; Tanabe, Umeda, & Fujita, 2004; Kumano, Tanabe, & Fujita, 2008; Abdolrahmani, Doi, Shiozaki, & Fujita, 2016). This process has been probed by examining neuronal responses to binocularly anticorrelated random-dot stereograms (aRDSs), in which the corresponding dots in the left-eye and right-eye images have opposite luminance contrast (Figure 1.4). aRDSs lack the globally consistent binocular match, and hence are devoid of the solution of the correspondence problem (Julesz, 1960). Therefore, neurons representing the solution should not be sensitive to binocular disparity in aRDSs (Figure 1.5). Neurons in the primary visual cortex and mid-level stages of the dorsal visual pathway (areas MT and MST) are sensitive to disparity in aRDSs with the tuning curves inverted from those for normal, correlated RDSs (cRDSs) (Cumming & Parker, 1997; Krug, Cumming, & Parker, 2004; Takemura, Inoue, Kawano, Quaia, & Miles, 2001). The inverted profile of disparity tuning suggests that the representation in these areas is based on the

cross-correlation between left-eye and right-eye images (Ohzawa, DeAngelis, & Freeman, 1990; Fleet, Wagner, & Heeger, 1996; Qian & Zhu, 1997). The disparity selectivity for aRDSs is attenuated in mid-level and higher cortical areas of the dorsal and ventral visual pathways (areas V4, IT, and AIP), suggesting that a solution to the correspondence problem is achieved in those areas (Janssen, Vogels, Liu, & Orban, 2003; Tanabe, Umeda, & Fujita, 2004; Theys, Srivastava, von Loon, Goffin, & Janssen, 2012).

In addition to the conversion from correlation-based to match-based representation, the visual system achieves another transformation of disparity information: from absolute to relative disparity. Disparity defined as 'a positional difference in retinal coordinates between left and right eyes' is called 'absolute disparity' ($\theta_L - \theta_R$ and $\phi_L - \phi_R$ in Figure 1.2). Since the retinal coordinates are centered at foveae in the eyes, absolute disparity depends on the observer's fixating point. Absolute disparity reflects depth of an object with respect to the fixation point (absolute depth). The visual system exploits not only absolute disparity but also a difference of absolute disparities between two objects, called 'relative disparity' (Westheimer, 1979; $(\theta_L - \theta_R) - (\phi_L - \phi_R)$ in Figure 1.2). Because any changes of absolute disparities are removed out by the subtraction of absolute disparities between the two objects, relative disparity remains constant for the changes of fixation point and provides relative depth between objects regardless of observer's vergence eye movements. Humans show greater sensitivity to relative than absolute disparity (Westheimer, 1979) and less sensitivity to changes in

absolute disparity without changes in relative disparity (Erkelens & Collewijn, 1985; Regan et al., 1986). Those observation suggests that the visual system makes use of relative disparity information for depth perception rather than relying only on absolute disparity. In the cortex, neurons in V1 are sensitive to absolute but not to relative disparity (Cumming & Parker, 1999). Neuronal sensitivity to relative disparity emerges in V2 (Thomas, Cumming, & Parker, 2002), and further developed across the ventral visual pathway (V4: Umeda et al, 2007; IT: Janssen et al., 2001). On the other hand, neurons in area MT, the middle stage in the dorsal visual pathway, are insensitive to relative disparity between two adjacent regions in visual field (Uka & DeAngelis, 2006) but have sensitivity to relative disparity in transparent planes (Krug et al., 2011). Overall, the visual system transforms correlation-based representation of absolute disparity into match-based representation of relative disparity (Figure 1.6).

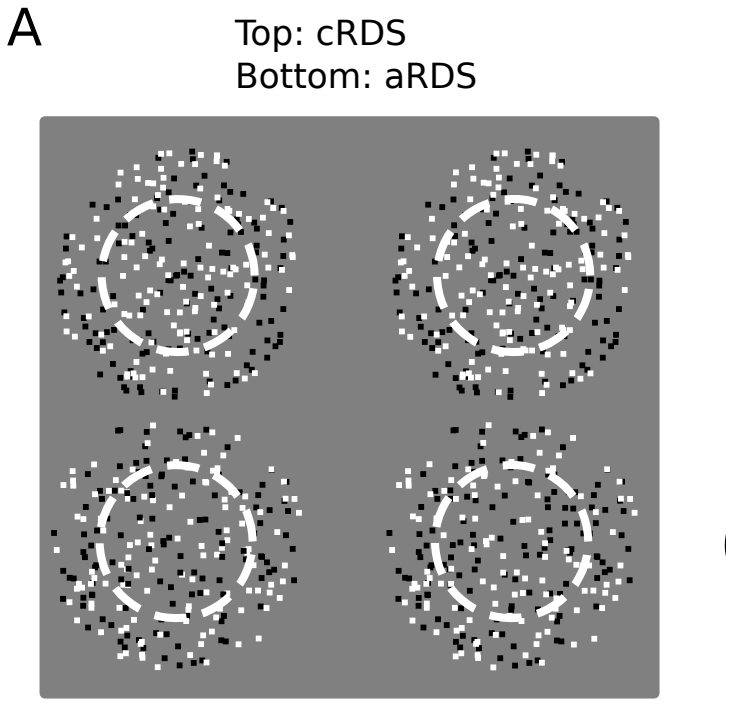

Figure 1.4: Correlated and anticorrelated RDSs. The central areas of the RDS at the top row constitute a correlated RDS (cRDS). The RDS at the bottom row has an anticorrelated RDS (aRDS) on the central region. In both RDSs, the surrounding annuli are cRDSs. When parallel-fused, the both RDSs have crossed disparity on the center. White dashed circles depict central areas.

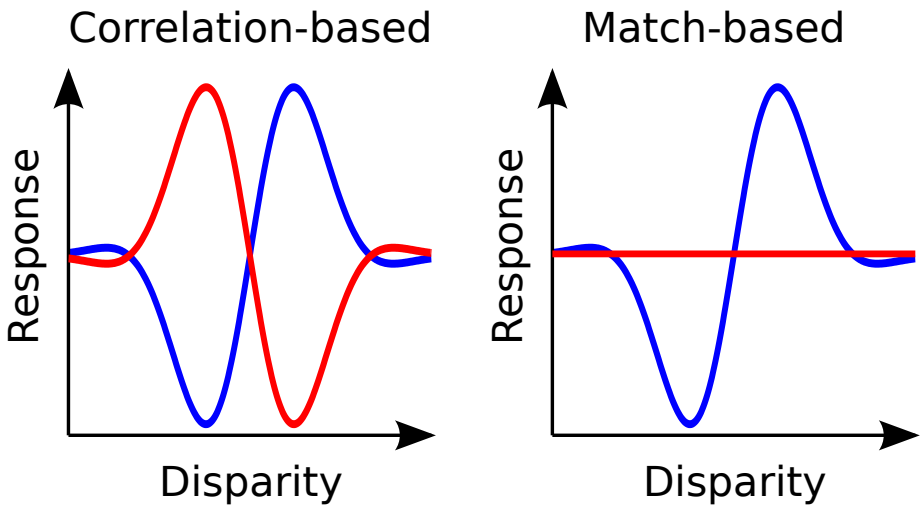

Figure 1.5: Correlation-based and match-based representation of disparity. Blue and red lines represent disparity tuning curves of neurons to cRDS and aRDS, respectively. Neurons which encode disparity based on binocular correlation show inverted disparity tuning curves to aRDSs (left plot). By contrast, neurons which encode disparity based on the solution of stereo correspondence problem do not show disparity modulation for aRDSs (right plot).

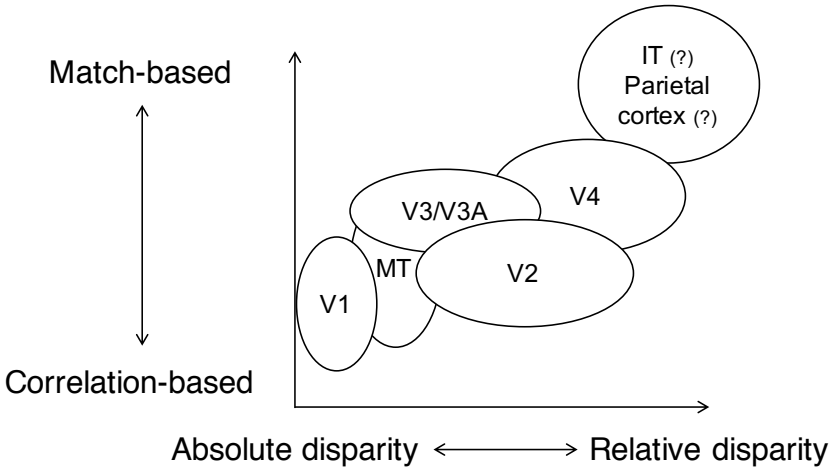

Figure 1.6: Neural representation of binocular disparity develops along with the hierarchy of cortical visual processing. Each visual area is plotted based on their neuronal sensitivity to disparity (correlation-based vs. match-based; absolute vs. relative disparity). The references are cited in the main text. Note that this is a conceptual illustration and the position of each area in this plot is inaccurate.

Contribution of correlation-based representation to depth perception

Recent psychophysical studies have demonstrated that under certain stimulus condition a patch of aRDS produces depth perception in reversed direction from the disparity-defined depth; crossed disparity evokes far perception, and uncrossed disparity evokes near perception (Tanabe, Yasuoka, & Fujita, 2008; Doi, Tanabe, & Fujita, 2011; Doi, Takano, & Fujita, 2013; bottom row in Figure 1.7). This reversal is consistent with the inverted disparity tuning of neurons that constitute correlation-based representation, providing evidence for the direct contribution of correlation-based representation to depth perception. A critical condition for producing this reversed depth perception is that a reference cRDS plane is placed immediately adjacent to a patch of aRDS. When the adjacent cRDS is replaced with an aRDS or a binocularly uncorrelated RDS (uRDS), reversed depth perception abolishes (Tanabe, Yasuoka, & Fujita, 2008; Doi, Tanabe, & Fujita, 2011; Cumming, Shapiro, & Parker, 1998; middle row in Figure 1.7). A small (0.35°) gap between the center and the surround also eliminates reversed depth perception (Hibbard, Scott-Brown, Haigh, & Adrain, 2014; Kamihirata, Oga, Aoki, & Fujita, 2015). Correlation-based representation thus mediates depth perception only when an adjacent reference plane is available. How a reference plane enables the visual system to produce reversed depth perception is unclear.

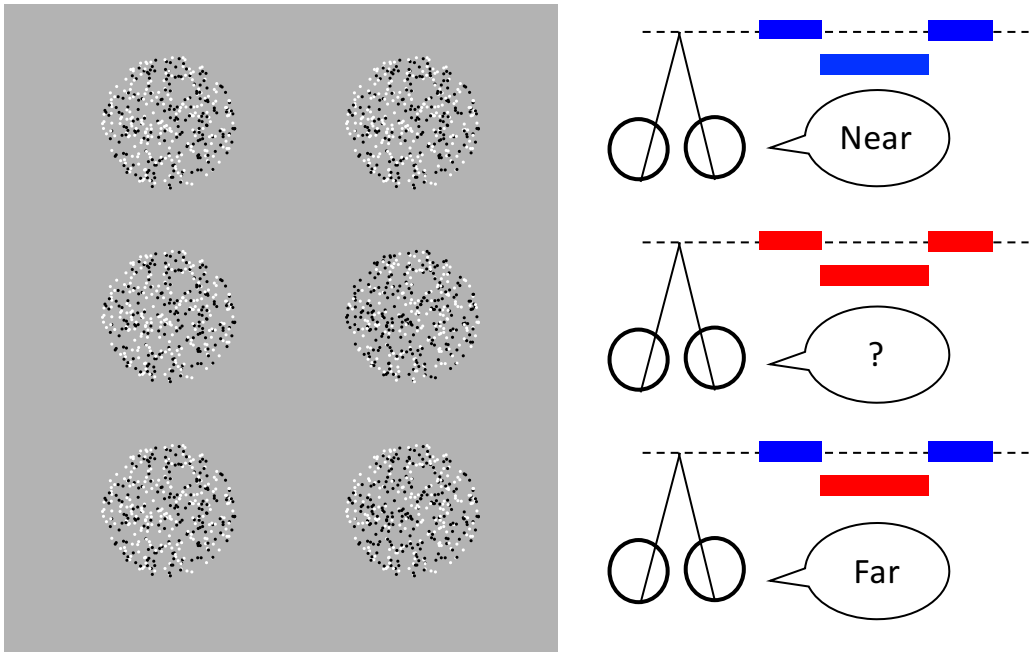

Figure 1.7: Reversed depth perception in aRDSs. The RDS in the top row is composed of a cRDS surrounded by a cRDS, causing correct depth perception. The RDS in the middle is an aRDS surrounded by an aRDS and does not elicit depth perception. The RDS in the bottom row is an aRDS surrounded by a cRDS, and produces reversed depth perception. In the all RDSs, the central areas have crossed disparity for parallel fusion.

The reference frame of reversed depth perception to aRDSs

Reversed depth perception in aRDSs shares two characteristics with perception of relative depth between two objects. First, binocular anticorrelation of a reference stimulus, which abolishes reversed depth perception (Tanabe, Yasuoka, & Fujita, 2008), increases discrimination thresholds for relative depth by more than an order of magnitude (Prince, Pointon, & Cumming, 2000; Cottereau, McKee, Ales, & Norica, 2012a). Second, an increase in the width of a spatial gap between a discrimination target and a reference stimulus, which eliminates reversed depth perception (Hibbard, Scott-Brown, Haigh, & Adrain, 2014; Kamihirata, Oga, Aoki, & Fujita, 2015), also raises discrimination thresholds for relative depth (Read, Phillipson, Serrano-Pedraza, Milner, & Parker, 2010; Cottereau, McKee, & Norica, 2012b). From these findings, I hypothesize that correlation-based representation forms reversed depth percept in a spatial frame relative to a reference stimulus (relative frame of reference) but not in a frame of reference with respect to the fixation plane (absolute frame of reference). This hypothesis predicts that a target stimulus of an aRDS evokes depth perception in the direction opposite to the relative disparity from a reference stimulus.

The physiological evidence, however, does not support this hypothesis. Previous neurophysiological studies show that correlation-based representation are mainly embedded in V1 and MT (Cumming & Parker, 1999; Krug, Cumming, & Parker, 2004). Moreover, neurons in MT are causally related to coarse depth perception (Uka &

DeAngelis, 2006). Psychophysical studies reveal that correlation-based representation dominates coarse depth discrimination (Doi et al., 2011). Hence, MT is the first candidate supporting reversed depth perception in aRDSs and contribution of correlation-based representation to depth perception. These results predict that correlation-based representation should produce depth in an absolute but not in a relative frame of references. In this scheme, the surround cRDS enable reversed depth perception by a center-surround interaction that enhance neural responses to the absolute disparity in the center aRDSs. Although there is no neurophysiological evidence that binocular correlation around receptive fields modulates the neuronal responses to aRDSs, MT neurons have center-surround receptive field organization potentially supporting such interactions (Born, 2000).

Aim of this research

In the previous studies reporting reversed depth perception, the absolute disparity of an aRDS always had the same sign as the relative disparity between an aRDS and its reference because the reference plane is always set at zero disparity (Tanabe et al., 2008; Doi et al., 2011; Doi et al., 2013). Thus, the results of these studies do not discriminate which reference frame, absolute or relative, underlies reversed depth perception. In my thesis, I addressed this question by manipulating the absolute disparities of the center aRDS and the surround cRDS independently. This stimulus manipulation enables us to distinguish reversal of depth based on relative disparity from

that based on absolute disparity. Consider, for instance, a concentric bipartite RDS consisting of a center disk of aRDS at a crossed disparity and a surround annulus of cRDS at a larger crossed disparity than the center (Figure 1.8). When an observer judges whether the center aRDS is nearer or further than the surround, depth reversal in an absolute frame of reference leads to depth judgment of 'far' (black arrow in Figure 1.8). In contrast, depth reversal in a relative frame of reference results in depth judgment of 'near' (white arrow in Figure 1.8). I demonstrate that depth perception for aRDSs is reversed with respect to the reference stimulus, suggesting that the correlation-based representation generates depth perception in a relative frame of reference.

