2006

Double-Matching In Anti-Correlated Random Dot Stereograms Of Panum's Limiting Case Reveals The Interactions Among The Elementary Disparity Signals Across Scale

Hwan Sean Lee
University of Alabama at Birmingham

DOUBLE-MATCHING IN ANTI-CORRELATED RANDOM DOT STEREOGRAMS OF PANUM'S LIMITING CASE REVEALS THE INTERACTIONS AMONG THE ELEMENTARY DISPARITY SIGNALS ACROSS SCALE

by

HWAN SEAN LEE

ALLAN C. DOBBINS, COMMITTEE CHAIR
CLAUDIO BUSETTINI
TIMOTHY J. GAWNE
THOMAS T. NORTON
ANDREW E. POLLARD
DONALD B. TWIEG

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2006

DOUBLE-MATCHING IN ANTI-CORRELATED RANDOM DOT STEREOGRAMS OF PANUM'S LIMITING CASE REVEALS THE INTERACTIONS AMONG THE ELEMENTARY DISPARITY SIGNALS ACROSS SCALE

HWAN SEAN LEE

BIOMEDICAL ENGINEERING

ABSTRACT

A fundamental problem in binocular vision is to understand the rules that govern matching of features in the two eyes. We adopted Panum's limiting case (PLC) into random-dot stereograms (RDS) to determine whether our visual system permits non-unique binocular matching for stereopsis. Using PLC RDS, we found that i.) the PLC-defined depth was compatible with the disparity-defined depth, which suggests that our visual system detects the relative disparity in PLC; ii.) the upper depth limits discriminating the PLC-defined surfaces and disparity-defined surfaces were almost identical, which suggests that there is a common mechanism – correlation-based stereopsis – responsible for the depth phenomenon. Furthermore, when one of the two matches in PLC was anti-correlated (of opposite contrast), the PLC RDS induced reversed depth, which is consistent with the neuronal and visuomotor responses. The reversed depth phenomenon is an indication that ordinary disparity detection produces the anti-correlated signal in PLC that is perceived as depth, an exception for anti-correlated stimuli.

To explain why we experience depth reversal in the anti-correlated PLC RDS but not in the conventional anti-correlated stimulus, we simulated the response of binocular disparity-sensitive neurons. We found that the disparity estimates in the correlated RDS were coherent across scale, whereas the disparity estimates were dispersed in anti-

correlated RDS. In contrast, the dispersion of the disparity estimates was limited to a range of spatial frequencies in the anti-correlated PLC RDS. Evidently, confined estimates were coherent enough to yield the sign of depth, but not coherent enough to support surface integration. In addition, we found that limiting the spatial frequency channels available for disparity detection reduced the depth discrimination accuracy.

Based on these observations, we propose a second stage mechanism that combines the elementary signals of the disparity detectors to produce the disparity map that resolves transparency in depth surfaces. The second stage mechanism incorporated the facts that 1) the removal of the monocular components from the elementary signal sharpens the disparity tuning; 2) inter-scale summation improves the signal-to-noise ratio and broadens the working disparity range. This model network further predicted the degree of decoherence in the anti-correlated RDS and the reversed depth phenomenon.

DEDICATION

I wish to dedicate this thesis to my grandmother, Jung-im, for her unstinting love and support.

ACKNOWLEDGEMENTS

My appreciation to Allan Dobbins, my advisor, for his guidance and support. He always reserved room for liberal minds and open discussion. The members of my dissertation committee, Andrew Pollard, Claudio Busettini, Donald Twieg, Thomas Norton, and Timothy Gawne have generously given their time, efforts, and expertise to better my work. I thank them for their contribution and their support. I am also grateful to my parents for their generous love and support. My greatest thanks go to Mika and Lynn for their remarkable patience and love.

TABLE OF CONTENTS

Page

TABLE OF CONTENTS (Continued)

Page

TABLE OF CONTENTS (Continued)

Page

LIST OF FIGURES

Figure Page

BINOCULAR MATCHING IN STEREOPSIS

ACCOUNTING FOR THE DEPTH SENSATION IN PANUM'S LIMITING CASE USING RANDOM-DOT STEREOGRAMS

LIST OF FIGURES (Continued)

Figure Page

LIST OF FIGURES (Continued)

Figure Page

CHAPTER 1

BINOCULAR MATCHING IN STEREOPSIS

The world that we see is three-dimensional although the retinal images from which we derive such perception are in two dimensions. Interpretation of the two retinal copies of the scene and recovering the third dimension, depth, relies on several visual cues. Among those, the most precise and robust cues come from our ability to combine images from the two eyes—binocular vision. Due to the lateral separation of the two eyes, the world seen by each retina is slightly different. Images of some scene elements are absent in one retina, whereas other common elements fall onto different locations. In figure 1.1, the bull's eye symbol is occluded by surface B and therefore visible only to the left eye. The human visual system is known to exploit the distance between the occluding surface and the partially occluded object from the retinal images to estimate the relative depth (Nakayama and Shimojo, 1990). On the other hand, the star symbol in figure 1.1 is visible to both eyes but images of the star falls on different locations in the two retinae. The difference between the two retinal positions is called retinal disparity or binocular disparity. The binocular disparity provides precise information about which points are closer and which are farther, and how much closer or farther they are. These two types of depth cues play important roles in stereoscopic depth perception.

Whether there are corresponding images in the two eyes or not, strategies of finding either the matched features or the unmatched features are required for the visual

system to retrieve the depths encoded on to the two retinae. The problem of finding the correct binocular match that originated from the same object is referred to as the correspondence problem. Because image features in the scene are often unique in luminance contrast, color, shape or orientation, it has been suggested that our visual system solves the correspondence problem by matching up those image features from the two retinae. In the case of random-dot stereogram (RDS, Julesz, 1960), however, image features are all identical dots, and matches can be established between any dots in the two images. Despite the computational complexity, our visual system easily finds the correct solution and one sees the depths in the RDS (figure 1.2)

Several rules have been proposed for matching binocular images in stereo algorithms (Marr, 1982) based on the observations of the real world. One of the most popular strategies that stereo matching algorithms adopt to reduce the number of spurious matches is called the uniqueness constraint (Marr and Poggio, 1979). The uniqueness constraint states that a single image component cannot be matched to more than one component in the other eye. The uniqueness constraint significantly reduces the number of false matches in RDS. However, this constraint poses a problem for the binocularly unmatched features.

The theme of this manuscript is to consider the governing rules of binocular matching in human stereo vision. In particular, we consider two cases that pose questions of how our visual system deals with the correspondence problem. The first case is the classical Panum's limiting case (PLC, Panum, 1858 in Ogle, 1962). In PLC, one sees a pair of vertical lines in one eye and a single vertical line in the other eye (figure 1.3). When they are binocularly combined, one sees two apparent lines in different depths.

The depth arises from binocular cue(s), but it is not clear whether we experience stereopsis in violation of unique matching, or whether another mechanism that does not require binocular correlation is responsible. We examined the compatibility between the PLC-defined depth and the correlation-based depth in RDS to investigate the common mechanism accounting for the depth impression.

The second case is the anti-correlated matches—the binocular match of opposite contrast features. The anti-correlated matches are considered to be false matches because they produce binocular lustre rather than consistent depth impression. In Chapter 3, we use a novel anti-correlated RDS employing the PLC configuration. The anti-correlated PLC RDS produced a striking impression of depth that was opposite to the assigned stimulus depth. We performed quantitative psychophysics to measure the depth reversal and provide an explanation based on physiological evidence.

These two cases represent serious violations of the existing fundamental constraints in computational stereo matching algorithms. In Chapter 4, we simulate these two cases in a disparity detector model that was proposed based on physiological data (Ohzawa, DeAngelis and Freeman, 1990). We found unique response patterns in simulation in response to the stimulus that give rise to depth impression and to the stimulus that does not. Through a psychophysics experiment, we validated one aspect of the response pattern in the simulation which coincided with the depth discrimination accuracy.

Finally, the observations are incorporated into a disparity mapping model network ,and a simple combining mechanism is proposed. The mechanism encompasses physiologically plausible computation and significantly improves the disparity mapping

in RDS. The produced map showed that the proposed mechanism solves the transparency of surfaces in depths in RDS and predicts the reversed depth phenomenon.

Specific introductions are included in each chapter. Some terminology related to stereopsis or stereoscopic depth perception is provided in the appendices.