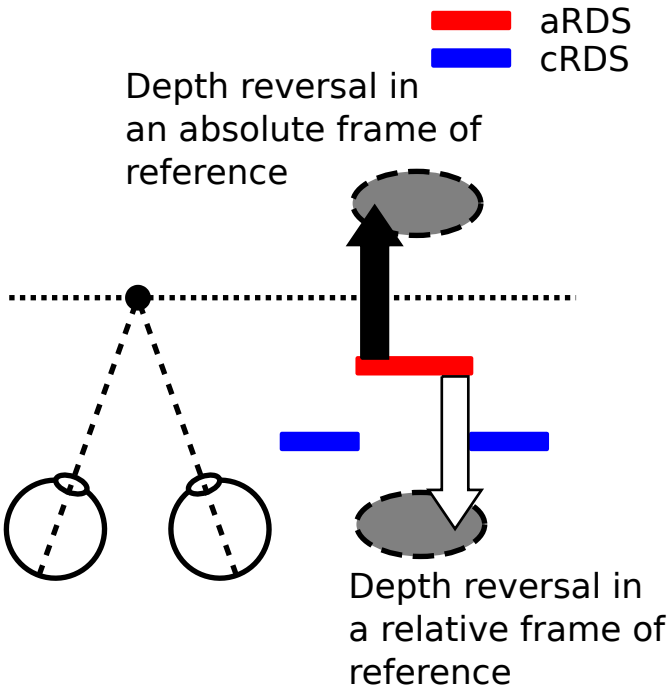

Figure 1.8: Schematic illustration of depth reversal in absolute and relative frames of reference. Here, both the center and the surround are at 'near' depth in an absolute frame of reference (i.e., both areas are nearer than the fixation plane indicated by a dashed line), but the center is at 'far' depth in a relative frame of reference (i.e., the center has a smaller crossed disparity than that of the surround cRDS). If reversed depth perception for aRDSs occurs in an absolute frame of reference, the center is perceived further than the fixation plane and the surround (black arrow). By contrast, if depth reversal occurs in a relative frame of reference, the center is perceived nearer than the surround (white arrow). The gray eclipses represent perceived depth positions.

Chapter 2: Depth perception in aRDSs with zero absolute and non-zero relative disparity

Introduction

As a first test to examine the frame of reference in which reversed depth occurs, I tested depth perception in aRDSs that had zero absolute but non-zero relative disparity. I fixed the center disparity at 0° while varying the surround disparity (Figure 2.1). In such stimuli, relative disparity between the center and the surround cRDS takes either negative or positive value and is not zero. If reversed depth perception for aRDSs occurs in a relative frame of reference, subjects should report depth opposite to the relative disparity between the center and the surround (‘near’ and ‘far’ choices for crossed and uncrossed surround disparities, respectively; see Figure 2.2). In contrast, if depth perception is dictated in an absolute frame of reference, observers would not experience reversed depth for these stimuli, because the center disparity was fixed at the horopter (Figure 2.1). In this experiment, depth reversal in a relative frame of reference leads inverted psychometric functions between cRDS and aRDS center stimuli (Left side in Figure 2.2). On the other hand, depth reversal in an absolute frame of reference leads the same psychometric functions for cRDS and aRDS center stimuli (Right side in Figure 2.2). Subjects’ reports of geometrically opposite depth should be attributed to

reversed depth perception in a relative frame of reference.

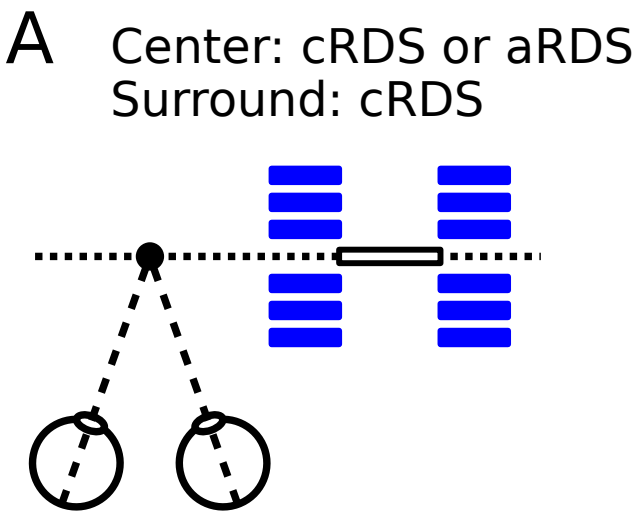

Figure 2.1: Stimulus configurations. The absolute disparity on the center (cRDS or aRDS) was fixed at zero, whereas the surround (cRDS) had absolute disparity varied from crossed to uncrossed disparity across trials.

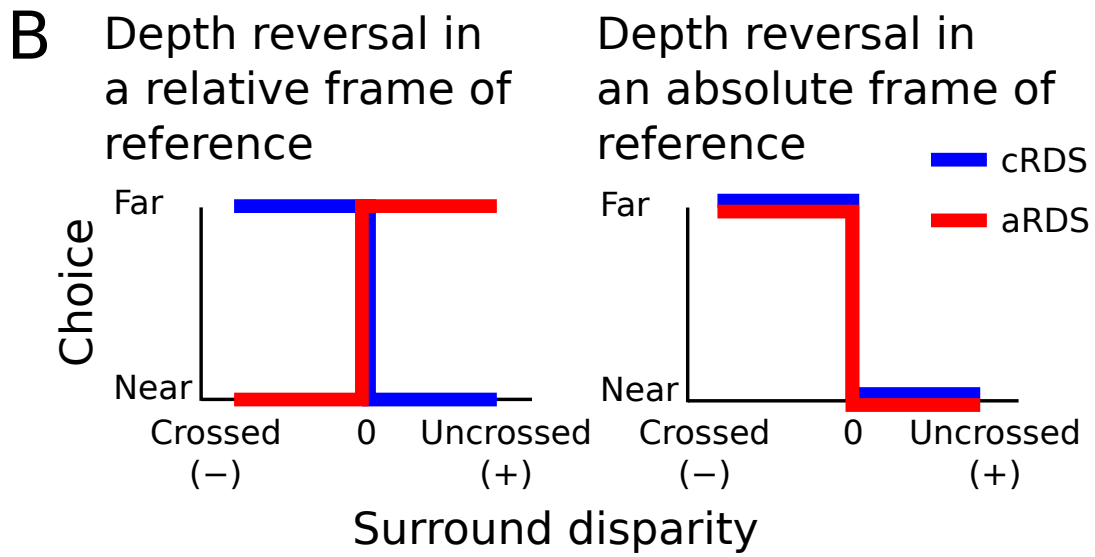

Figure 2.2: Experimental predictions. The predicted performances of subjects are plotted as proportion of far choices against surround disparity. Blue and red lines represent psychometric functions for cRDSs and aRDSs, respectively. If reversed depth perception occurs in a relative frame of reference, subjects will perceive reversed depth in the stimuli (left). By contrast, depth reversal in an absolute frame of reference will not lead reversal of depth perception (right).

Methods

Subjects

Eight subjects, including an author (SCA), participated in the experiments. All but one (SCA) subjects were naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. I obtained a written informed consent from each subject, and performed all experiments in accordance with the Declaration of Helsinki.

Each subject performed a screening test before starting the initial session. In this test, subjects reported the relative depth between the center and surround of cRDSs. The disparity of the center (center disparity) was either -0.16° or 0.16°, while the disparity of the surround (surround disparity) was fixed at zero. The eight subjects were able to discriminate the depth correctly in more than 90% of the trials. Three other subjects failed to reach the correct rate of 90%, and were not tested further.

Apparatus

The subjects viewed visual stimuli on a full-flat cathode-ray tube (CRT) monitor (Multiscan G520, Sony, Tokyo) placed 57 cm away from the eyes. The head was stabilized on a chin rest during experiments. The monitor had a spatial resolution of 1152 × 864 pixels and subtended 38.8° × 29.5° of the visual field. The monitor refresh rate was 85 Hz. Visual stimuli were generated using OpenGL and the OpenGL Utility

Toolkit (GLUT), and presented with a graphic board (NVIDIA Quadro FX3700, Elsa Japan, Tokyo). I applied antialiasing to present visual stimuli at a subpixel resolution. I used liquid crystal active shutter glasses (RE7-CANE, Elsa, Aachen) to achieve dichoptic stimulus presentation. I minimized interocular crosstalk of visual stimuli by using only red phosphors, which have the shortest decay time among the three types of phosphors used in the monitor. I assessed the amount of crosstalk by measuring the luminance of the ghost image (Tanabe et al., 2004). The crosstalk was less than 2%.

Visual stimuli

Visual stimuli were dynamic random dot stereograms (RDSs) composed of a center disk and a surrounding annulus. The diameter of the center disk was 4.8°. The width of the surrounding annulus was 1.6°. RDSs were composed of an equal number of bright and dark square dots (bright, 2.5 cd/m^2; dark, 0.0 cd/m^2; measured through a shutter glass). Those dots were presented on a mid-luminance background (1.3 cd/m^2; measured through a shutter glass). The size of each dot was 0.07° × 0.07°. In each frame, the dots occupied 25% of the area of an RDS. Positions of dots were randomized every two frames, resulting in the dot-pattern refresh rate of 42.5 Hz. The center of an RDS was positioned at 4.8° below the fixation target at the center of the screen.

Tasks

The subjects were required to discriminate whether the center disk of an RDS was in

front of (‘near’) or behind (‘far’) the surrounding annulus. Each experiment consisted of 3 blocks of trials. To keep their vergence constant, I presented nonius lines that consisted of horizontal (length: 0.8°) and vertical (length: 0.4°) lines. In each eye image, a vertical line was placed above (in right eye image) or below (in left eye image) the center of a horizontal line. Thus, the nonius lines formed a cross when correctly fused. The nonius lines were presented throughout the block, and the subjects were instructed to fuse the left and right nonius lines and maintain their fixation so that the nonius lines appeared to form a cross. The subjects took a rest for at least a few minutes between blocks.

Each trial began with the presentation of an RDS. After a fixed duration of 1 s, the stimulus disappeared and the subjects reported their choice (near vs. far) by pressing a key within a choice period of 1 s. Upon the key press the choice was displayed in words, and was kept on the screen until the end of the choice period so that the subjects can verify their own response. Within the choice period, the subjects were allowed to change their choice by pressing a key again. I did not provide any feedback on the correctness of the choice. When the subjects did not press a key during the choice period, the trial of the same stimulus condition was repeated later in the same block. The repeated trial was randomly inserted into the sequence of the remaining trials. The next trial started 1 s after the end of the choice period.

In this experiments, I examined whether humans perceive reversed depth in aRDSs that

are predicted to evoke reversed depth perception if the depth reversal occurs in a relative, but not in an absolute, frame of reference. The center was either a cRDS or an aRDS on a given trial. The surround was always a cRDS. The absolute disparity of the center, hereafter referred to as center disparity, was fixed at zero. The absolute disparity of the surround, surround disparity, was varied across trials (-0.32, -0.16, -0.08, -0.04, 0.04, 0.08, 0.16, or 0.32°; negative and positive values indicate crossed and uncrossed disparities, respectively). In each block, each surround disparity was combined with the cRDS center in 5 trials and with the aRDS center in 10 trials. In addition to these trials, I included control trials to verify that the subjects performed the task by using visual information both from the center and the surround. In the control trials, the center was cRDS of non-zero center disparity (-0.32, -0.08, 0.08, or 0.32°) and the annulus cRDS of zero surround disparity. If the subjects ignored the center disparity during the experiments, they could not judge the depth correctly in the control trials. Each control stimulus condition was repeated twice in each session. In total, each block consisted of 128 trials which were randomly interleaved (8 surround disparities × 5 repetitions for correlated center disk; 8 surround disparities × 10 repetitions for anticorrelated center disk; 4 center disparities × 2 repetitions for control). The stimulus was presented for 1 s in each trial. Each subject performed 3 blocks.

Results and Discussion

When the center disk was a cRDS, the proportion of far choices sharply dropped from 1

to 0 at the transition from crossed to uncrossed surround disparity (blue lines in Figure 2.3). For cRDSs the subjects perceived depth consistent with geometrically defined depth. In stark contrast, the proportion of far choices for an aRDS center increased at the transition in many subjects (red lines in Figure 2.3), indicating reversal of perceived depth direction from geometrically defined depth. Psychometric functions of two subjects (SCA, MM) agree well with the prediction based on depth reversal in a relative reference frame (Compare Figure 2.2 with Figure 2.3). The psychometric functions for aRDSs were, however, diverse across subjects. Choices of some subjects were biased idiosyncratically towards near or far in aRDS conditions (near: FT and SID; far: SAH; Figure 2.4). When the surround had large disparity magnitude (±0.32°), the proportion of far choices approached 0.5 (subject MM, YK, and SAH) or dropped at zero (MAM, and YST). Despite these discrepancies, all subjects shared the tendency that proportion of far choices increased as the surround disparity was shifted from crossed to uncrossed in a fine range (from -0.08° to 0.08°). The inverted psychometric function for aRDS is evident in proportion of far choices averaged across subjects (Figure 2.4). In the control trials, all subjects correctly reported geometrically defined depth, indicating that the subjects performed the task by exploiting both the center and surround disparities (Figure 2.6; proportion of correct: mean±s.d, 0.99 ± 0.01).

To quantify the magnitude of depth reversal induced by anticorrelation, I calculated the proportion of correct responses for cRDSs and aRDSs (Figure 2.5). Correct responses were defined according to the sign of relative disparity between the center and the

surround; ‘far’ and ‘near’ choices are correct for surround crossed and uncrossed stimuli, respectively. Thus, proportion of correct responses lower than chance levels indicates reversal of perceived depth. For cRDSs, the proportion of correct responses was almost 1 in all subjects (0.96 to 1; blue bars in Figure 2.5). For aRDSs, the proportion of correct responses for all surround disparities was lower than chance in all but one subject (YST) (0.18 to 0.41; red bars in Figure 2.5; binominal test, $p<0.05$). Thus, most subjects reversed their perceived depth for aRDSs with zero absolute disparity, suggesting that depth reversal takes place with respect to the reference plane.

Most of the subjects perceived reversed depth in aRDSs having zero absolute and non-zero relative disparity. Since the absolute disparity of the aRDSs was fixed at zero, the reversed depth in this experiment cannot be attributed to the reversal of absolute depth. Thus, I suggest that reversed depth perception in aRDSs relies on relative reference frame of depth. Two alternative explanation accounts for the results. First, subjects can produce similar psychometric functions by relying not on relative disparity but on depth perception in RDSs without solution of the stereo correspondence problem. Second, specific changes in vergence of subjects can result in similar psychometric functions even when the reversal of perceived depth occurs in absolute reference frame of depth. I tested these possibilities in the following experiments (Experiments in Chapter 2 and 3).

Depth perception of one subject (YST) did not reversed for aRDSs in terms of

proportion of correct responses. He reported that the center was nearer than the surround when the surround had large uncrossed disparity (0.32°; Figure 2.3). As a result, his proportion of correct fell into chance level (Figure 2.5). Subject MAM had a similar psychometric function although her proportion of correct was lower than chance level (Figure 2.3, 2.4). I do not have clear interpretation on their strong near bias.

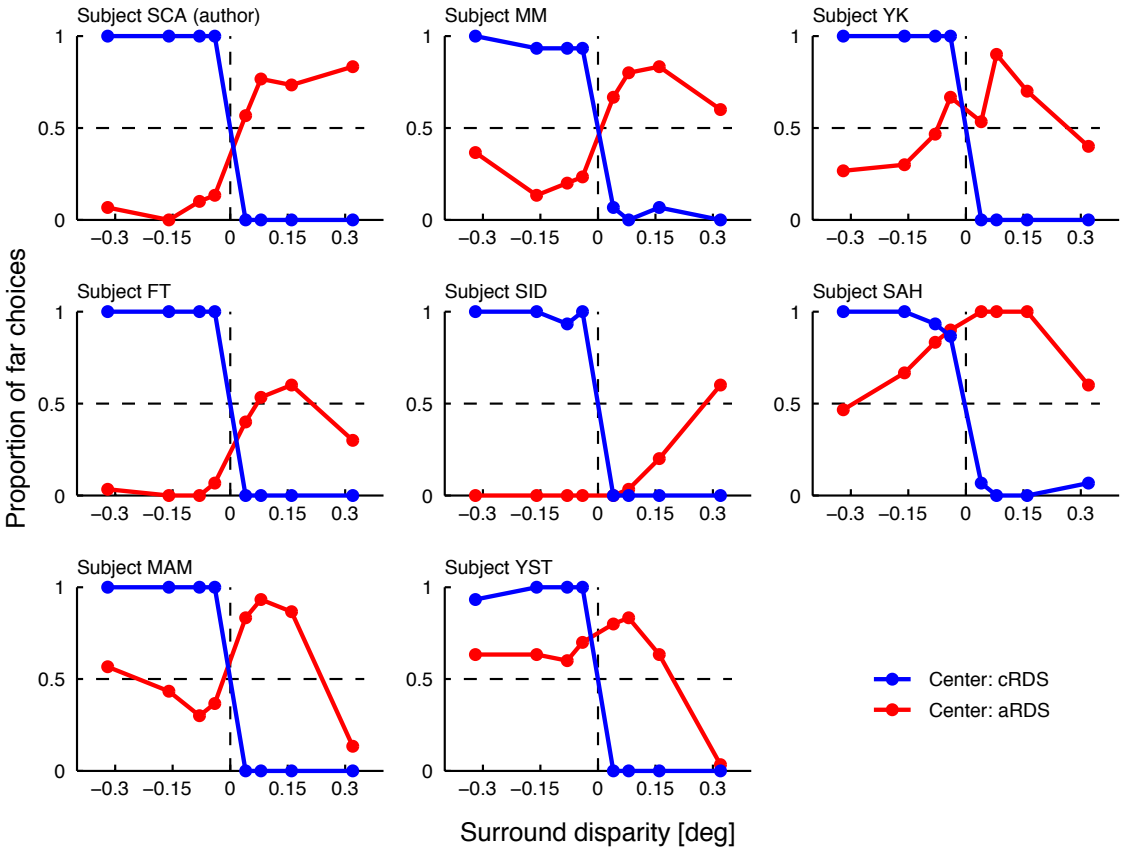

Figure 2.3: Psychometric functions of eight subjects. Proportion of far choices is plotted as a function of surround disparity. Blue and red points represent subjects' performances in cRDS and aRDS conditions, respectively.