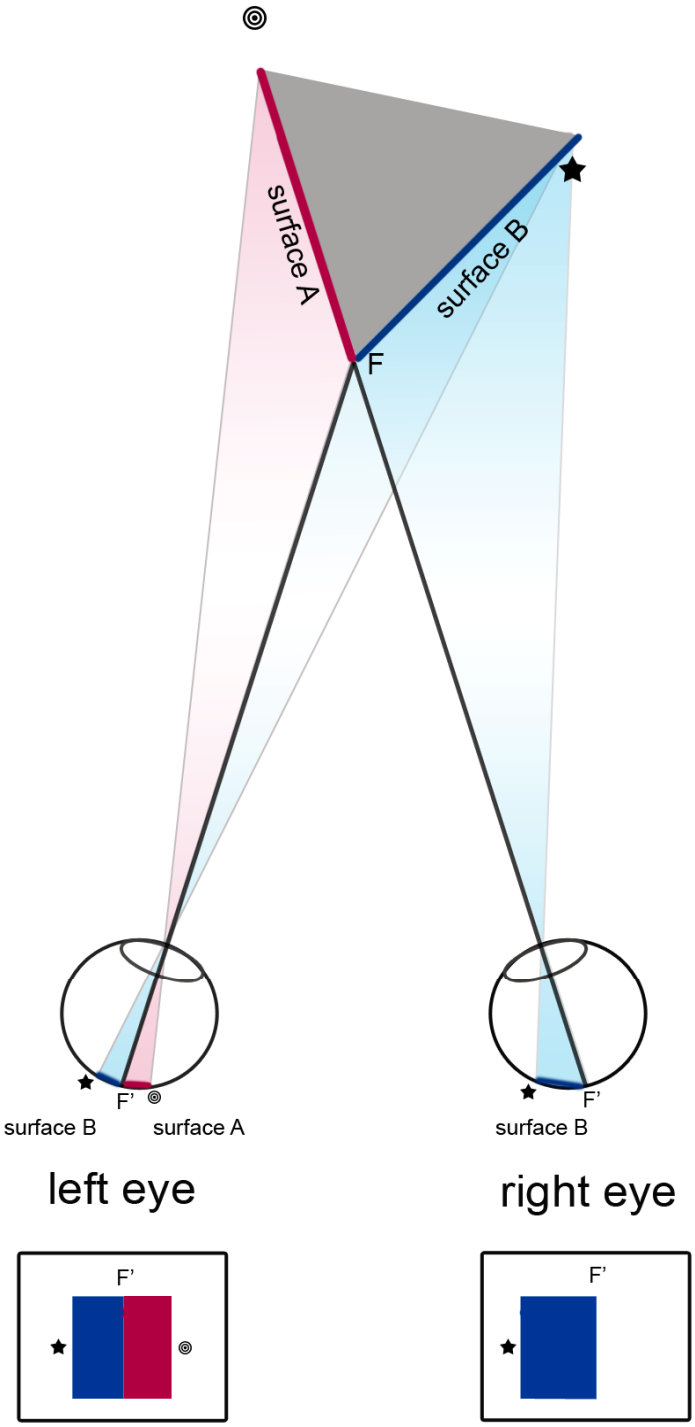


Figure 1.1. Stereoscopic depth cues. The surface A and the bull's eye symbol at the far left are only visible to the left eye. The star symbol on the far right corner of the triangular pillar is visible to both eyes but the locations where the images of the star fall on to are different in left and right retinae. (F: fixation point, F': fovea, star: corresponding image)

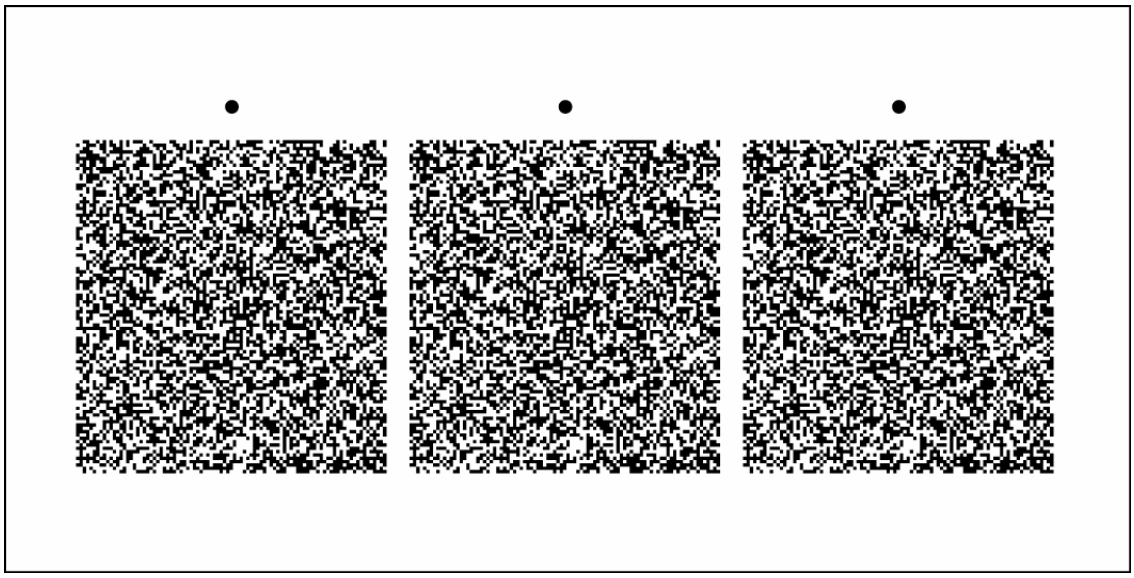


Figure 1.2. A typical random-dot stereogram introduced by Julesz (1960). The left and the middle pair and the middle and the right pair have opposite depth configuration so that the central square in the stereogram appears in opposite depth (nearer or farther than the surrounding square) when fused. Free-fusing the dots above the three images first helps seeing the hidden square.

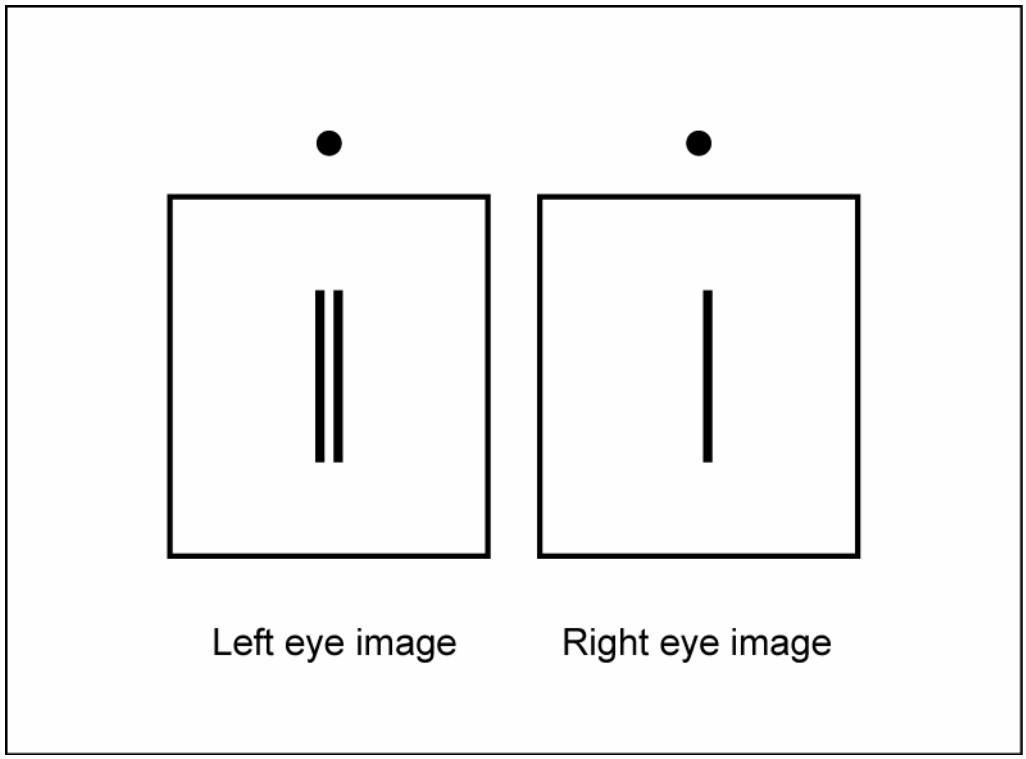


Figure 1.3. Panum's limiting case. When the two images are binocularly presented, one sees two lines at different depths. One may appreciate the depth impression by free-fusion, a way of seeing depths in a stereogram with bare eyes. For free-fusion, viewers fixate either closer or farther than the depth of this page.

CHAPTER 2

ACCOUNTING FOR THE DEPTH SENSATION IN PANUM'S LIMITING CASE USING RANDOM-DOT STEREOGRAMS

In Panum's Limiting Case (Panum, 1858 in Ogle, 1962), one views a pair of vertical lines in one eye and a single vertical line in the other eye (figure 1.3). When the two images in the two eyes are binocularly combined, one sees two lines at different depths. The left line appears farther away when the line pair is presented to the left eye. The right line appears farther away when the line pair is presented to the right eye. The depth difference between the two lines increases as the distance between the line pair in one eye increases within a small range. Panum's limiting case (PLC) naturally occurs when one line is occluding the other line for one eye but not for the other (figure 2.1). The depth sensation in PLC arises from binocular cue(s) but what accounts for the depth sensation is not clear. We first list three major explanations which try to account for perception of depth direction and depth magnitude in PLC.

2.1 Theories Accounting for Depth Impression in PLC

2.1.1 Double-matching Theory (Double-fusion theory)

Images of an object at depths different from the fixation plane fall on different locations on the two retinae. The difference between the locations of the retinal images is called binocular disparity and gives rise to a compelling sensation of depth. To recover depths from the two retinal images, the images projected from the same object must be

identified from each retina and combined. Double-matching theory (Hering, 1879 in Howard and Rogers, 2002) assumes that the two lines in one eye have a common binocular match with the single line in the other eye. In figure 2.2, the binocular disparity of the first match between the two lines at the corresponding positions (match 1) is zero. The binocular disparity of the other match (match 2)—the match between the right line in the left eye image and the single line in the right eye image—is equal to the distance between the line pair in the left eye. According to double-matching theory, the relative disparity (disparity difference) between the two matches is equivalent to the degree of separation between the line pair. The binocular disparity provides information about which points are closer and which are farther, and how much closer or father they are, information about both the depth order and the depth magnitude.

The double-matching theory accounts for the most precise depth impression; however, it bears one problem. One of the most popular strategies that stereo matching algorithms have adopted is the so-called uniqueness constraint (Marr and Poggio, 1976). The uniqueness constraint states that a single image component cannot be combined simultaneously with more than one component in the other eye. The uniqueness constraint significantly reduces the number of spurious matches when the two images contain multiple identical components.

2.1.2 Occlusion/camouflage Configuration Theory

An opaque nearby surface creates an occluded zone behind that surface. This zone is subdivided into a binocular occlusion zone that is visible by either eye and two monocular occlusion zones that are visible only to the left or right eye. In figure 2.3, the

gray-faded zone behind the surface is one of the monocular occlusion zones. When an object is placed within a monocular occlusion zone, the image of this object on the retina is always situated on the nasal side of the image that is projected from the occluding surface.

There is another visual zone in front of the surface, in which an object of identical texture may be camouflaged. This zone is also subdivided into the binocular and monocular camouflage zones. When an object is within one of the monocular camouflage zones (pink shade in figure 2.3), the image of this object is always situated on the temporal side of the image of the binocular surface on the retina. The configuration of the monocular image and the binocular surface provides information about whether the object is in front of or behind the surface. This is referred to as "configuration theory." Although monocular occlusion/camouflage can provide the depth order between a binocular surface and the monocular features, this depth cue cannot provide precise depth estimates. This is because the image of the monocular feature on the retina can be projected from any point in depth along the line of sight within the monocular occlusion/camouflage zone (positions of green dots in figure 2.3, for example).

2.1.3 Da Vinci Stereopsi

Conventional stereopsis involves establishing the proper matches between the images in the two eyes so that the binocular disparities are determined for those matching features. Nakayama and Shimojo (1990) described a new type of stereopsis in which the unmatched image features play a role to give rise to depth sensation based on differential occlusion. In their stimulus (figure 2.4a), a monocular bar next to a binocular surface

appeared at different depth when the bar position was consistent with being occluded by the binocular surface. The perceived depth was quantitatively related to the separation of the bar from the surface, which approximated the minimum of the possible depths along the line of sight that is occluded for the other eye (figure 2.4b). The phenomenon was named "da Vinci stereopsis," crediting the first description of monocular occlusion to Leonardo da Vinci.

In the study of Nakayama and Shimojo (1990), when the monocular bar was placed on the nasal side of the binocular surface in the stimulus, observers reported seeing no consistent depth. This is the monocular camouflage configuration and expected to give rise to “Near” depth perception according to the configuration theory or double-matching theory. The camouflage configuration is more stringent than the occlusion configuration because the features must have identical luminance, color, and texture statistics in relation to the background and it is the rarest of cases in natural scene. Nakayama and Shimojo (1990) called such configuration ecologically ‘invalid’ and further found that such configuration yields binocular luster, switching of the perception between left and right eye’s view (Shimojo and Nakayama, 1990). In summary, da Vinci stereopsis accounts for some degree of quantitative depth perception when and only when the configuration of the lines is consistent with occlusion.