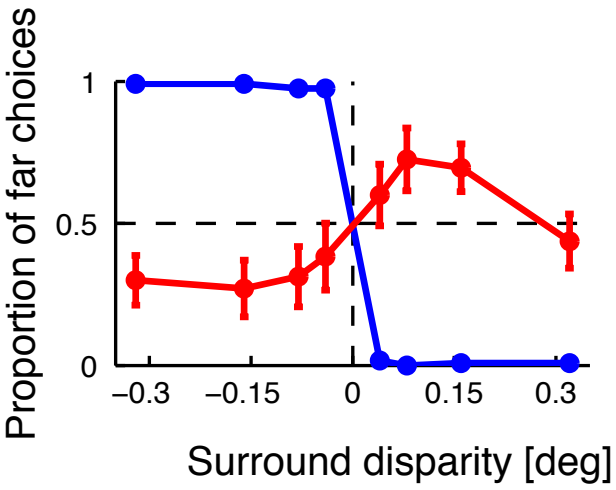

Figure 2.4: Proportion of far choices averaged across subjects (N = 8). Blue and red points represent subjects' performances in cRDS and aRDS conditions, respectively. The error bars represent SEM.

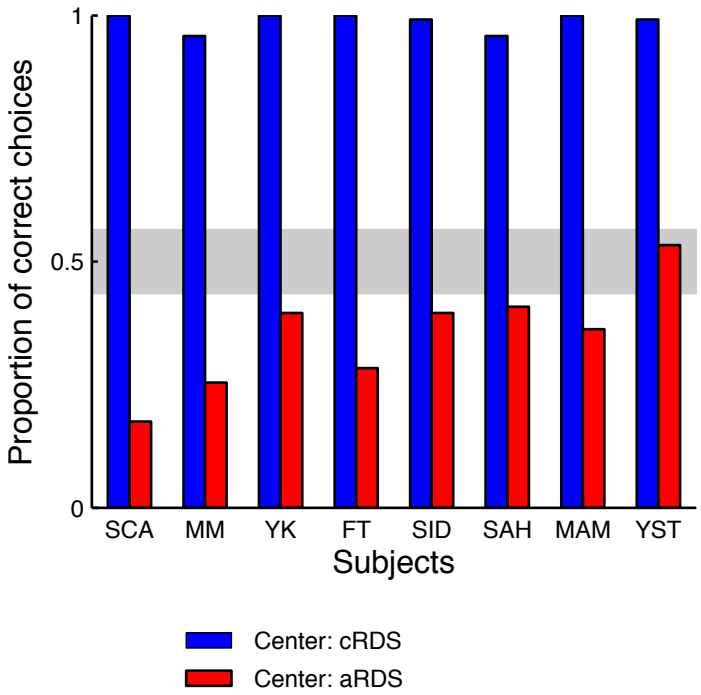

Figure 2.5: Proportion of correct choices for each subject. Blue and red bars represent proportion of correct responses in cRDS and aRDS conditions, respectively. Gray band represents chance level range calculated for all trials in aRDS conditions (240 trials, binominal test, $p < 0.05$).

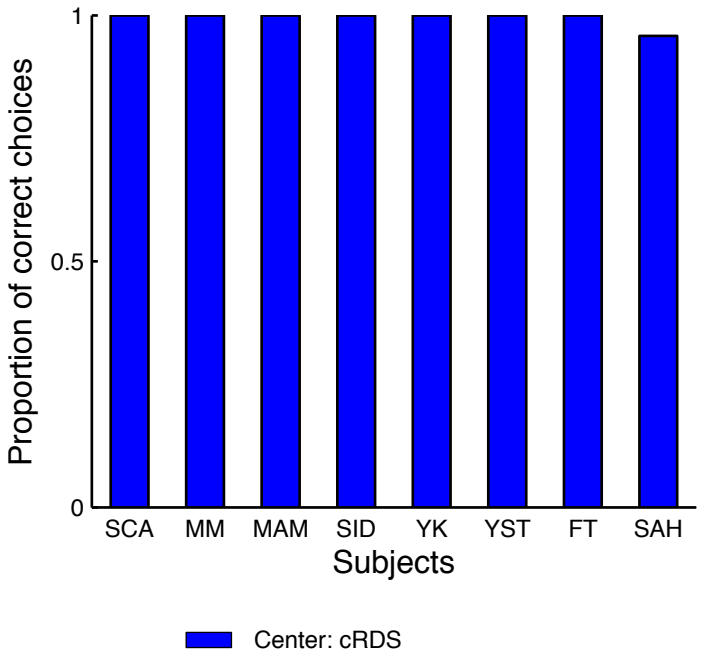

Figure 2.6: Proportion of correct choices for each subject in control trials. In control trials, the center cRDS had non-zero disparity whereas the disparity on the surround cRDS was fixed at zero.

Chapter 3: Depth perception in RDSs without binocular correspondence

Introduction

Besides the reversal of relative depth, depth judgments relying solely on the surround disparity could lead reversed depth perception in the experiment in Chapter 2. Suppose that subjects ignored the center disparity and reported the depth indicated by surround disparity (i.e., 'near' for crossed surround disparity and 'far' for uncrossed surround disparity). In such case, the resulting proportion of correct choices would be lower-than-chance. The control trials, in which the center had non-zero and the surround had zero absolute disparities, doubt the possibility that the subjects ignored the center disparity. These control trials, however, included only cRDSs. aRDSs do not evoke sensation of crisp surfaces (Tanabe et al., 2008) and thus possibly let the subjects adopt different task strategy from cRDSs. Thus, the control trials do not exclude the possibility that only the surround disparity play a role in depth discrimination specifically for aRDS. Here I examined depth perception for a binocularly uncorrelated RDS (uRDS) which is surrounded by a cRDS annulus of varying absolute disparity. uRDSs do not evoke sensation of crisp surfaces and do not contain disparity. If the subjects ignored the center disparity and reported the depth of the surround for RDSs

without crisp surface (e.g., aRDSs or uRDSs), subjects will also report reversed depth for uRDSs.

Methods

Subjects

Six subjects, also participating in the experiment in Chapter 2, participated in the experiments. All but one (SCA) subjects were naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. I obtained a written informed consent from each subject, and performed all experiments in accordance with the Declaration of Helsinki.

Apparatus

The subjects viewed visual stimuli on a full-flat cathode-ray tube (CRT) monitor (Multiscan G520, Sony, Tokyo) placed 57 cm away from the eyes. The head was stabilized on a chin rest during experiments. The monitor had a spatial resolution of 1152 × 864 pixels and subtended 38.8° × 29.5° of the visual field. The monitor refresh rate was 85 Hz. Visual stimuli were generated using OpenGL and the OpenGL Utility Toolkit (GLUT), and presented with a graphic board (NVIDIA Quadro FX3700, Elsa Japan, Tokyo). I applied antialiasing to present visual stimuli at a subpixel resolution. I used liquid crystal active shutter glasses (RE7-CANE, Elsa, Aachen) to achieve

dichoptic stimulus presentation. I minimized interocular crosstalk of visual stimuli by using only red phosphors, which have the shortest decay time among the three types of phosphors used in the monitor. I assessed the amount of crosstalk by measuring the luminance of the ghost image (Tanabe et al., 2004). The crosstalk was less than 2%.

Visual Stimuli

Visual stimuli were dynamic random dot stereograms (RDSs) composed of a center disk and a surrounding annulus. The diameter of the center disk was 4.8°. The width of the surrounding annulus was 1.6°. RDSs were composed of an equal number of bright and dark square dots (bright, 2.5 cd/m^2; dark, 0.0 cd/m^2; measured through a shutter glass). Those dots were presented on a mid-luminance background (1.3 cd/m^2; measured through a shutter glass). The size of each dot was 0.07° × 0.07°. In each frame, the dots occupied 25% of the area of an RDS. Positions of dots were randomized every two frames, resulting in the dot-pattern refresh rate of 42.5 Hz. The center of an RDS was positioned at 4.8° below the fixation target at the center of the screen.

Tasks

The subjects were required to discriminate whether the center disk of an RDS was in front of (‘near’) or behind (‘far’) the surrounding annulus. Each experiment consisted of 3 blocks of trials. To keep their vergence constant, I presented nonius lines that consisted of horizontal (length: 0.8°) and vertical (length: 0.4°) lines. In each eye image,

a vertical line was placed above (in right eye image) or below (in left eye image) the center of a horizontal line. Thus, the nonius lines formed a cross when correctly fused. The nonius lines were presented throughout the block, and the subjects were instructed to fuse the left and right nonius lines and maintain their fixation so that the nonius lines appeared to form a cross. The subjects took a rest for at least a few minutes between blocks.

Each trial began with the presentation of an RDS. After a fixed duration of 1 s, the stimulus disappeared and the subjects reported their choice (near vs. far) by pressing a key within a choice period of 1 s. Upon the key press the choice was displayed in words, and was kept on the screen until the end of the choice period so that the subjects can verify their own response. Within the choice period, the subjects were allowed to change their choice by pressing a key again. I did not provide any feedback on the correctness of the choice. When the subjects did not press a key during the choice period, the trial of the same stimulus condition was repeated later in the same block. The repeated trial was randomly inserted into the sequence of the remaining trials. The next trial started 1 s after the end of the choice period.

In this experiments, I examined the depth perception for a binocularly uncorrelated RDS (uRDS) surrounded by a cRDS to rule out the possibility that the subjects ignore the center disparity for RDSs without binocular correspondence. The stimulus set was similar to the one used in the experiment in Chapter 2 but differed in two respects. First,

I replaced the aRDS center with an uRDS center in which dot patterns were independently generated for left-eye and right-eye images. Second, surround disparity of zero was included only for uRDS center stimuli to test whether assignment of disparity to the surround changes depth perception to uRDSs. Thus, the surrounding annulus had one of 9 disparities (-0.32, -0.16, -0.08, -0.04, 0, 0.04, 0.08, 0.16, or 0.32°) for the uRDS center. Each surround disparity was repeated 5 times for cRDS center and 10 times for uRDS center. The same control trials as in the experiment in Chapter 2 were included. Thus, each session consisted of randomly interleaved 138 trials (8 surround disparities × 5 repetitions for cRDS center; 9 surround disparities × 10 repetitions for uRDS center; 4 center disparities × 2 repetitions for control trials). The duration of stimulus presentation was 1 s in each trial. Each subject performed 3 blocks.

Results and Discussion

When the center was a cRDS, the subjects almost perfectly reported geometrically correct depth, i.e., near for crossed disparities and far for uncrossed disparities (blue lines in Figures 3.1 and 3.2; blue bars in Figure 3.3). When the center was a uRDS, the proportion of far choices was independent on the sign of surround disparity in four out of six subjects (SCA, MM, YK, and SID); they preferred ‘near’ interpretation over the tested range of surround disparities (black lines in Figure 3.1). The proportion of far choices of the remaining subjects (MAM and YST) decreased as the surround disparity was changed from crossed to uncrossed (black lines in Figure 3.1). uRDSs did not give

rise an inverted psychometric function for proportion of far choices averaged across subjects (Figure 3.2). All subjects correctly reported geometrically defined depth in the control trials as in the experiment in Chapter 2 (Figure 3.4; Mean proportion of correct: 1 ± 0.00, mean$\pm$S.D.), indicating that the subjects performed the task by exploiting disparities of both the center and the surround for cRDSs. To define correctness of responses, the center was regarded as a zero disparity plane for uRDSs. Data from zero surround disparity conditions were excluded for the calculation of proportion of correct. The proportion of correct was either within the chance level (SCA, MM, YK, and SID) or higher than the chance level (MAM and YST) (black bars in Figure 5; binominal test, SCA: $p = 0.56$, MM: $p = 0.65$, YK: $p = 0.40$, SID: $p = 0.95$, MAM: $p < 0.001$, YST: $p < 0.001$). No subject showed lower-than-chance performance (reversed depth perception). These results were in sharp contrast with the results for aRDSs in the experiment in Chapter 2 where the proportion of correct was lower than chance (red bars in Figure 2.4). Thus, I conclude that the reversed depth perception in the experiment in Chapter 2 was unlikely to be the results of task strategy in which the subjects ignore the center disparity and report the surround depth for aRDSs.

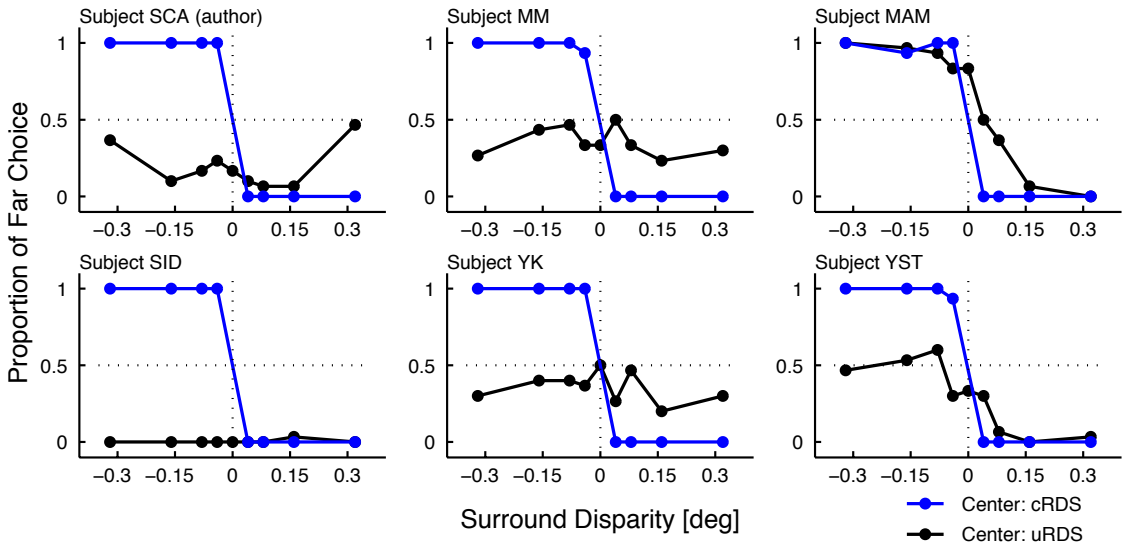

Figure 3.1: Psychometric functions. Proportion of far choice is plotted as a function of surround disparity for six subjects. Blue and black points represent subjects' performances in cRDS and uRDS conditions, respectively.

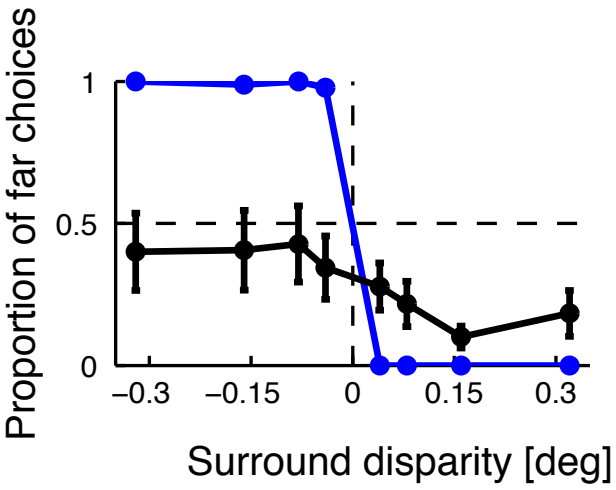

Figure 3.2: Proportion of far choices averaged across subjects (N = 6). Blue and black oints represent subjects' performances in cRDS and uRDS conditions, respectively. The error bars represent SEM.