2.2 Conflicting Evidence in PLC

Certain aspects of depth phenomenon in PLC exclusively support or disprove one or more explanation(s). One of the main arguments concerning the depth phenomenon in PLC is whether the degree of separation between the line pair in one eye

is quantitatively related to the apparent depth between the two lines. If proven, this will rule out the configuration theory. In addition, if the perceived depth in PLC closely matches that of the binocular disparity, it would support the double-matching theory. Another question is whether there is valid depth impression when one of the line pair in one eye that is not fixated is placed on the nasal side of the fixated one in the stimulus (camouflage configuration). The evidence against this would exclusively support the da Vinci stereopsis.

The evidence concerning these arguments, however, is diverse and inconclusive. At small degrees of separations, some researchers reported metric (Gettys and Harker, 1967; Gillam, Cook, and Blackburn, 2003) or parametric (Westheimer, 1986) depth impression in PLC in both occlusion and camouflage configurations. Other researchers have disputed the quantitative depth phenomenon in PLC when the fixation is strictly controlled (Shimono, Tam and, Nakamizo, 1999). Although many researchers reported depth impression in both occlusion and camouflage configurations, some other researchers reported that the camouflage configuration yields no consistent or veridical depth impression (Shimono et al., 1999; Nakayama and Shimojo, 1990). In addition, the depth experience among participants in PLC stimulus varies widely, and naive observers are reported as experiencing difficulties seeing depth in PLC (Gettys and Harker, 1967; Westheimer, 1986). However, practice which does not necessarily require feedback improves depth perception significantly (Westheimer, 1986).

2.3 Adoption of PLC into random-dot stereograms

In our pilot study, we observed similar inconsistencies in the classic PLC stimulus. For example, the depth was more vivid in one direction (farness/nearness) than the other while the best direction varied among participants, and the depth was often difficult to determine when the degree of separation was greater than 20~30 min of arc. In such cases, having the fixation shifted from one match to the other helped discriminate the depths (Westheimer, 1986).

The present experiments are designed to clarify some of these issues in PLC. The general aim is to discriminate among several theories of the depth phenomenon in PLC. To determine which account is responsible, we adopted the PLC configuration into a random-dot stereogram (RDS). RDSs are noise images that contain systematic binocular correlations. When the images in the two eyes are properly registered, they give rise to the perception of surfaces in depths (Julesz, 1960). The RDS of the PLC configuration (PLC RDS) we created is similar to the one invented by Braddick (Marr, 1982), in which observers perceive two transparent surfaces, one in front of the other, when binocularly combined (figure 2.5). The differential depths of the two surfaces were very compelling, and the quality of depth was consistent among participants. First-time viewers did not require practice and readily saw the depth planes in our experimental setting.

2.3.1 Depth Direction and Magnitude in PLC RDS

In PLC, one can choose to fixate on one of the lines in apparent depths. Choosing which one to fixate on does not switch the depth order of the closer and the farther lines, but puts the unfixated line off the fixation plane. In our PLC RDS, the observer sees one

surface at the depth of the fixation cross (on-fixation plane) and the other surface in front of or behind the fixated surface (off-fixation plane).

The relative depth of the off-fixation plane depends on the direction and the degree of the separation in the PLC configuration (figure 2.6). The direction of separation determines whether the off-fixation plane appears in front of or behind the on-fixation plane (figure 2.6a). In figure 2.7, the upper and the lower halves contain random dots whose separation is in opposite directions. Viewers see the two off-fixation planes in the upper and lower halves at different depths simultaneously, one half in front of and the other half behind the common surface at the depth of the fixation cross.

The depth difference between the two transparent surfaces in our PLC RDS is related to the degree of separation between the dot pair in PLC configuration (figure 2.6b). In the PLC RDS of figure 2.8, the degree of separation in the lower half is twice the degree of separation in the upper half. Both of the two half planes appear either in front of or behind the surface at fixation depth, but appear at different depth magnitudes.

2.3.2 Compatibility Between PLC RDS and Double-layered RDS

The magnitude of the depth between the two surfaces in the PLC RDS appears compatible with that of a RDS composed of two layers of random-dots at different depths (double-layered RDS). In the PLC RDS, the magnitude of depth is related to the degree of separation between the dot pair in one eye, whereas in the double-layered RDS, it is determined by the relative disparity between the two random-dot layers. Figure 2.9 illustrates the compatibility of the perceived depth in the two configurations. In figure 2.9, the upper half contains PLC RDS and the lower half contains double-layered RDS. The

degree of separation in the PLC configuration is equal to the relative disparity between the two random-dot layers in double-layered RDS. When binocularly combined, the on-fixation planes and the off-fixation planes in the upper and lower halves appear to meet seamlessly at the equator. Except for the texture change, the border is indistinguishable in terms of depth.

The depth impression in conventional RDS is attributed to establishing proper matches between the left and right eye images and the detection of binocular disparities among the matches. The similarity between the PLC-defined depth and the disparity-defined depth in the two configurations suggests that a common mechanism may be possibly responsible for the depth sensation in these two configurations. If true, the two configurations would share similar limitations in seeing the depth surfaces. Furthermore, the quantitative depth sensation in the two cases should be comparable. To test these two hypotheses, we measured how closely one can match the depth differences in the two configurations to determine if the perceived depth in PLC configuration is truly compatible with that of disparity-defined RDS.