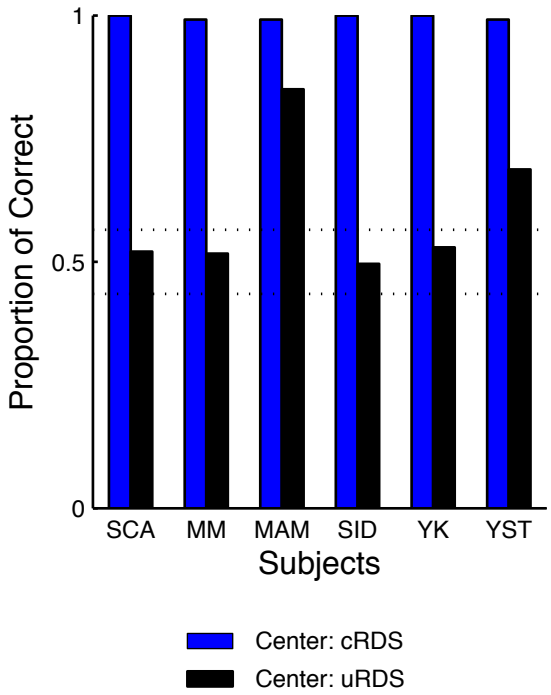

Figure 3.3: Proportion of correct responses for each subject. Blue and black bars represent proportion of correct responses in cRDS and uRDS conditions, respectively. Gray band represents chance level range calculated for all trials in uRDS conditions (240 trials, binominal test, $p < 0.05$).

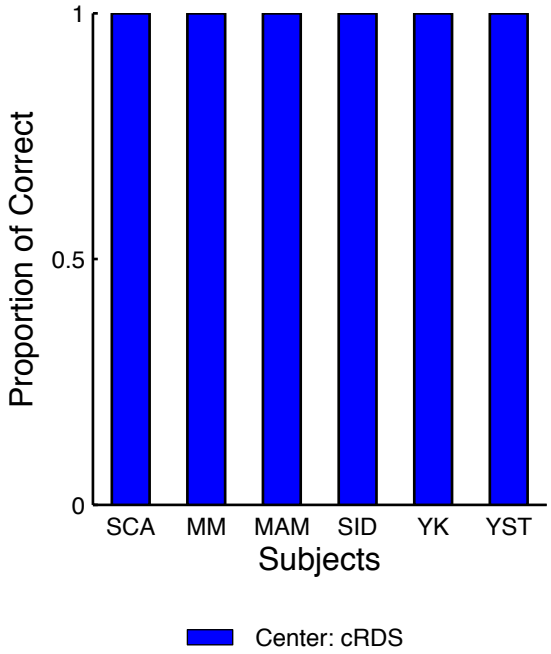

Figure 3.4: Proportion of correct responses for each subject in control trials. In control trials, the center cRDS had non-zero disparity whereas the disparity on the surround cRDS was fixed at zero.

Chapter 4: Depth perception in aRDSs with a short presentation period

Introduction

If vergence eye movement occurs during stimulus presentation, it will change absolute disparities of the stimuli, and potentially jeopardize the experimental rationale distinguishing between absolute and relative frame of references in the experiment in Chapter 2. For example, when a subject makes a convergent eye movement in response to the surround of crossed disparity, both the absolute and the relative disparities of the center indicate 'far' depth. In such a case, the subject's reports of opposite depth for aRDSs are consistent with reversed depth in both relative and absolute frames of reference. Although the subjects were required to maintain fixation at the nonius line, I did not measure the eye movement. In this experiments, I presented stimuli for a duration shorter than the latency of vergence eye movement (94 ms; Masson et al., 1997) to avoid the possible confounding effects, and tested whether the results of Chapter 2 held under this condition.

Methods

Subjects

Six subjects, also participating in the experiment in Chapter 2, participated in the experiments. All but one (SCA) subjects were naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. I obtained a written informed consent from each subject, and performed all experiments in accordance with the Declaration of Helsinki.

Apparatus

The subjects viewed visual stimuli on a full-flat cathode-ray tube (CRT) monitor (Multiscan G520, Sony, Tokyo) placed 57 cm away from the eyes. The head was stabilized on a chin rest during experiments. The monitor had a spatial resolution of 1152 × 864 pixels and subtended 38.8° × 29.5° of the visual field. The monitor refresh rate was 85 Hz. Visual stimuli were generated using OpenGL and the OpenGL Utility Toolkit (GLUT), and presented with a graphic board (NVIDIA Quadro FX3700, Elsa Japan, Tokyo). I applied antialiasing to present visual stimuli at a subpixel resolution. I used liquid crystal active shutter glasses (RE7-CANE, Elsa, Aachen) to achieve dichoptic stimulus presentation. I minimized interocular crosstalk of visual stimuli by using only red phosphors, which have the shortest decay time among the three types of phosphors used in the monitor. I assessed the amount of crosstalk by measuring the

luminance of the ghost image (Tanabe et al., 2004). The crosstalk was less than 2%.

Visual Stimuli

Visual stimuli were dynamic random dot stereograms (RDSs) composed of a center disk and a surrounding annulus. The diameter of the center disk was 4.8°. The width of the surrounding annulus was 1.6°. RDSs were composed of an equal number of bright and dark square dots (bright, 2.5 cd/m^2; dark, 0.0 cd/m^2; measured through a shutter glass). Those dots were presented on a mid-luminance background (1.3 cd/m^2; measured through a shutter glass). The size of each dot was 0.07° × 0.07°. In each frame, the dots occupied 25% of the area of an RDS. Positions of dots were randomized every two frames, resulting in the dot-pattern refresh rate of 42.5 Hz. The center of an RDS was positioned at 4.8° below the fixation target at the center of the screen.

Tasks

The subjects were required to discriminate whether the center disk of an RDS was in front of (‘near’) or behind (‘far’) the surrounding annulus. Each experiment consisted of 4 blocks of trials. To keep their vergence constant, I presented nonius lines that consisted of horizontal (length: 0.8°) and vertical (length: 0.4°) lines. In each eye image, a vertical line was placed above (in right eye image) or below (in left eye image) the center of a horizontal line. Thus, the nonius lines formed a cross when correctly fused. The nonius lines were presented throughout the block, and the subjects were instructed

to fuse the left and right nonius lines and maintain their fixation so that the nonius lines appeared to form a cross. The subjects took a rest for at least a few minutes between blocks.

Each trial began with the presentation of an RDS. After a fixed duration of 94 ms, the stimulus disappeared and the subjects reported their choice (near vs. far) by pressing a key within a choice period of 500 ms. Upon the key press the choice was displayed in words, and was kept on the screen until the end of the choice period so that the subjects can verify their own response. Within the choice period, the subjects were allowed to change their choice by pressing a key again. I did not provide any feedback on the correctness of the choice. When the subjects did not press a key during the choice period, the trial of the same stimulus condition was repeated later in the same block. The repeated trial was randomly inserted into the sequence of the remaining trials. The next trial started 1 s after the end of the choice period.

In this experiments, I addressed if the results in Chapter 1 were confounded by vergence eye movement during the stimulus presentation period. The center disparity would deviate from its nominal value if the vergence angle moves away from the fixation plane in response to disparity of the surround. This effect potentially complicates the interpretation of the results in Chapter 1, which relies on the assumption that the center disparity is fixed at zero. In this experiments, I repeated the same experiment as Chapter 1 with the stimulus duration of 94 ms, which is shorter than the latency of human

vergence eye movement and thus eliminates possible changes in vergence angle during stimulus presentation (Masson, Busettini, & Miles, 1997). The stimulus set and the number of repetitions in each session were identical to those used in the experiment in Chapter 2. In total, each block consisted of 128 trials which were randomly interleaved (8 surround disparities × 5 repetitions for correlated center disk; 8 surround disparities × 10 repetitions for anticorrelated center disk; 4 center disparities × 2 repetitions for control). The stimulus was presented for 1 s in each trial. Each subject performed 3 blocks. Each subject performed 4 blocks.

Results and Discussion

The subjects reported geometrically correct depth for cRDSs (blue lines in Figures 4.1 and 4.2; blue bars in Figure 4.3). For aRDSs, the general profile of the psychometric curves was mirror symmetrical to that for cRDSs. The proportion of far choices increased as the surround disparity shifted from crossed to uncrossed (red lines in Figures 4.1 and 4.2). All subjects correctly reported geometrically defined depth in the control trials as in the experiment in Chapter 2 (Figure 4.4; proportion of correct: mean±s.d, 0.98 ± 0.03), indicating that the subjects performed the task by exploiting disparities of both the center and the surround. The proportion of correct choices was lower than chance in all subjects (0.21-0.38, red bars in Figure 4.3; binominal test, $p < 0.05$). The subjects perceived reversed depth in aRDSs with zero absolute and non-zero relative disparities even when the stimulus duration was too short to elicit vergence eye

movement. Thus, reversed depth for aRDSs of zero absolute disparity did not result from change in absolute disparity induced by vergence eye movement.

All subjects perceived reversed depth in aRDSs with zero absolute and non-zero relative disparity when the stimulus duration was too short to evoke vergence eye movement (Masson et al., 1997). Thus, it is not likely that changes of vergence cause their performances in the experiment in Chapter 2. One subject (YST) perceived reversed depth with short stimulus duration, while not with long stimulus duration in the experiment in Chapter 2. The difference is due to disappearance of his bias of near judgment at large uncrossed disparity on the surround (0.32°). I do not have clear interpretation for the reason why long stimulus duration caused the near bias in the experiment in Chapter 2.

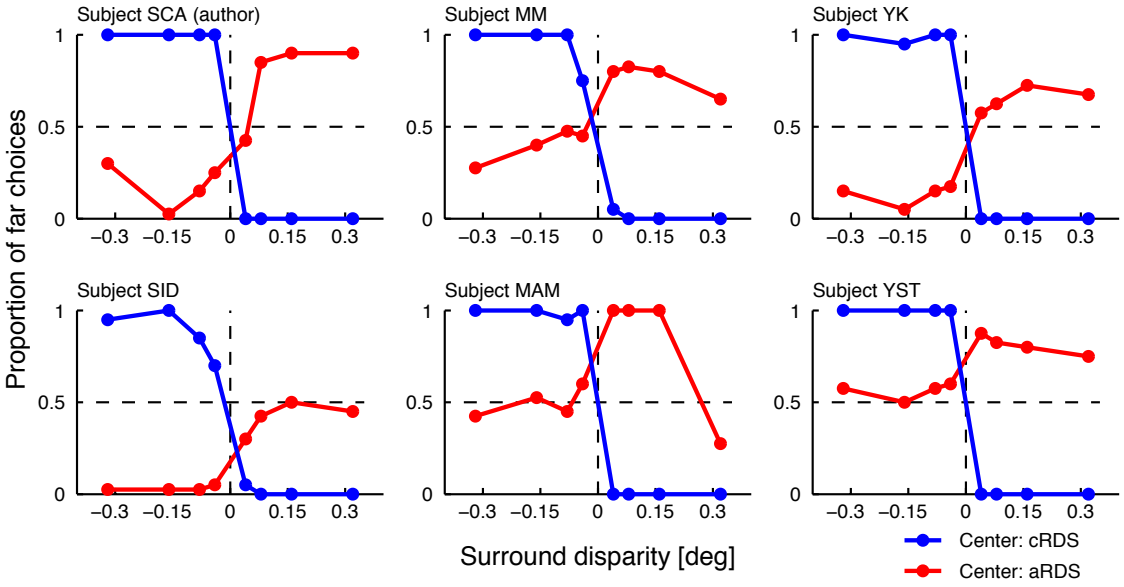

Figure 4.1: Psychometric functions of eight subjects. Proportion of far choice is plotted as a function of surround disparity. Blue and red points represent subjects' performances in cRDS and aRDS conditions, respectively.

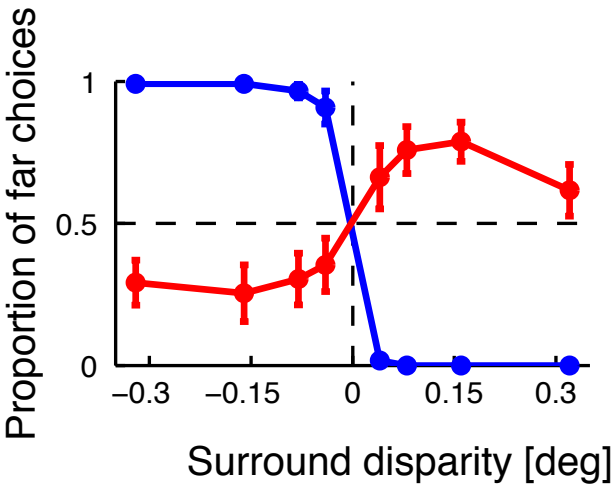

Figure 4.2: Proportion of far choices averaged across subjects (N = 6). Blue and red points represent subjects' performances in cRDS and aRDS conditions, respectively. The error bars represent SEM.

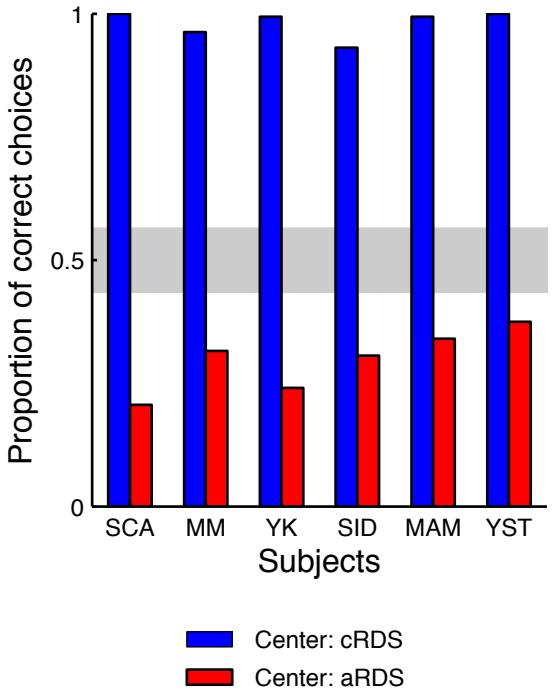

Figure 4.3: Proportion of correct responses for each subject. Blue and red bars represent proportion of correct responses in cRDS and aRDS conditions, respectively. Gray band represents chance level range calculated for all trials in aRDS conditions (320 trials, binominal test, $p < 0.05$).

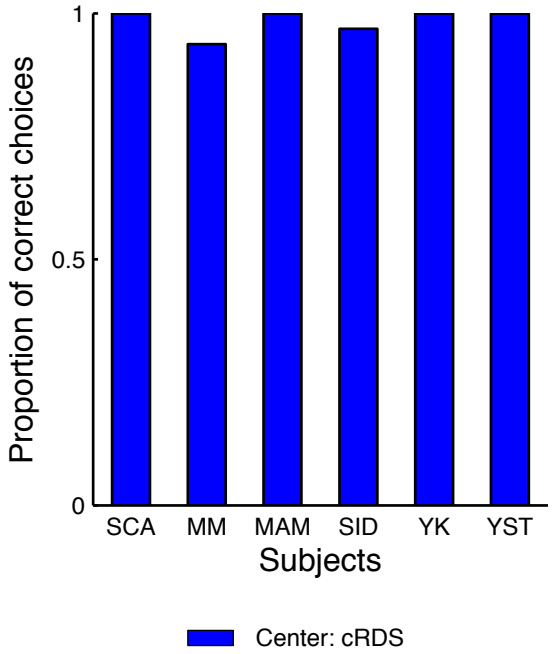

Figure 4.4: Proportion of correct responses for each subject in control trials In control trials, the center cRDS had non-zero disparity whereas the disparity on the surround cRDS was fixed at zero..

Chapter 5: Depth perception in aRDSs with various combinations of center and surround disparities

Introduction

Although an ideal relative reference frame of depth is independent of absolute disparities of stimuli, perception of depth in a relative frame of reference in practice is affected by absolute disparities of the stimuli. Specifically, humans have the highest sensitivity to relative disparity when the stimuli are in the vicinity of horopter (Blakemore, 1970). In previous studies (Tanabe et al., 2008; Doi et al., 2011; Doi et al., 2013) and the experiment in Chapter 2, either the center or the surround had zero absolute disparities and thus the stimuli were in the vicinity of horopter. Therefore, it is possible that reversed depth perception is limited to stimuli having depth around the fixation plane. Here I examined whether reversal of relative depth is generalized to various pedestal disparities of the stimuli (Figure 5.1A). If reversed depth perception in a relative frame of reference occurs in stimuli away from horopter, different surround disparities will result in horizontal shift of psychometric functions against center disparity but will not change their shape: when the surround has crossed or uncrossed

disparity, the psychometric function will shift to leftward or rightward, respectively (Figure 5.1B).

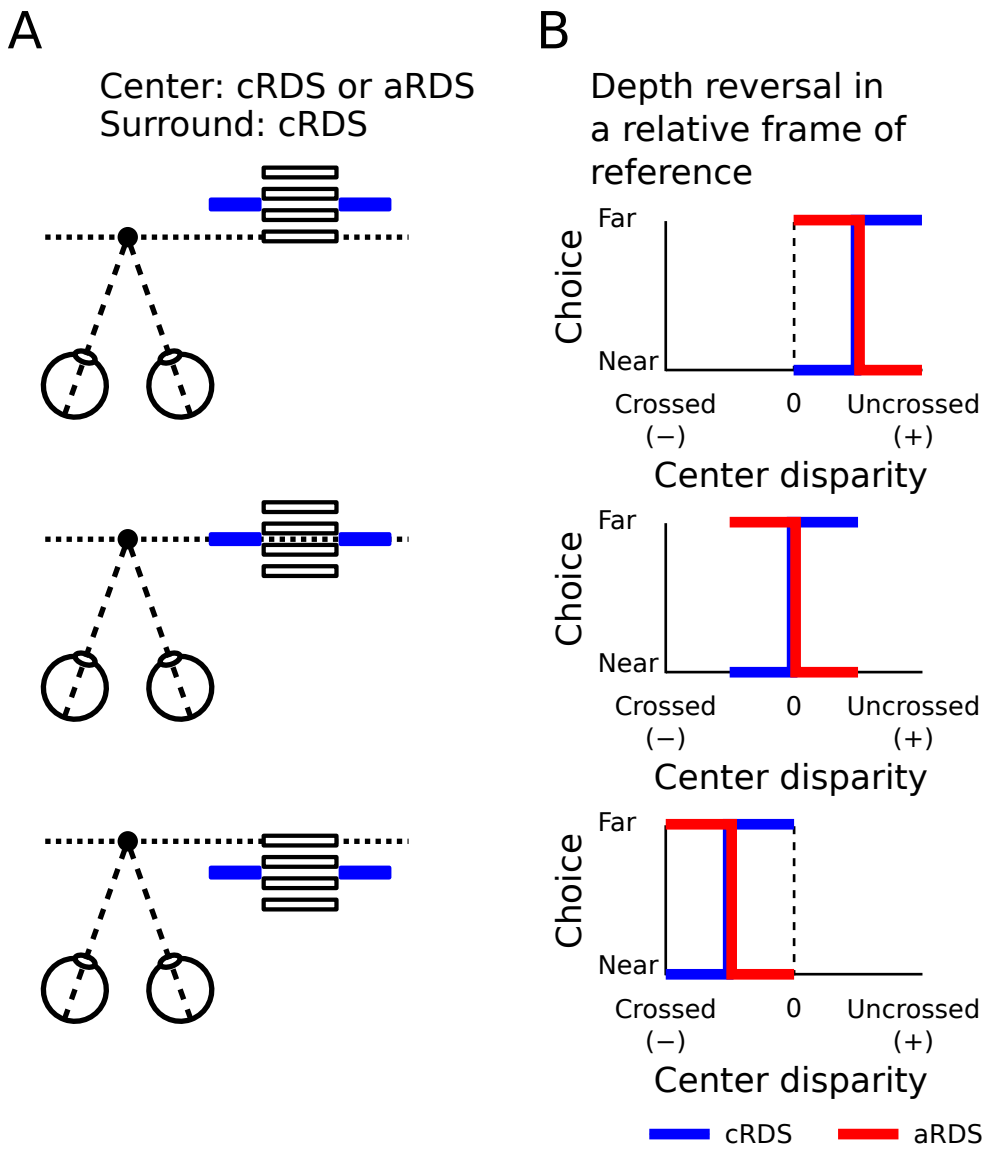

Figure 5.1: (A) Stimulus configuration. The surround disparity was either uncrossed (top), zero (middle), or crossed (bottom). The absolute disparity in the center (cRDS or aRDS) was fixed at zero, whereas the surround (cRDS) had absolute disparity varying from crossed to uncrossed disparity. (B) Experimental predictions. The predicted proportion of far choices against center disparity is plotted for each surround disparity separately (top: uncrossed surround disparity, middle: zero surround disparity, bottom: crossed surround disparity. Blue and red lines represent psychometric functions for cRDSs and aRDSs, respectively.

Methods

Subjects

Six subjects, also participating in the experiment in Chapter 2, participated in the experiments. All but one (SCA) subjects were naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. I obtained a written informed consent from each subject, and performed all experiments in accordance with the Declaration of Helsinki.

Apparatus

The subjects viewed visual stimuli on a full-flat cathode-ray tube (CRT) monitor (Multiscan G520, Sony, Tokyo) placed 57 cm away from the eyes. The head was stabilized on a chin rest during experiments. The monitor had a spatial resolution of 1152 × 864 pixels and subtended 38.8° × 29.5° of the visual field. The monitor refresh rate was 85 Hz. Visual stimuli were generated using OpenGL and the OpenGL Utility Toolkit (GLUT), and presented with a graphic board (NVIDIA Quadro FX3700, Elsa Japan, Tokyo). I applied antialiasing to present visual stimuli at a subpixel resolution. I used liquid crystal active shutter glasses (RE7-CANE, Elsa, Aachen) to achieve dichoptic stimulus presentation. I minimized interocular crosstalk of visual stimuli by using only red phosphors, which have the shortest decay time among the three types of phosphors used in the monitor. I assessed the amount of crosstalk by measuring the

luminance of the ghost image (Tanabe et al., 2004). The crosstalk was less than 2%.

Visual Stimuli

Visual stimuli were dynamic random dot stereograms (RDSs) composed of a center disk and a surrounding annulus. The diameter of the center disk was 4.8°. The width of the surrounding annulus was 1.6°. RDSs were composed of an equal number of bright and dark square dots (bright, 2.5 cd/m^2; dark, 0.0 cd/m^2; measured through a shutter glass). Those dots were presented on a mid-luminance background (1.3 cd/m^2; measured through a shutter glass). The size of each dot was 0.07° × 0.07°. In each frame, the dots occupied 25% of the area of an RDS. Positions of dots were randomized every two frames, resulting in the dot-pattern refresh rate of 42.5 Hz. The center of an RDS was positioned at 4.8° below the fixation target at the center of the screen.

Tasks

The subjects were required to discriminate whether the center disk of an RDS was in front of (‘near’) or behind (‘far’) the surrounding annulus. Each experiment consisted of 3 blocks of trials. To keep their vergence constant, I presented nonius lines that consisted of horizontal (length: 0.8°) and vertical (length: 0.4°) lines. In each eye image, a vertical line was placed above (in right eye image) or below (in left eye image) the center of a horizontal line. Thus, the nonius lines formed a cross when correctly fused. The nonius lines were presented throughout the block, and the subjects were instructed

to fuse the left and right nonius lines and maintain their fixation so that the nonius lines appeared to form a cross. The subjects took a rest for at least a few minutes between blocks.

Each trial began with the presentation of an RDS. After a fixed duration of 94 ms, the stimulus disappeared and the subjects reported their choice (near vs. far) by pressing a key within a choice period of 500 ms. Upon the key press the choice was displayed in words, and was kept on the screen until the end of the choice period so that the subjects can verify their own response. Within the choice period, the subjects were allowed to change their choice by pressing a key again. I did not provide any feedback on the correctness of the choice. When the subjects did not press a key during the choice period, the trial of the same stimulus condition was repeated later in the same block. The repeated trial was randomly inserted into the sequence of the remaining trials. The next trial started 1 s after the end of the choice period.

In this experiments, I manipulated both the center and the surround disparities to test whether reversed depth perception in aRDSs is generalized across various pedestal disparities. The center was either a cRDS or an aRDS. The surround was always a cRDS and had one of three absolute disparities (-0.16, 0.00, or 0.16°). The center disparity was chosen so that the relative disparity between the center and surround was either -0.16, -0.04, 0.04, or 0.16°. Each combination was repeated 5 times for the cRDS center and 10 times for the aRDS center. Each session consisted of 180 trials arranged

in a random order (4 center disparities × 3 surround disparities × 5 repetitions for cRDS center, 4 center disparities × 3 surround disparities × 10 repetitions for aRDS center). The stimulus was presented for 94 ms in each trial. Each subject performed 3 blocks.

Results and Discussion

When the center disk was a cRDS, judgments of depth direction agreed with the sign of relative disparity in all six subjects. When I plotted the proportion of far choices as a function of center disparity separately for the three surround disparities, the psychometric functions shifted horizontally with an invariant shape along the abscissa, i.e., the center disparity axis (blue lines in Figures 5.2A and 5.3A). The amount of shifts was in accord with changes in surround disparity. Indeed, the psychometric functions closely overlapped when plotted as a function of relative disparity (blue lines in Figures 5.2B and 5.3B). The psychometric functions for the aRDS center also shifted horizontally with an invariant shape when plotted as a function of the center disparity (red lines in Figures 5.2A and 5.3A). In contrast to the psychometric functions for cRDSs, the proportion of far choices decreased as the center disparity changed from crossed via zero to uncrossed disparities (red lines in Figures 5.2A and 5.3A). Like the psychometric functions for cRDSs, those for aRDSs overlapped when plotted as a function of relative disparity (red lines in Figures 5.2B and 5.3B). I quantified reversal of perceived depth for the three surround disparities by calculating the proportion of correct choices. Correct responses were defined based on the sign of relative disparity

between the center and the surround. For cRDSs, the proportion of correct choices was close to one for all three surround disparities (blue lines in Figure 5.4). For aRDSs, all subjects showed lower-than-chance performance at least one surround disparity (Figure 9; binominal test, $p<0.05$, Bonferroni corrected). The proportion of correct choices of three (SCA, YK, YST) and two subjects (MAM, MM) was lower than chance for all and two surround disparities, respectively (Figure 5.4; binominal test, $p<0.05$, Bonferroni corrected). Importantly, the proportion of correct choices averaged across subjects did not depend on surround disparity (Kruskal-Wallis test, $p=0.68$). Thus, anticorrelation reversed depth perception according to relative disparity at stimuli displaced in depth from the horopter.

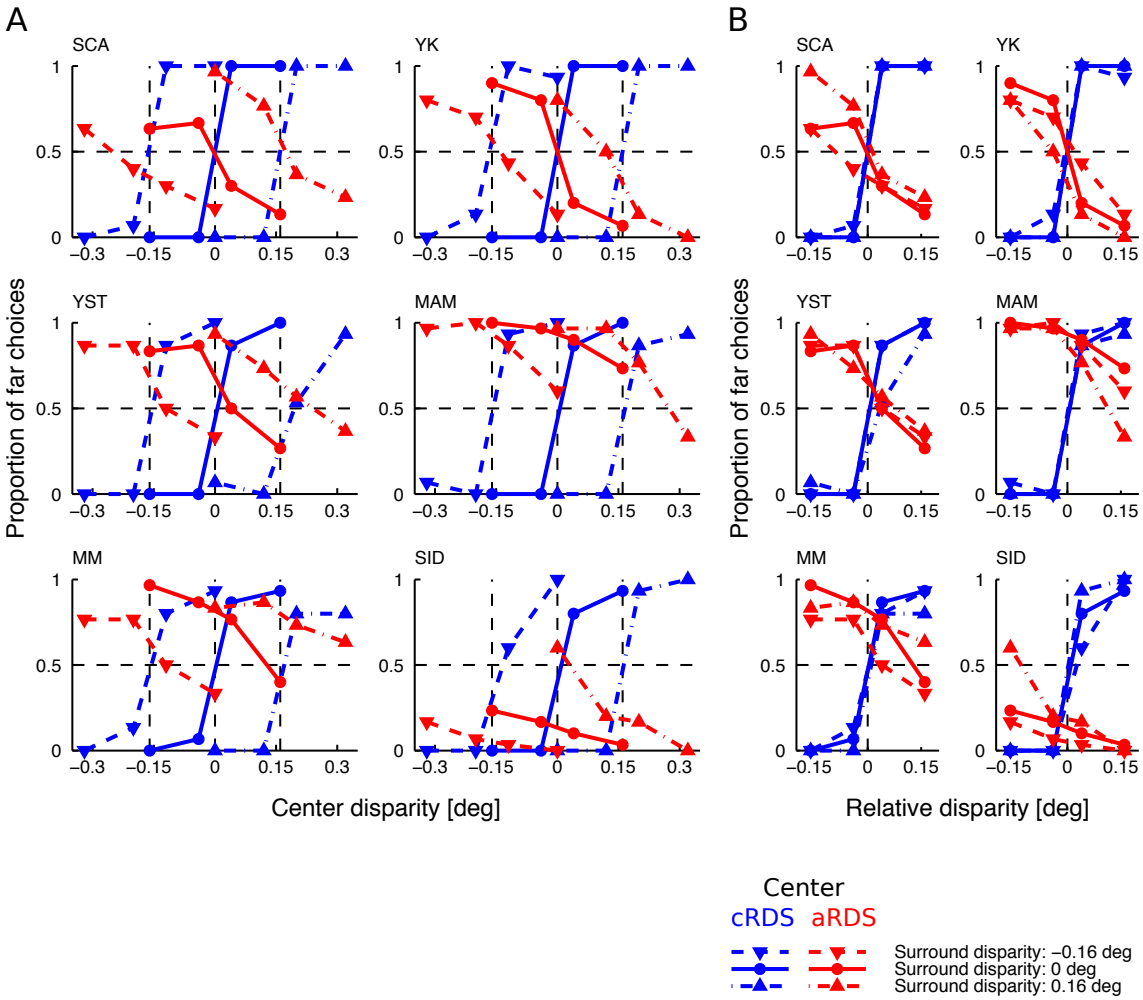

Figure 5.2: Psychometric functions of six subjects. (A) Proportion of far choice against center disparity. Blue and red points represent subjects' performances in cRDS and aRDS conditions, respectively. Downward triangles, circles, and upward triangles represent subjects' performances when the surround disparity is -0.16, 0, and 0.16°, respectively. (B) Proportion of far choice against relative disparity for the same six subjects. Blue and red points represent performances in cRDS and aRDS conditions, respectively. Downward triangles, circles, and upward triangles represent subjects' performances when the surround disparity is -0.16, 0, and 0.16°, respectively.

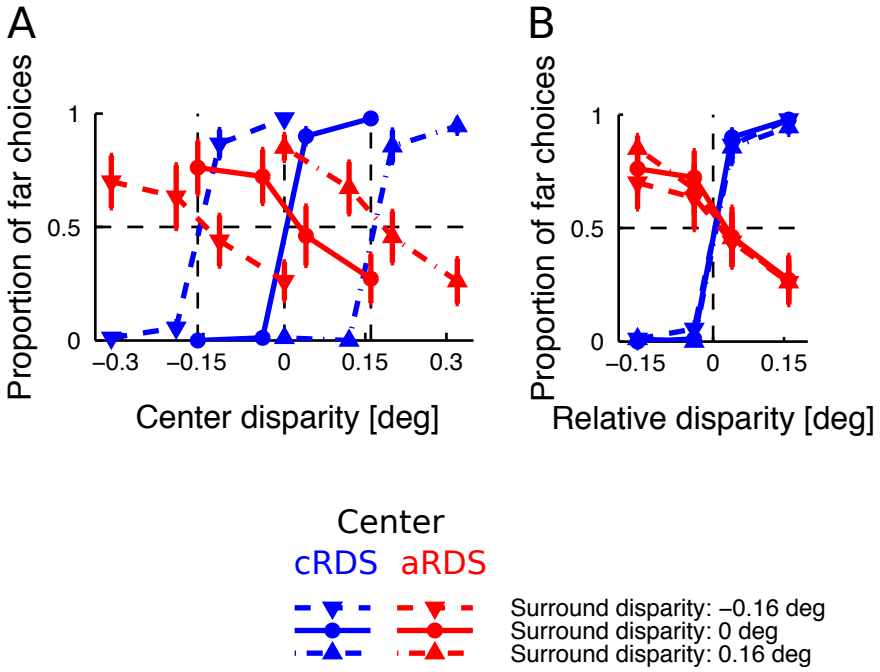

Figure 5.3: Proportion of far choices averaged across subjects (N = 6). (A) Proportion of far choices against center disparity. (B) Proportion of far choices against relative disparity. The error bars represent SEM. The other conventions of the plotting are the same as in Figure 5.2.

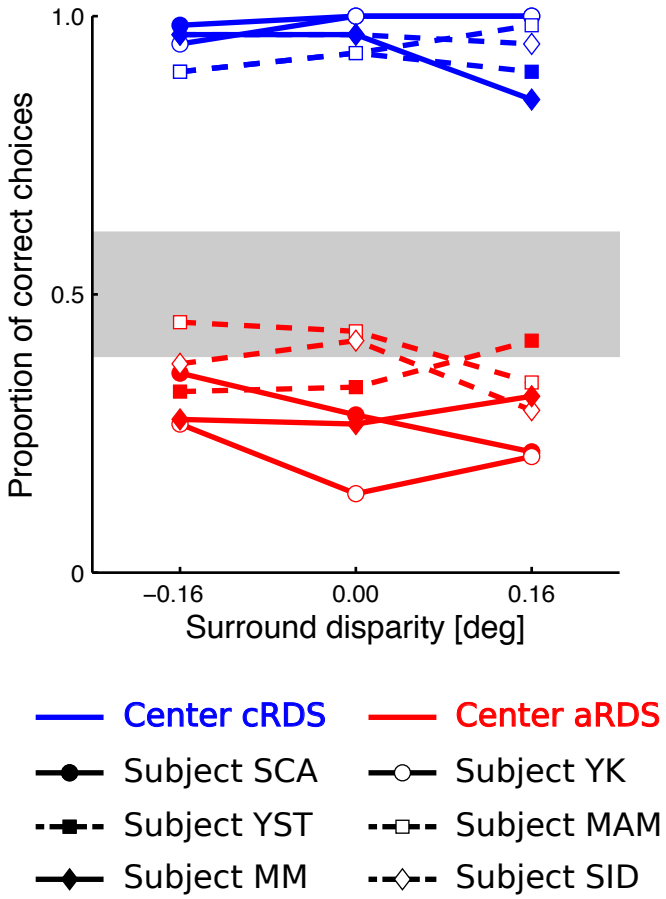

Figure 5.4: Proportion of correct responses against surround disparity. Blue and red lines represent proportion of correct responses in cRDS and aRDS conditions, respectively. Different symbols represent performances of different subjects. Gray band represents chance level range in aRDS conditions (120 trials, binominal test, $p < 0.05$, Bonferroni corrected).