2.4 EXPERIMENT I: Quantitative Depth Measure of PLC RDS

The depth between the two surfaces in PLC RDS increased as the degree of separation between the dot pair in one eye increased. We measured the magnitude of the depth between the surfaces by matching it to the depth between the two surfaces in the double-layered RDS. The participants were instructed to check the proper fixation with the nonius line method before brief presentation of the stimulus (200 ms).

2.4.1 Methods

General settings. Three individuals (including the author) participated in this experiment. The participants had normal or corrected-to-normal vision and had previous experience viewing RDSs. Each participant sat in a dark room at a distance of 65 cm from the display, which was centered at eye level with the aid of a chin-rest.

Displays were generated from custom-made programs by using OpenGL® graphics libraries on a Silicon Graphics Indigo2 workstation and were viewed on a 21" flat-screen color monitor (SGI GDM-5011P, resolution of 1024 x 768, 96 Hz refresh rate; 48 frames/s per eye) through ferroelectric liquid crystal shutter goggles (Crystal Eyes2, Stereographics Corp., 30% open transmittance and open:closed transmittance ratio of 1000:1). This setting was used in all the following psychophysics experiments.

RDS stimulus design. The stimulus contained two types of RDS, the double-layered RDS in the upper half and the PLC RDS in the lower half of the display. Figure 2.10 shows an example of this stimulus, and figure 2.11 shows the arrangement of the depth surfaces. Both RDSs were narrow strips extending 34 degrees of visual angle (dva) in width and 3.3 dva in height and contained equal numbers of white random-dots (4 min of arc x 4 min of arc) covering 12% of a gray background. The two half RDSs were separated by 6 min of arc blank zone. There was a red fixation cross (20 min of arc in width and 12 min of arc in height) in the center of the display, which was presented intermittently before stimulus presentation. The fixation cross was accompanied with nonius lines (10 min of arc) directly above and below to monitor proper fixation.

Procedures of experiment. The depth between surfaces in PLC RDS was kept constant while the participant repeated adjusting the depth between the surfaces in the double-layered RDS to match the magnitude of the two depths. The degree of separation in the PLC configuration was 10, 12, 14, 16 and 18 min of arc, and each value was repeated twice in random order for each participant. The direction of separation was always nasal in the stimulus so that the off-fixation plane appeared in front of the on-fixation plane. For each degree of PLC separation, the relative disparity in double layered RDS started with 6 min of arc in uncrossed direction (left eye image shifted leftward, and the right eye image shifted rightward).

A keystroke initiated a trial displaying a fixation cross in the center and two RDSs, one in the upper half and the other in the lower half of the screen. Each RDS contained a single layer of random-dots in this initial display, which appeared as a single surface at the depth of the fixation cross. After 500 ms, a second random-dot layer was overlaid on to the first layer for 200 ms. Stimulus presentation was brief to reduce the effect of vergence movements on depth perception. The fixation cross was turned off during this 200 ms.

The second layer in the lower half contained dots whose positions were in the lateral vicinity of the random-dots of the first layer (PLC RDS), whereas the second layer in the upper half contained a different set of random-dots whose positions were independent of the first layer (double-layered RDS). The second overlaying layers in both configurations appeared as transparent surfaces in front of the first layers. The first layer at the fixation depth remained for another 200 ms after the second layers disappeared. The participant then used another keystroke to introduce a disparity change of 2 min of

arc in the desired direction. With a short delay, another run of 500ms-200ms-200ms stimulus presentation started. The participants were allowed to make as many changes as needed to match the two second surfaces as close in depth as possible. When the depth differences of the two surfaces in the upper and lower RDSs appeared to be matched, the participants used a keystroke to report the matched disparity. Another value of separation in PLC configuration was set to the PLC RDS, and participants repeated the described procedure.

2.4.2 Results

The perceived depths from the two configurations were highly comparable. Figure 2.12 shows the average response of three participants. The relative disparities that the participants matched to the given separations between 10 to 18 min of arc were very close to but slightly smaller than the degree of separation in the PLC configuration. The small deviation from the exact match of the magnitude was almost constant (2 min of arc) and consistent across three participants. The PLC-defined RDS gave rise to a depth that was highly compatible to the disparity-defined RDS.

2.4.3 Camouflage Configuration and the Constant Depth Reduction in PLC

We chose nasal PLC separations in the stimulus (camouflage configuration) for depth matching, although the depth impression was equally good for both directions when viewing duration was not limited (see figure 2.7). However, with brief presentation, participants had greater confidence in the camouflage configuration than the occlusion configuration.

The metric depth impression in our PLC RDS with camouflage configuration contradicts the predictions of both configuration theory and da Vinci stereopsis. The theory of da Vinci stereopsis considers the camouflage configuration as invalid for depth perception. Furthermore, the quantitative depth that is closely matched to that of disparity-defined depth suggests that there is a common mechanism for estimating depth in these two configurations.

When the PLC-defined surface is joined next to the disparity-defined surface, the boundaries merge and the two surfaces appear as one flat seamless surfaces (see figure 2. 9). With a small gap between the two RDS types, however, scrutinizing each surface reveals there is a small difference in depth. Our results also reflected this small difference (figure 2.12). Because the difference was constant in the average of the three participants, we verified that it was not due to an artifact of our stimulus software.

In PLC, the disparity detectors tuned for low spatial frequencies would detect the dot pair as one entity if the frequency was lower than twice the distance between the two dots. For these detectors, the centroid of the two dots causes the disparity estimate to be half the disparity obtained with a fine disparity detector. We think the reduction in the perceived depth magnitude in PLC RDS is due to the disparity pooling of these two signals, a phenomenon which has been observed under other circumstances (Parker and Yang, 1989).