Chapter 6: Reversed depth perception in cRDSs surrounded by aRDSs

Introduction

In the experiments in Chapters 2, 3, 4, and 5, aRDSs were presented at the central region of RDSs, and the surrounding annulus were cRDSs. Subjects were required to report whether the center was nearer or further than the surround in each trial. Thus, the surround cRDSs act as references of depth. It may be possible that specific spatial configuration of stimuli (e.g., correlated stimuli is placed around the discrimination target as references of depth) is necessary for the construction of a relative frame of reference and for the contribution of correlation-based representation to depth perception. If that is the case, anticorrelation of the surrounding annulus instead of the center disk will abolish the reversal of perceived depth. To test this possibility, I examined depth perception for cRDSs surrounded by aRDSs.

Methods

Subjects

Four subjects, including an author (SCA), participated in the experiments. All but one

(SCA) subjects were naïve to the purpose of the experiments. All subjects had normal or corrected-to-normal vision. I obtained a written informed consent from each subject, and performed all experiments in accordance with the Declaration of Helsinki.

Apparatus

The subjects viewed visual stimuli on a full-flat cathode-ray tube (CRT) monitor (Multiscan G520, Sony, Tokyo) placed 57 cm away from the eyes. The head was stabilized on a chin rest during experiments. The monitor had a spatial resolution of 1152 × 864 pixels and subtended 38.8° × 29.5° of the visual field. The monitor refresh rate was 85 Hz. Visual stimuli were generated using OpenGL and the OpenGL Utility Toolkit (GLUT), and presented with a graphic board (NVIDIA Quadro FX3700, Elsa Japan, Tokyo). I applied antialiasing to present visual stimuli at a subpixel resolution. I used liquid crystal active shutter glasses (RE7-CANE, Elsa, Aachen) to achieve dichoptic stimulus presentation. I minimized interocular crosstalk of visual stimuli by using only red phosphors, which have the shortest decay time among the three types of phosphors used in the monitor. I assessed the amount of crosstalk by measuring the luminance of the ghost image (Tanabe et al., 2004). The crosstalk was less than 2%.

Visual Stimuli

Visual stimuli were dynamic random dot stereograms (RDSs) composed of a center disk and a surrounding annulus. The diameter of the center disk was 4.8°. The width of the

surrounding annulus was 1.6°. RDSs were composed of an equal number of bright and dark square dots (bright, 2.5 cd/m^2; dark, 0.0 cd/m^2; measured through a shutter glass). Those dots were presented on a mid-luminance background (1.3 cd/m^2; measured through a shutter glass). The size of each dot was 0.07° × 0.07°. In each frame, the dots occupied 25% of the area of an RDS. Positions of dots were randomized every two frames, resulting in the dot-pattern refresh rate of 42.5 Hz. The center of an RDS was positioned at 4.8° below the fixation target at the center of the screen.

Tasks

The subjects were required to discriminate whether the center disk of an RDS was in front of ('near') or behind ('far') the surrounding annulus. Each experiment consisted of 4 blocks of trials. To keep their vergence constant, I presented nonius lines that consisted of horizontal (length: 0.8°) and vertical (length: 0.4°) lines. In the left-eye and right-eye images, the nonius lines formed T- and inverted T-shapes, respectively. The nonius lines were presented throughout the block, and the subjects were instructed to fuse the left and right nonius lines and maintain their fixation so that the nonius lines appeared to form a cross. The subjects took a rest for at least a few minutes between blocks.

Each trial began with the presentation of an RDS. After a fixed duration of 94 ms, the stimulus disappeared and the subjects reported their choice (near vs. far) by pressing a key within a choice period of 500 ms. Upon the key press the choice was displayed in

words, and was kept on the screen until the end of the choice period so that the subjects can verify their own response. Within the choice period, the subjects were allowed to change their choice by pressing a key again. I did not provide any feedback on the correctness of the choice. When the subjects did not press a key during the choice period, the trial of the same stimulus condition was repeated later in the same block. The repeated trial was randomly inserted into the sequence of the remaining trials. The next trial started 1 s after the end of the choice period.

In this experiments, the center was always a cRDS. The surround was either a cRDS or an aRDS, and had one of three absolute disparities (-0.16, 0.00, or 0.16°). The center disparity was chosen so that the relative disparity between the center and surround was either -0.16, -0.04, 0.04, or 0.16°. Each combination was repeated 5 times for the cRDS center and 10 times for the aRDS center. Each session consisted of 180 trials arranged in a random order (4 center disparities × 3 surround disparities × 5 repetitions for cRDS surround, 4 center disparities × 3 surround disparities × 10 repetitions for aRDS surround). The stimulus was presented for 94 ms in each trial. Each subject performed 3 blocks.

Results and Discussion

When the surround disk was a cRDS, judgments of depth direction agreed with the sign of relative disparity in all four subjects as in the experiment in Chapter 5 (blue lines in Figure2 6.1 and 6.2). The psychometric functions for the aRDS surround also shifted

horizontally with an invariant shape when plotted as a function of the center disparity (red lines in Figures 6.1A and 6.2A). Like the psychometric functions for cRDSs, those for aRDSs overlapped when plotted as a function of relative disparity (red lines in Figures 6.1B and 6.2B). I quantified reversal of perceived depth for the three surround disparities by calculating the proportion of correct choices. Correct responses were defined based on the sign of relative disparity between the center and the surround. For cRDSs, the proportion of correct choices was close to one for all three surround disparities (blue lines in Figure 6.3). For aRDSs, the proportion of correct choices of three subjects (SCA, MF, MAM) was lower than chance for all surround disparities (Figure 6.3; binominal test, $p<0.05$, Bonferroni corrected). The remaining subject (HOI) showed almost chance performances for aRDS-surround (Figure 6.3). The HOI's psychometric functions for aRDSs, however, showed gradual decreases as the relative disparity shifted from negative to positive (Figure 6.1), resulting to moderated inversion from psychometric functions for cRDSs in all surround disparities. Thus, anticorrelation of the surround annulus caused reversed depth perception according to relative disparity as in the experiment in Chapter 5. Reversed depth perception is unlikely to depend on spatial alignment of correlated depth references. Rather, the visual system is likely to construct a relative frame of reference when a correlated stimulus is presented somewhere in the visual stimuli.

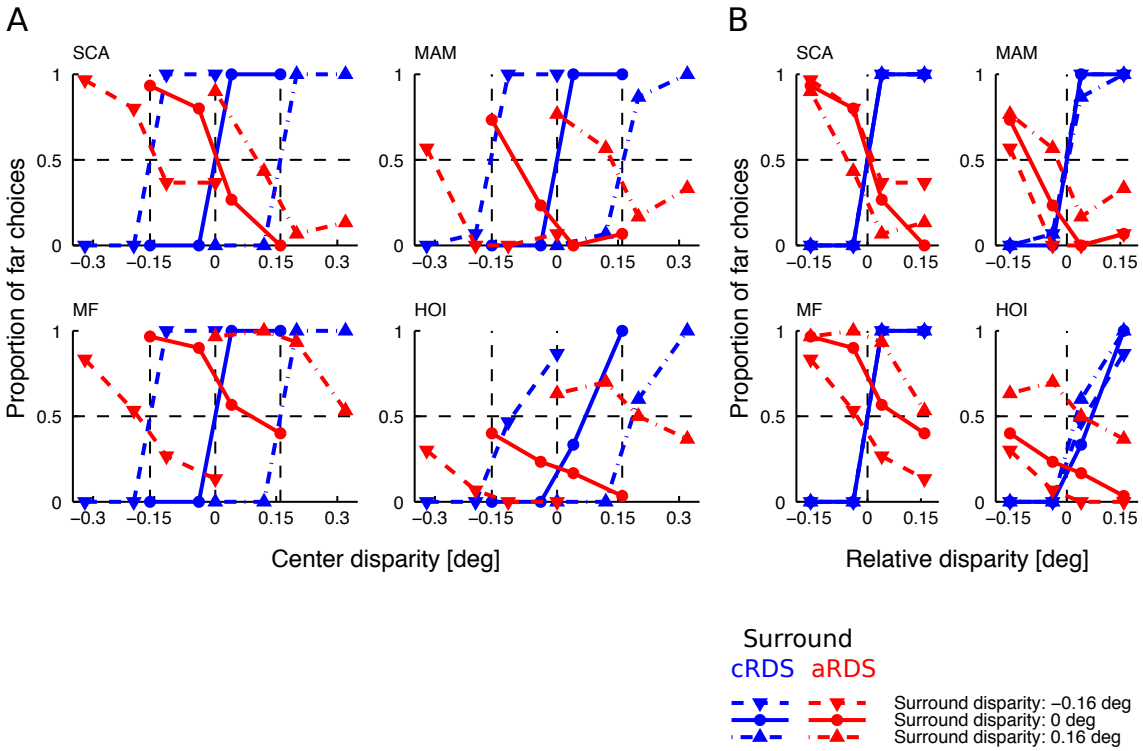

Figure 6.1: Psychometric functions of four subjects. (A) Proportion of far choice against center disparity. Blue and red points represent subjects' performances in surround-cRDS and surround-aRDS conditions, respectively. Downward triangles, circles, and upward triangles represent subjects' performances when the surround disparity is -0.16, 0, and 0.16°, respectively. (B) Proportion of far choice against relative disparity for the same six subjects. Blue and red points represent performances in cRDS and aRDS conditions, respectively. Downward triangles, circles, and upward triangles represent subjects' performances when the surround disparity is -0.16, 0, and 0.16°, respectively.

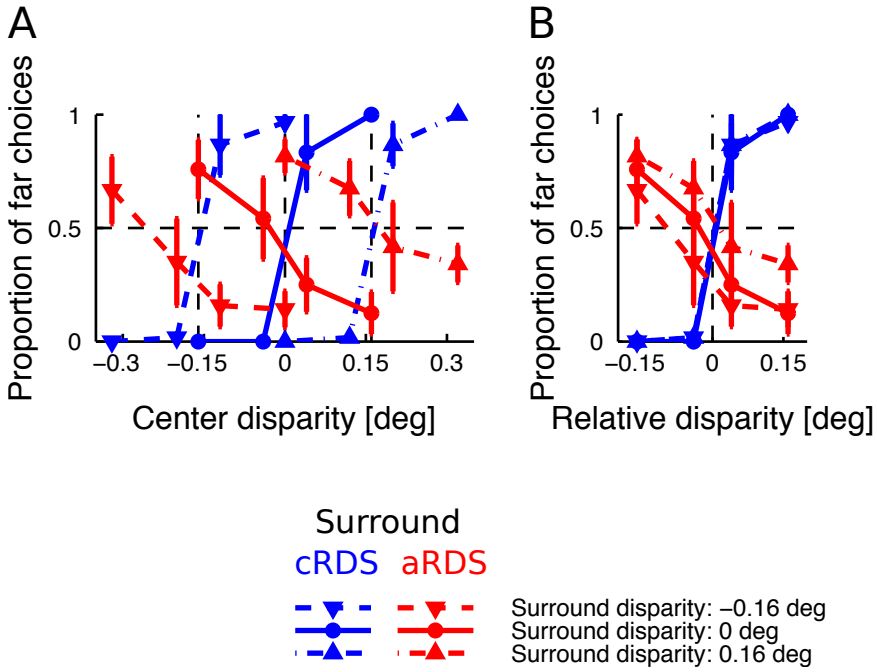

Figure 6.2: Proportion of far choices averaged across subjects (N = 4). (A) Proportion of far choices against center disparity. (B) Proportion of far choices against relative disparity. The error bars represent SEM. The other conventions of the plotting are the same as in Figure 6.1.

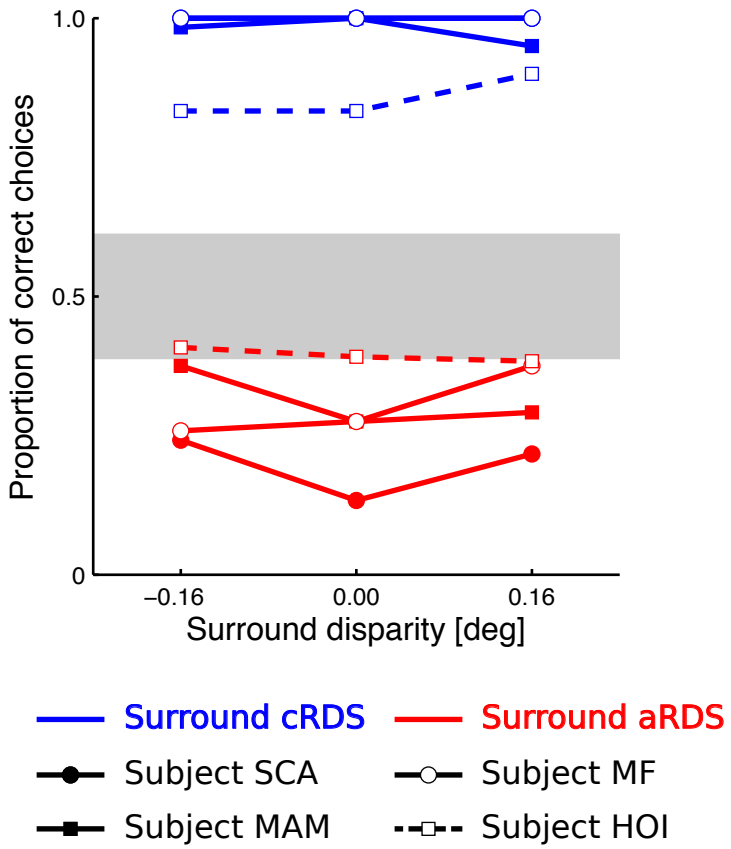

Figure 6.3: Proportion of correct responses against surround disparity. Blue and red lines represent proportion of correct responses in cRDS and aRDS conditions, respectively. Different symbols represent performances of different subjects. Gray band represents chance level range in aRDS conditions (120 trials, binominal test, $p < 0.05$, Bonferroni corrected).

Chapter 7: Model of neural representation underlying reversed depth perception in aRDSs in a relative frame of reference

Our results suggest that neurons constituting the correlation-based representation encode relative disparity. Such response properties can be generated by modification of a previously proposed model of relative disparity encoding (Figure 7.1; Thomas, Cumming, & Parker, 2002). The model proposed by Thomas et al. produces selectivity to relative disparity by combining neurons selective to absolute disparity. The first step of the model consists of two pairs of absolute-disparity detectors. In each pair, one disparity detector is sensitive to absolute disparity of the center of stimuli, whereas the other is sensitive to absolute disparity of the surround. The two pairs share the difference of preferred disparities between the center and the surround disparity detectors (Δd), but have different phases in the Gabor tuning curves. Next, the outputs of the center and the surround disparity detectors in each pair (c and s) are summed and squared, producing an interaction between disparities in the center and the surround ($(\mathrm{c}+\mathrm{s})^2 = c^2 + s^2 + 2cs$). Finally, the outputs of the pairs are summed in a next-stage neuron (R). The model shows strong response when the difference of disparities between the center and the surround is Δd, resulting in the response peak elongated

along diagonal in center-surround disparity space (top-left panel in Figure 7.3). The diagonal response profile leads to the relative disparity selectivity across a range of absolute disparities in the center and the surround (bottom-left panel in Figure 7.3).

Thomas et al. (2002) did not examine the responses of absolute disparity detectors to aRDSs. Here, we extended the model to examine the response of the model to aRDSs in light of the present finding of the depth reversal in a relative frame of reference. I incorporated the V1-like disparity selective neurons (Cumming & Parker, 1997), in which the cell's disparity tuning to aRDSs has inverted peak and reduced amplitude from that to cRDSs (Figure 7.2), to the initial absolute disparity detectors. The disparity tuning curve of each absolute disparity detector was modeled by a Gabor function. When an RDSs in the receptive field of a disparity detector was anticorrelated, amplitude of the Gabor function had negative value less than 1. With this assumption, the absolute disparity detectors had inverted disparity tunings with reduced response amplitude to aRDSs, like as the actual neurons in V1 (Cumming & Parker, 1997).