2.5 EXPERIMENT II: The Upper Depth Limit of the PLC RDS

From the result of the preceding depth matching experiment, we hypothesized that there is a common mechanism for perceiving depths in the PLC-defined surfaces and

the disparity-defined surfaces. If this were, the two cases would share a similar upper depth limit.

We created disparity-defined gratings and PLC-defined gratings in RDS, which appeared as horizontal stripes at two alternating depths when binocularly combined (figure 2.13). These gratings elicit the perception of vertical motion when the assigned depths are shifted either upward or downward. We asked the subjects to determine the direction of motion while increasing the depths of the stripes away from the fixation depth. Human observers are very keen to determine the direction of motion. The participants readily detected the correct direction of motion although the details of the depth structure were vague. Perception of motion in these RDSs requires stereopsis, and there was no monocular cue for motion; thus, the disparity threshold for motion detection reflects the upper depth limit at which one can discriminate the depth-defined regions in RDS.

2.5.1 Methods

RDS stimulus design. The depth-defined gratings were created with two configurations: i.) PLC RDS and ii.) double-layered RDS. In each configuration, there was an extra flat plane that appeared to be sandwiched between the alternating stripes when the RDS was binocularly fused. This plane was composed of random-dots whose positions were identical in both eye images (zero disparity). In PLC RDS (figure 2.14a), the alternating stripes were created by placing additional dots with small lateral shifts with respect to the dots in the flat plane. Half of the dots in the stripes was plotted in the left eye image and the other half was plotted in the right eye image to balance the total number of dots in the

two eye images. The stripes that contained dots positioned at the nasal side in the stimulus appear in front of the flat plane, and the dots positioned at the temporal side appear behind the flat plane. In the double-layered RDS (figure 2.14b), the alternating stripes contained random-dots whose positions were either shifted rightward in the left-eye image and shifted leftward in the right-eye image (crossed disparity) or shifted leftward in the left-eye image and shifted rightward in the right-eye image (uncrossed disparity) for the same amplitude.

The left and right eye images contained an equal number of black random-dots (4 min of arc x 4 min of arc) covering 12% of a median gray background and were replotted every 100 ms. The display extended 34 degrees in width and 20 degrees in height and included a red cross (20 min of arc in width and 12 min of arc in height) in the center of the display. The cross had a yellow center point, and nonius lines (10 min of arc) were placed above and below the cross to monitor fixation. The height of the stripes was 3.3 dva and moved either upward or downward at a speed of 5 dva/sec.

The offsets between the two planes (the separation in PLC or the relative disparity) create flanks next to the lateral edges of the flat plane. This square-wave pattern moves with the stripes. To avoid such motion cue that is available without stereopsis, the monocular flanks were clipped to fit to the width of the display.

Procedures of experiment. Participants initiated a trial with a keystroke which displayed a fixation cross in the center followed by the dynamic RDS stimulus that lasted for 3000 ms. The direction of motion of the disparity-defined stripes (upward or downward) was chosen randomly. The participant reported the direction of motion with a keystroke.

The initial disparities of the stripes were 12 min of arc in both crossed and uncrossed directions. This amplitude increased with two successive correct judgments and decreased with one incorrect judgment by 4 min of arc ('two-up/one-down' staircase procedure). A trial was completed with 15 reversals between correct and incorrect judgments.

2.5.2 Results

At small disparities, the stripes at alternating depths appeared crisp with well-defined horizontal edges. As the disparity increased, the solidness of the stripes diminished and eventually collapsed so that the depths of the individual stripes and the boundaries became ambiguous. The participants were no longer able to determine the direction of motion.

The results of the PLC RDS and double-layered RDS are shown in figure 2.15. The mean detection limit, which reflects the upper depth limit of perceiving the disparity-defined gratings, was about 44 min of arc for double-layered RDS. The mean detection limit for PLC RDS was 46 min of arc. For every participant, the magnitudes of the detection limits in the two configurations were very close. The similarity in the upper depth limits of the two cases provides another piece of evidence that supports a common mechanism responsible for depth perception in PLC-defined surfaces and disparity-defined surfaces.

2.6 EXPERIMENT III: Relationship Between the Stripe Height and the Depth Limit

In a pilot study, when we used a different height of the stripes, the upper depth limit was significantly altered. There are two possible aspects that would affect the upper depth limit of seeing the depth-defined motion: i.) the number of depth edges (proportional to spatial frequencies); and ii.) the contiguous area of the constant disparity elements. By varying the height of the disparity-defined stripes, we tested these alternatives in the double-layered RDS. Reducing the stripe height while keeping the display dimensions constant would increase the number of effective depth transitions. If the former is responsible for the change in depth limit, the upper depth limit would be increased, whereas, if the latter is responsible, the upper depth limit would be reduced.

2.6.1 Methods

The RDS stimulus used in this experiment was similar to the double-layered RDS in the preceding experiment described in section 2.4.1 (figure 2.14b), except that the dot density was 10%. The heights of the stripes were 1.67 dva, 3.33 dva, and 6.67 dva, which were repeated twice in random order. The experimental procedures were identical to the preceding upper depth limit experiment.

2.6.2 Results

Figure 2.16 shows the results from this experiment. The mean upper limits were 39 min of arc, 51 min of arc, and 60 min of arc for 1.67 dva, 3.33 dva, and 6.67 dva, respectively. For all participants, the upper disparity limit of seeing the disparity-defined grating increased as the height of the stripes in the grating increased. The relationship

between the stripe height and the upper limit appeared to be approximately linear on the semi-log plot, i.e., each doubling of the period resulted in about 10 min of arc increase in the upper depth limit on average (figure 2.17). Strip height greater than 13.34 dva was not tested due to the monitor size.