My extension of Thomas' model enabled the model to generate relative disparity selectivity with inverted disparity tuning for aRDSs. When the center was an aRDS, the outputs of the center disparity detectors, as well as the center-surround interaction terms, was inverted. This modification changed the model response to aRDSs to a diagonal response profile in which the peak and the trough were swapped compared with the response for center-cRDS stimuli (top-right pane in Figure 7.3). As a result, the model

showed inverted disparity tuning curves, preserving a horizontal shift along with the surround disparities (bottom-right panel in Figure 7.3). A simple depth discrimination scheme based on the model reproduces psychophysical performances in the experiment in Chapter 5. I assumed that a decision making process judged the relative depth of the center from the surround (‘near’ or ‘far’) by comparing the outputs of near and far relative-disparity detectors (Shiozaki et al., 2012). Both cRDS- and aRDS-center stimuli caused the same horizontal shift of psychometric functions along with surround disparities, but the curves in the two conditions had opposite slope, as in the results of Chapter 5 (Figure 7.4). In addition, aRDSs resulted in shallower slope of the predicted psychometric functions than cRDS, as in the results of Chapter 5. The model creates neuronal representation supporting reversed depth perception in aRDSs in a relative frame of reference.

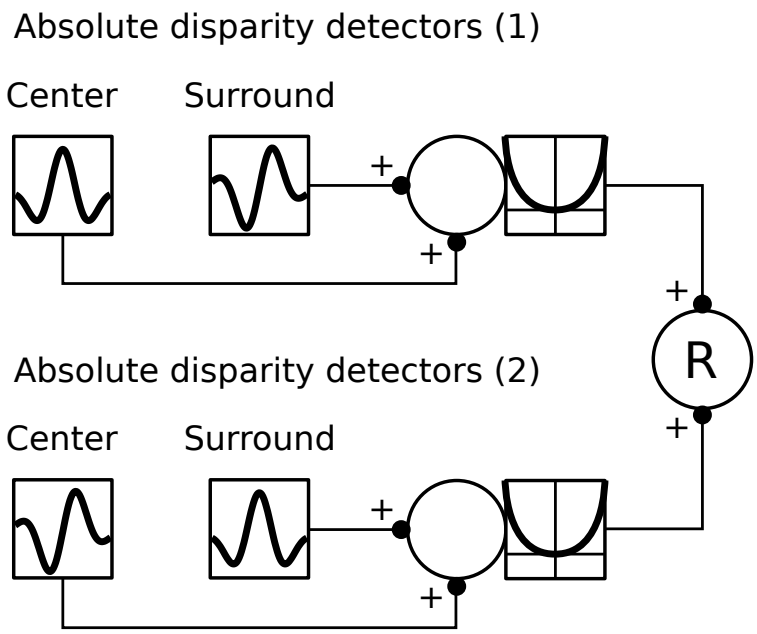

Figure 7.1: Structure of the model to generate correlation-based representation of relative disparity.

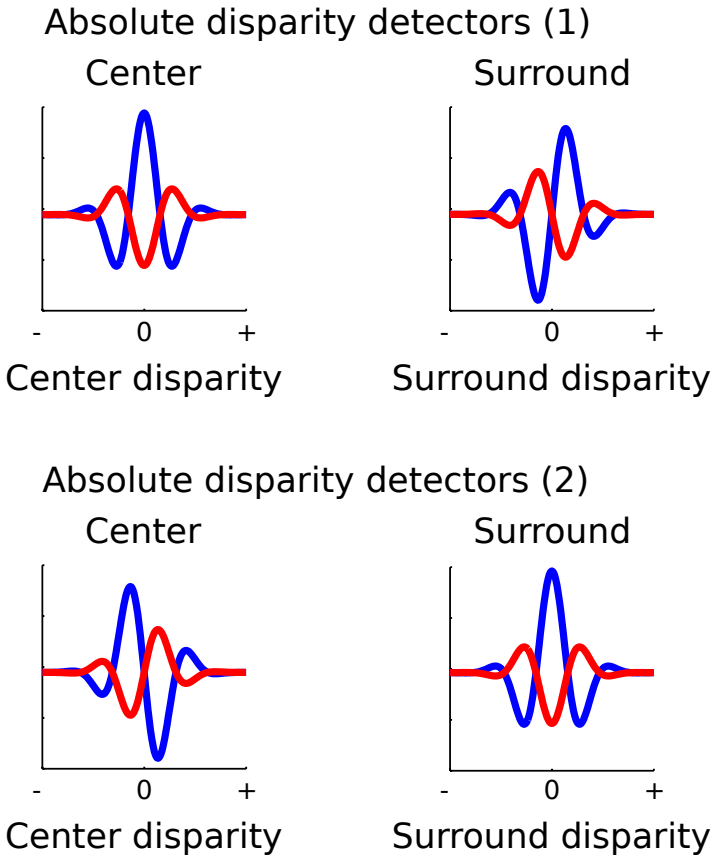

Figure 7.2: Disparity tunings of absolute disparity detectors. Blue and red curves represent disparity tunings for cRDS and aRDS, respectively.

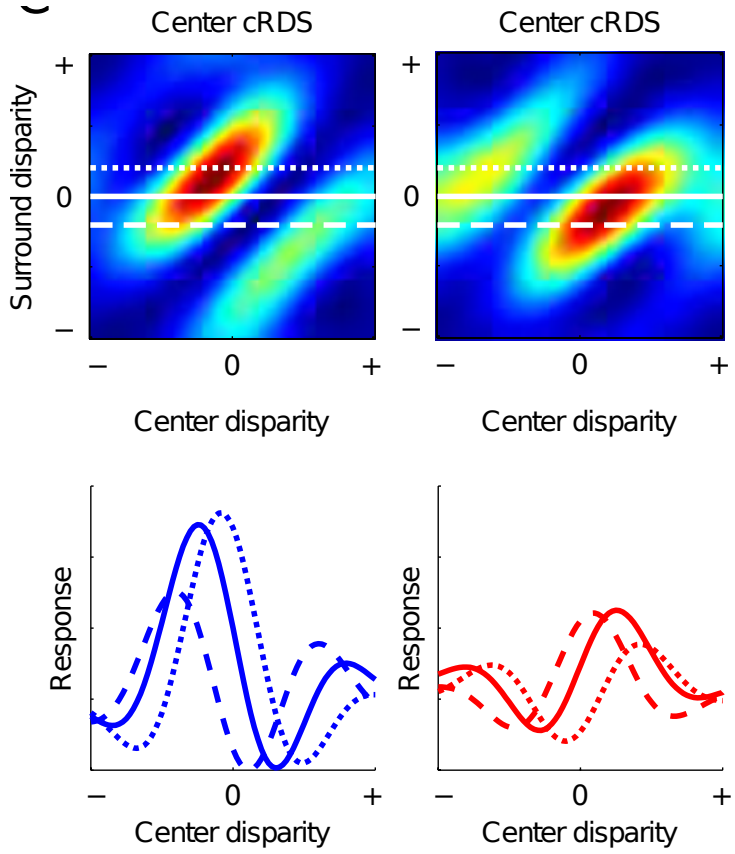

Figure 7.3: Responses of the model to center-cRDS and -aRDS stimuli. The top panels show response of the model in center-surround disparity space. The bottom panels show the model's disparity tuning curves against center disparity for different surround disparities. The left and right panels show responses of the model for cRDS- and aRDS-center, respectively.

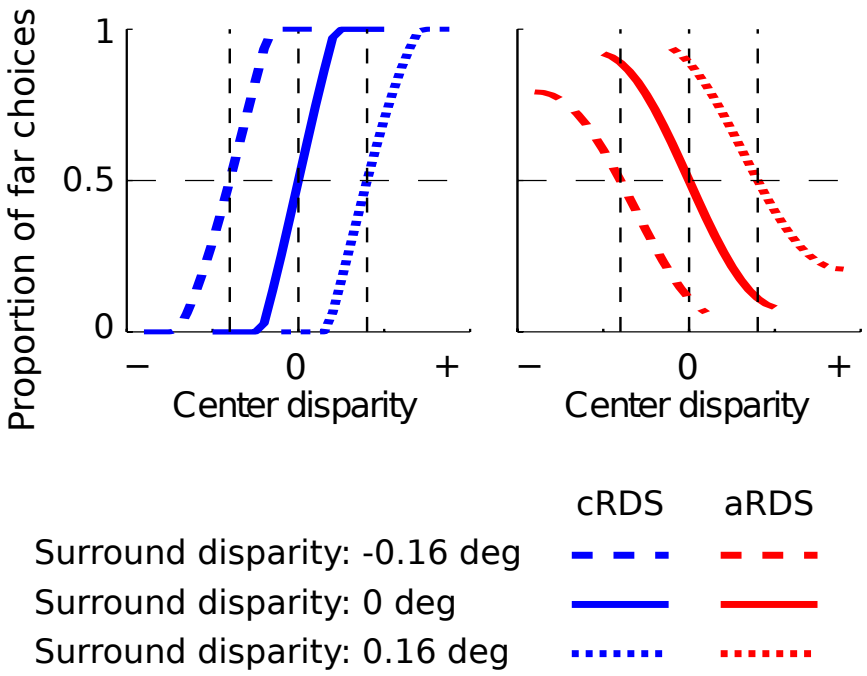

Figure 7.4: Depth discrimination performances based on the model. Proportion of far choices against center disparity derived from the model is plotted. The left and right panels show performances for cRDS- and aRDS-center stimuli, respectively. Three lines represent performances in different surround disparity conditions. I assume that the decisions of depth (‘near’ or ‘far’) are made based on comparison of near- and far-detectors. The near and far detectors are constructed with the Thomas’ model preferring near and far relative disparities, respectively.

Chapter 8: General Discussion

Summary and conclusion

In this research, I investigated the reference frame of depth supporting the depth perception produced by binocular cross-correlation. I showed that aRDSs having zero absolute and non-zero relative disparity produced reversed depth perception (Chapter 2). Since absolute disparity of aRDSs was fixed at zero, reversed depth perception was attributed to the reversal of relative disparity with respect to the surround cRDSs. The reversed depth perception in aRDSs with zero absolute disparity was not a spurious product of exploiting of surround disparity or vergence eye movements (Chapters 3 and 4). The perceived depth in aRDSs reversed according to relative, rather than absolute, disparity when the center and the surround had non-zero absolute disparities (Chapter 5). Subjects perceived reversed depth in cRDSs when the surround reference was anticorrelated (Chapter 6). From these findings of reversed depth perception in aRDSs, I conclude that correlation-based representation contributes to depth perception by forming depth in a relative, but not in an absolute, frame of reference. Correlation-based representation of relative disparity may underlie the reversed depth perception in aRDSs. I showed that such representation was generated by a physiologically inspired model of relative disparity selective neurons (Chapter 7). Taken together, I suggest that

correlation-based representation in our brain encodes relative, not only absolute, disparity (Figure 8.1).

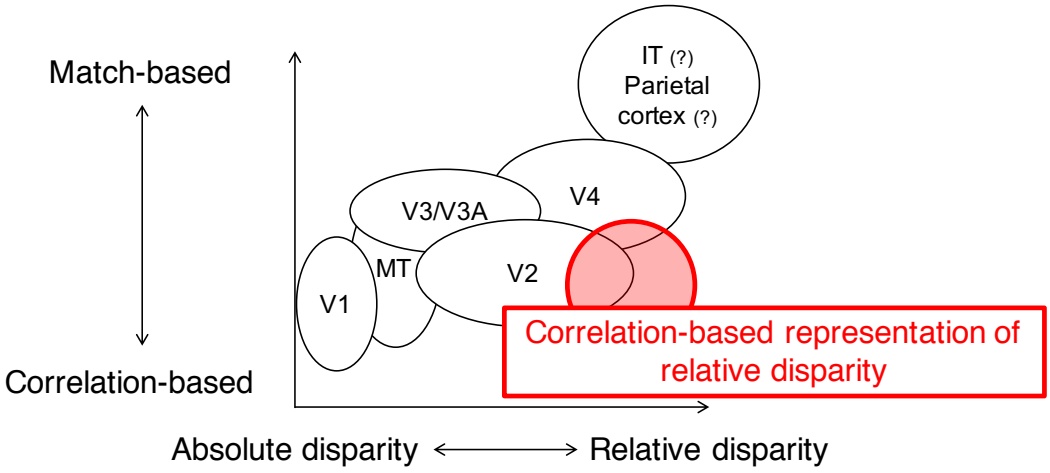

Figure 8.1: Cortical representation of disparity suggested by this research. Correlation-based representation of relative disparity can support the reversed depth perception in a relative frame of reference.

The role of depth reference in depth perception based on correlation-based representation

I showed that correlation-based representation forms depth in a relative frame of reference using an adjacent cRDS as the reference. Studies on fine stereoscopic depth perception, which relies on a relative frame of reference (Westheimer, 1979; McKee & Levi, 1987), demonstrated that a reference stimulus must have certain properties to provide a relative frame of reference. Fine stereoscopic depth perception is substantially degraded when a reference stimulus is placed far from a target stimulus (Cottereau et al., 2012b) or when a reference is binocularly uncorrelated (Prince et al., 2000; Cottereau et al., 2012a). Thus, an adjacent, binocularly correlated reference is important for creating a relative frame of reference in fine stereoscopic depth perception. Similarly, reversed depth perception in aRDSs requires an adjacent, correlated reference. Subjects perceive reversed depth in aRDSs when a correlated reference is placed adjacent to aRDSs (this study; Tanabe et al., 2008; Doi et al., 2011, 2013). The reversed depth perception is abolished when the reference is an aRDS (Julesz, 1960; Cogan et al., 1993; Cumming et al., 1998; Read & Eagle, 2000) or the target and the reference are separated by a spatial gap (Hibbard et al., 2014; Kamihirata et al., 2015). Correlation-based representation thus contributes to depth perception only when a relative frame of reference is constructed.

The notion that a relative frame of reference is necessary for contribution of

correlation-based representation to depth perception is consistent with other studies reporting the contribution of correlation-based representation on depth perception. Neri, Parker, & Blakemore (1999) presented a patch of RDSs that contains both binocularly correlated and anti-correlated dots and placed cRDSs adjacent to the patch as a reference. They employed psychophysical reverse correlation to show that depth perception reflects the disparity of anti-correlated dots in a manner consistent with the contribution of correlation-based representation; inverted psychophysical kernels were observed for aRDSs. Other studies (Read & Eagle, 2000; Hayashi, Miyawaki, Maeda, & Tachi, 2003) reported reversed depth perception for aRDSs that did not accompany any surrounding stimuli. In these experiments, however, a binocularly correlated stimulus was superimposed on aRDSs as a fixation point. This stimulus could act as a depth reference for constructing a relative frame of reference. Overall, our proposal gives a parsimonious explanation for the mixed reports on the depth perception to aRDSs and the contribution of correlation-based representation to depth perception.

Reversed depth perception for short stimulus duration

In Chapter 4, I showed that subjects also perceived reversed depth for aRDS with zero-absolute disparity with brief stimulus presentation (94 ms). This result confirms that the reversed depth perception in aRDSs is a consequence of reversal of relative depth. In the experiment in Chapter 2, changes of subjects' gaze direction to the surround cRDSs make the surround absolute disparity zero. If that happens, absolute

disparity on the center aRDS is no longer zero, and both reversal of absolute and relative depth can account for the reports of opposite depth. The results in Chapter 4 decline this alternative by demonstrating that subjects perceived reversed depth in aRDSs with zero-absolute disparity for stimulus duration shorter than the latency of human vergence eye movement to RDSs (Masson et al., 1997).

Vergence eye movement can occur after the disappearance of stimuli (Richards & Foley, 1971; Jones, 1977). The post-stimulus vergence eye movement can provide perceptual signal for depth discrimination, leaving the possibility that reversed depth perception in aRDSs is not the result of reversal of relative depth. However, I view this possibility unlikely for two reasons. First, there is no explanation that vergence-based signals produce reversal of depth perception in aRDSs. Although aRDSs elicit vergence eye movement to opposite direction from the disparity assigned at the aRDSs (Masson et al., 1997), absolute disparity on aRDSs was fixed at zero in the experiments in Chapter 2, 3, and 4. Absolute disparity on the surrounding cRDSs was changed across trials. Therefore, absolute disparity on the stimuli evoked vergence eye movement to the same direction in both cRDSs and aRDSs conditions. Second, vergence-based signals do not mediate depth perception for diplopic targets (Lugtigheid et al., 2014). Thus, it is unlikely that vergence-based signals dominate depth perception in aRDSs. For these reasons, I conjecture that vergence eye movement could not provide perceptual signals supporting reversed depth perception in aRDSs.

Reversed depth perception in center-cRDS and surround-aRDS stimuli

I found that reversed depth was perceived even when a cRDS was surrounded by an aRDS. This result indicates that construction of a relative frame of reference does not rely on specific spatial configurations. In all the experiments, subjects were instructed to answer "whether the center is nearer or farther than the surround". Thus, the center and the surround RDSs act as a target and a reference of discrimination, respectively. If the construction of a relative frame of reference depends on a specific spatial configuration of the discrimination stimuli (e.g., a cRDS should be placed around the discrimination target), the visual system cannot construct a relative frame of reference in the surround aRDS conditions. In such cases, the depth discrimination in the experiment in Chapter 6 should rely on the absolute disparity of the discrimination target (i.e., the center cRDSs). The subjects' reports of depth, however, depends solely on the relative disparity between the center and the surround. Hence, a relative frame of reference was available for the visual system when the surrounding references were aRDSs. cRDSs presented at the central regions as discrimination targets may provide a relative frame of reference. It is likely that the visual system requires correlated stimuli not only around discrimination target stimuli but somewhere in the presented stimuli for construction of a relative frame of reference.

I suggest that neurons supporting reversed depth perception in aRDSs show inverted

disparity tuning to cRDSs surrounded by aRDSs. To my knowledge, no study examines how neuronal responses to cRDSs are modified by binocular correlation around the receptive fields. From my results I suggest that neuronal response amplitude to cRDSs is decreased by anticorrelation of the surround stimulus in the visual areas supporting reversed depth perception in aRDSs. The model introduced in Chapter 7 produced neuronal disparity tuning properties that support reversed depth for center-cRDS and surround-aRDS stimuli; the model gave rise to an inverted disparity tuning curve in a cRDS surrounded by an aRDSs. This is because I did not assume any center-surround asymmetry in the model; the center and the surround disparity detectors shared identical disparity tunings and no coefficient to weight the output of the center or the surround disparity detectors was introduced in the binocular summation stage. I suggest that the neural mechanism producing inverted disparity tuning curves for aRDSs also produces inverted disparity tuning for cRDSs surrounded by aRDSs.

Cortical visual areas supporting the reversed depth perception in aRDSs

Neurons which support the reversal of relative depth for aRDSs should have two properties: selectivity for relative disparity and inverted disparity tuning to aRDSs. Neurophysiological studies have investigated selectivity to relative disparity and responses to aRDSs across many visual cortical areas in monkeys, although no study has tested the two properties in the same neurons. Here, I discuss the possible neural

correlates of reversed depth perception in aRDSs, based on the literature of electrophysiology using monkeys. I assume that the degree of correlation-matching and absolute-relative disparity sensitivity are independent at single neurons level.

V1 neurons show an inverted disparity tuning curve for aRDSs (Cumming & Parker, 1997), but lack neuronal selectivity for relative disparity in V1 (Cumming & Parker, 1999). Neuronal activity in V1 was not correlated to animals' perceptual decision in disparity discrimination tasks in terms of choice probability (Nienborg & Cumming, 2006). For these reason, I reject the possibility that V1 is the neural correlates of reversed depth perception in aRDSs.

Area MT is regarded as the first candidate of neural correlates of reversed depth perception aRDSs. First, nearly half of the neurons in MT have inverted disparity tuning curves to aRDS (Krug et al., 2004). Second, neuronal activities of MT are causally related to coarse disparity discrimination, in which correlation-based representation dominates depth perception (Uka & DeAngelis, 2006; Doi et al., 2011). Neurons in MT are, however, selective to absolute but not to relative disparity between two adjacent surfaces (Uka & DeAngelis, 2006). Therefore, my results contradict the hypothesis that MT underlies reversed depth perception in aRDSs.

Neuronal sensitivity to relative disparity emerges in V2; a subpopulation of neurons in V2 is selective to relative disparity (Thomas et al., 2002). In addition, V2 holds sensitivity to aRDSs as the same level as V1 (Allouni et al., 2005, SfN Abstract).

Neuronal activity in V2 is correlated to perceptual choices in a disparity discrimination task (Nienborg & Cumming, 2006). My model assuming direct inputs from V1-like disparity selective cells successfully produces neural representation supporting the reversal of relative depth (Chapter 7). These lines of evidence suggest that V2 may underlie reversed depth perception in aRDSs. V2 cells, however, show only modest selectivity to relative disparity. A few neurons in V2 have perfect selectivity to relative disparity (shift ratio of 1.0), and about half of the recorded cells are selective to absolute disparity (Thomas et al., 2002). Thus, to produce reversed depth perception in aRDSs with V2 cells, the visual system should exploit a small subset of neurons exhibiting relative disparity selectivity.

V4 plays a distinguished role in depth perception in a relative frame of reference. Neurons in V4 are selective to relative disparity (Umeda et al., 2007) and are causally related to fine depth discrimination, which relies on relative disparity (Shiozaki et al., 2013). A subpopulation of cells in V4 has inverted disparity tuning to aRDSs, although the response amplitude is weaker than V1 or MT (Tanabe et al., 2004; Kumano et al., 2007). These observations raise the possibility that V4 underlies reversed depth perception in aRDSs. Neural activity pooled across V4 population, however, does not produce inverted disparity tuning to aRDSs (Abdolrahmani et al., 2016). Thus, it is unlikely that V4 underlies reversed depth perception in aRDSs. To relate V4 to correlation-based representation, one needs to assume a specific read-out, such as selective pooling, from neuronal subpopulations showing inverted disparity tuning to

aRDSs.

Neuronal sensitivity to disparity in aRDSs disappears in higher visual areas such as AIP in the parietal cortex (Theys et al., 2012) or IT in the temporal cortex (Janssen et al., 2003). Thus, it is unlikely that reversed depth perception in aRDSs is the products of neural activity in such higher visual areas. Correlation-based representation is held in the early and mid-level visual areas (V1, V2, MT, V4) and thus reversed depth perception in aRDSs reflects neuronal activity in the early or mid-level stages of disparity processing.

References

Abdolrahmani, M., Doi, T., Shiozaki, H. M., & Fujita, I. (2016). Pooled, but not single-neuron, responses in macaque V4 represent a solution to the stereo correspondence problem. Journal of Neurophysiology.

Allouni, A. K., Thomas, O. M., Solomon, S. G., Krug, K., & Parker, A. J. (2005). Local and global binocular matching in V2 of the awake macaque. Society for Neuroscience Abstract, 510.8.

Barlow, H. B., Blakemore, C., & Pettigrew, J. D. (1967). The neural mechanism of binocular depth discrimination. Journal of Physiology, 193, 327-342.

Blake, R., & Wilson, H. (2011). Binocular vision. Vision Research, 51, 754-770.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. The Journal of Physiology, 211, 599-622.

Born, R. T. (2000). Center-surround interactions in the middle temporal visual area of the owl monkey. Journal of Neurophysiology, 84, 2658-2669.

Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anticorrelated stereograms: effects of spatial density and interocular delay. Vision Research, 33, 1959–1975.

Cottereau B. R., McKee S. P., Ales J. M., Norcia A. M. (2012a). Disparity-specific spatial interactions: evidence from EEG source imaging. The Journal of Neuroscience, 32, 826-840.

Cottereau B. R., McKee S. P., Norcia A. M. (2012b). Bridging the gap: global disparity processing in the human visual cortex. Journal of Neurophysiology, 107, 2421-2429.

Cumming, B. G., & DeAngelis, G. C. (2001) The physiology of stereopsis. Annual Review of Neuroscience, 24, 203-238.

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

Cumming, B. G. B., Shapiro, S. E. S., & Parker, A. J. A. (1998). Disparity detection in anticorrelated stereograms. Perception, 27, 1367–1377.

Cumming, B. G., & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. The Journal of Neuroscience, 19, 5602–5618.

Doi T., & Fujita I. (2014) Cross-matching: A modified cross-correlation underlying threshold energy model and match-based depth perception. Frontiers in Computational Neuroscience, 8:127, 1-15.

Doi T., Takano M., & Fujita I. (2013). Temporal channels and disparity representations in stereoscopic depth perception. Journal of Vision, 13(13), 26.1-25.

Doi, T., Tanabe, S., & Fujita, I. (2011). Matching and correlation computations in stereoscopic depth perception. Journal of Vision, 11(3), 1.1-16.

Erkelens, C. J., & Collewijn, H. (1985). Motion perception during dichoptic viewing of moving random-dot stereograms. Vision Research, 25, 583-588.

Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Neural encoding of binocular disparity: energy models, position shifts and phase shifts. Vision Research, 36, 1839-1857.

Gonzalez, F., & Perez, R. (1998). Neural mechanisms underlying stereoscopic vision.

Progress in Neurobiology, 55, 191-224.

Hayashi, R., Miyawaki, Y., Maeda, T., & Tachi, S. (2003). Unconscious adaptation: a new illusion of depth induced by stimulus features without depth. Vision Research, 43, 2773–2782.

Hibbard, P. B., Scott-Brown, K. C., Haigh, E. C., & Adrain, M. (2014). Depth perception not found in human observers for static or dynamic anti-correlated random dot stereograms. PLoS One, 9, e84087.

Howard, I. P., & Rogers, B. J. (1995). Binocular vision and stereopsis. Oxford University Press (Oxford, England).

Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2001) Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. The Journal of Neuroscience, 21, 9419-9429.

Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2003). At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. Neuron, 37, 693-701.

Jones, R. (1997). Anomalies of disparity detection in the human visual system. Journal of Physiology, 264, 621-640.

Julesz, B. (1960). Binocular Depth Perception of Computer-Generated Patterns. Bell System Technical Journal, 39, 1125-1162.

Kamihirata, H., Oga T., Aoki, S. C., & Fujita, I. (2015). A gap between adjacent surfaces deteriorates depth perception based on binocular correlation computation. Journal of Physiological Sciences, 65, Suppl. 1:S155.

Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. Journal of Neurophysiology, 92, 1586-1596.

Krug, K., & Parker, A. J. (2011). Neurons in dorsal visual area V5/MT signal relative disparity. The Journal of Neuroscience, 31, 17892–17904.

Kumano, H., Tanabe, S., & Fujita, I. (2008). Spatial frequency integration for binocular correspondence in macaque area V4. Journal of Neurophysiology, 99, 402-408.

Lugtigheid, A. J., Wilcox, L. M., Allison, R. S., & Howard, I. P. (2013). Vergence eye movements are not essential for stereoscopic depth. Proceedings of the Royal Society B, 281, 2013218.

McKee, S. P., & Levi, D. M. (1987). Dichoptic hyperacuity: the precision of nonius alignment. Journal of the Optical Society of America, 4, 1104-1108

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. Freeman, New York.

Marr, D., & Poggio T. A. (1979). Computational theory of human stereo vision. Proceedings of the Royal Society B, 204, 301-328.

Masson, G. S., Busettini, C., & Miles, F. A. (1997). Vergence eye movements in response to binocular disparity without depth perception. Nature, 389, 283-286.

Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. Nature, 401, 695–698.

Nider, A. (2003). Stereoscopic vision: Solving the correspondence problem. Current Biology, 13, R394-R396.

Nikara, T., Bishop, P. O., & Pettigrew, J. D. (1968). Analysis of retinal correspondence by studying receptive fields of binocular single units in cat striate cortex. Experimental Brain Research, 6, 353-372.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth

discrimination in the visual cortex: neurons ideally suited as disparity detectors. Science, 249, 1037-1041.

Qian, N., & Zhu, Y. (1997). Physiological computation of binocular disparity. Vision Research, 37, 1811-1827.

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. Nature Review Neuroscience, 8, 379-391.

Poggio, G. F., & Poggio, T. (1984). The analysis of stereopsis. Annural Review of Neuroscience, 7, 379-412.

Prince, S. J., Pointon, A. D., Cumming, B. G., & Parker, A. J. (2000). The precision of single neuron responses in cortical area V1 during stereoscopic depth judgments. The Journal of Neuroscience, 20, 3387-3400.

Regan, D., Erkelens, C. J., and Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. Investigative Ophthalmology & Visual Science, 27, 584-597.

Read, J. C. A. (2014) The place of human psychophysics in modern neuroscience. Neuroscience, 296, 116-129.

Read, J. C., & Eagle, R. A. (2000). Reversed stereo depth and motion direction with anti-correlated stimuli. Vision Research, 40, 3345-3358.

Read, J. C. A., Phillipson, G. P., Serrano-Pedraza, I., Milner, A. D., & Parker, A. J. (2010). Stereoscopic vision in the absence of the lateral occipital cortex. PLoS One, 5, e12608.

Richards, W., & Foley, J. M. (1971). Interhemispheric processing of binocular disparity. Journal of the Optical Society of America, 61, 419-421.

Shiozaki, H. M., Tanabe, S., Doi, T., & Fujita, I. (2012). Neural activity in cortical area V4 underlies fine disparity discrimination. The Journal of Neuroscience, 32, 3830–3841.

Takemura, A., Inoue, Y., Kawano, K., Quaia, C., & Miles, F. A. (2001). Single-unit activity in cortical area MST associated with disparity-vergence eye movements: evidence for population coding. Journal of Neurophysiology, 85, 2245-2266.

Tanabe, S., Umeda, K., & Fujita, I. (2004). Rejection of false matches for binocular correspondence in macaque visual cortical area V4. The Journal of Neuroscience, 24, 8170–8180.

Tanabe, S., Yasuoka, S., & Fujita, I. (2008). Disparity-energy signals in perceived stereoscopic depth. Journal of Vision, 8(3), 22.1–10.

Theys, T., Srivastava, S., van Loon, J., Goffin, J., & Janssen, P. (2012). Selectivity for three-dimensional contours and surfaces in the anterior intraparietal area. Journal of Neurophysiology, 107, 995–1008.

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. Nature Neuroscience, 5, 472–478.

Uka, T., & DeAngelis, G. C. (2006). Linking neural representation to function in stereoscopic depth perception: roles of the middle temporal area in coarse versus fine disparity discrimination. The Journal of Neuroscience, 26, 6791-6802.

Umeda, K., Tanabe, S., & Fujita, I. (2007). Representation of stereoscopic depth based on relative disparity in macaque area V4. Journal of Neurophysiology, 98, 241–252.

Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. Experimental Brain Research, 36, 585–597.

Curriculum vitae

Shuntaro C. Aoki

Laboratory for Cognitive Neuroscience

Graduate School of Frontier Biosciences

Osaka University

Office address: CiNet Building Rm. 2B2, 1-4 Yamadaoka, Suita, Osaka 565-0871 Japan
Tel: +81-6-6879-4648
E-mail: s_aoki@bpe.es.osaka-u.ac.jp

Education

M.S., Graduate School of Frontier Biosciences, Osaka University, 2009
B.A., Department of Psychology, Faculty of Letters, Kyoto University, 2007

Publications

Aoki SC, Shiozaki HM, Fujita I "Binocular cross-correlation computation produces depth perception in a relative frame of reference" in prep.

Presentations

Kamihirata Y, Oga T, Aoki SC, Fujita I "相互相関計算に基づく奥行き知覚は刺激間距離により減弱する" The 92nd Annual Meeting of The Physiological

Society of Japan (Kobe, Japan), March 2015 (in Japanese)

Aoki SC, Shiozaki HM, Fujita I “両眼相関に基づく奥行き知覚は相対視差に依存する” The 18th Meeting of Vision Science Forum (Maebashi, Japan), August 2014 (in Japanese)

Yoshioka T, Aoki SC, Fujita I “Effects of binocular luminance correlation in stimulus surround to disparity selectivity in V4 cells (V4 野細胞の視差選択性に対する刺激周辺部の両眼間輝度相関の影響)” 包括脳夏のワークショップ (Nagoya, Japan), August 2013 (in Japanese)

Higuchi Y, Tanaka Y, Aoki SC, Fujita I “先行図形呈示による後続図形の大きさ知覚への影響” The 17th Meeting of Vision Science Forum (Shiga, Japan), August 2013 (in Japanese)

Aoki SC, Shiozaki HM, Fujita I “Reversed depth perception in binocularly anticorrelated stimuli occurs in a relative depth reference frame.” Neuro 2013, the 36th Annual Meeting of the Japanese Neuroscience Society (Kyoto, Japan), June 2013

Higuchi Y, Tanaka Y, Aoki SC, Fujita I “Size illusion in the figural after-effect paradigm.” Neuro 2013, the 36th Annual Meeting of the Japanese Neuroscience Society (Kyoto, Japan), June 2013

Aoki SC, Shiozaki HM, Fujita I “Reversed depth perception in binocularly anti-correlated stimuli: Does it occur in an absolute or relative frame of reference?” Neuroscience 2012, the 42nd Annual Meeting of the Society for Neuroscience (New Orleans, LA, USA), Oct 2012

Aoki S, Shiozaki H, Fujita I “Relative disparity computation underlies the effects of surround area binocular correlation on depth perception.” Vision Science Society 10th Annual Meeting (Naples, FL, USA), May 2010

Aoki S, Shiozaki H, Fujtia I “A computational model of relative disparity discrimination underlying effects of surround areas on reversed depth perception.” Winter Meeting of the Vision Society of Japan (Tokyo, Japan), January 2010 (in Japanese)

Aoki S, Shiozaki H, Fujita I “A computational model of spatial integration of binocular disparity underlying effects of a reference area on depth discrimination.” 10th Summer Workshop on Mechanism of Brain and Mind (Sapporo, Japan), August 2009 (in Japanese)