The initial aim of this experiment was to supplement a relevant fact that the upper depth limit for seeing the depth-defined structures varies with the dimensions of the depth structure. In any case, the results suggest that for depth contrast (the contrast required to detect motion), spatial pooling of the contiguous disparity signal is more important than the number (or total length) of the disparity boundaries.

2.7 The Quantitative Depth Impression in Camouflage Configuration in PLC

In this chapter, we provided two kinds of evidence that support double-matching in PLC. First, the quantitative depth impression in PLC-defined surfaces was closely matched to that of disparity-defined surfaces. Second, the upper depth limit for seeing PLC-defined depth and disparity-defined structure was approximately equal. Together, these suggest there is a common mechanism processing the depth information in these two cases.

We used the camouflage configuration in the depth matching experiment, which ruled out da Vinci stereopsis being responsible for the depth impression because placing the monocular image in the nasal side of the binocular image in the stimulus is invalid according to the da Vinci stereopsis mechanism. The configuration theory also does not account for the quantitative depth impression we observed in our PLC RDS. Therefore,

we conclude that a conventional stereopsis mechanism is responsible for depth perception in PLC RDS.

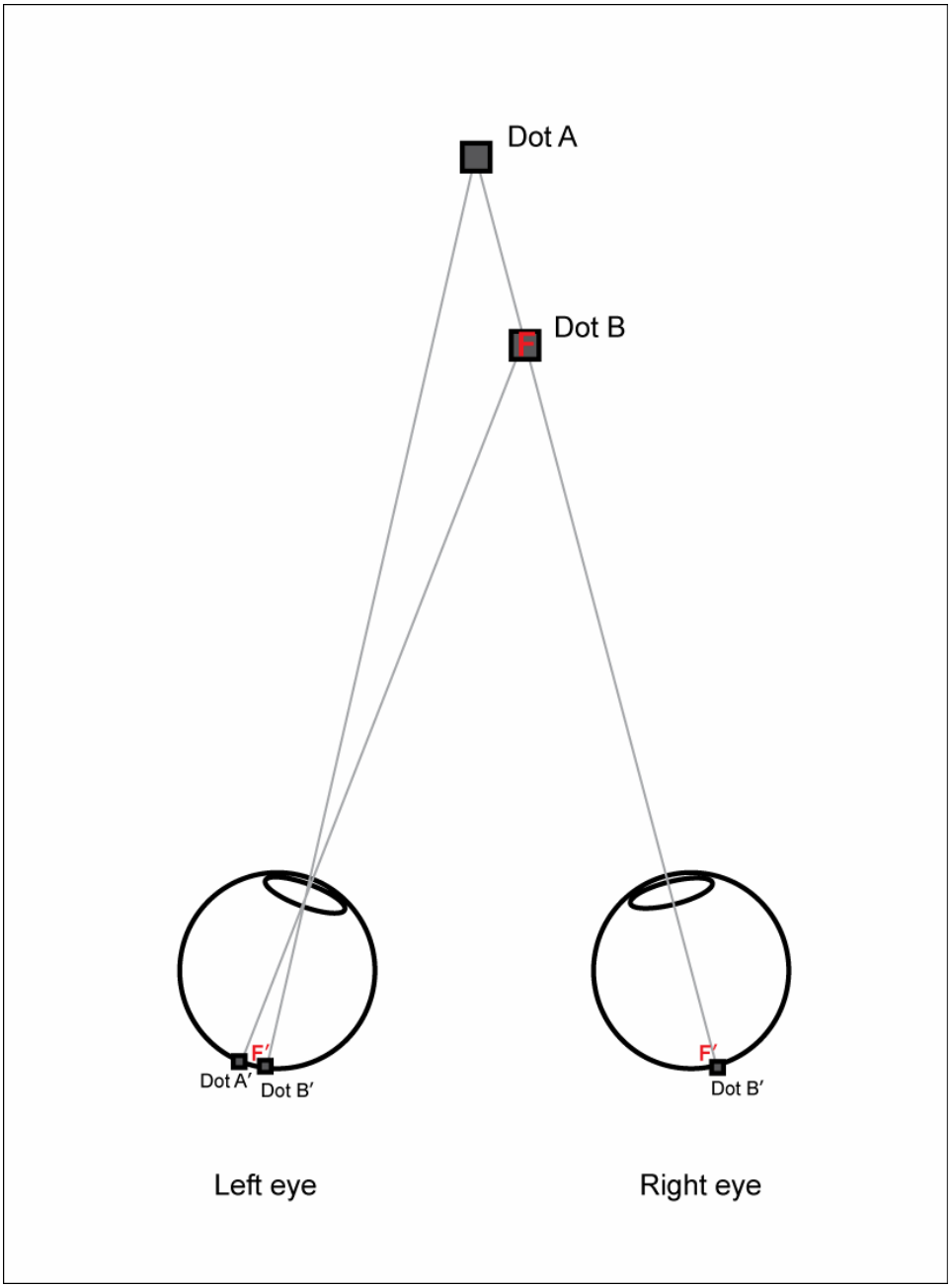


Figure 2.1. PLC and its retinal projections. PLC resembles a naturally occurring situation in which one line occludes the other line for one eye but not for the other. In this figure, the eyes fixate on point F, and dot A is occluded by dot B, and therefore, is invisible to the right eye.

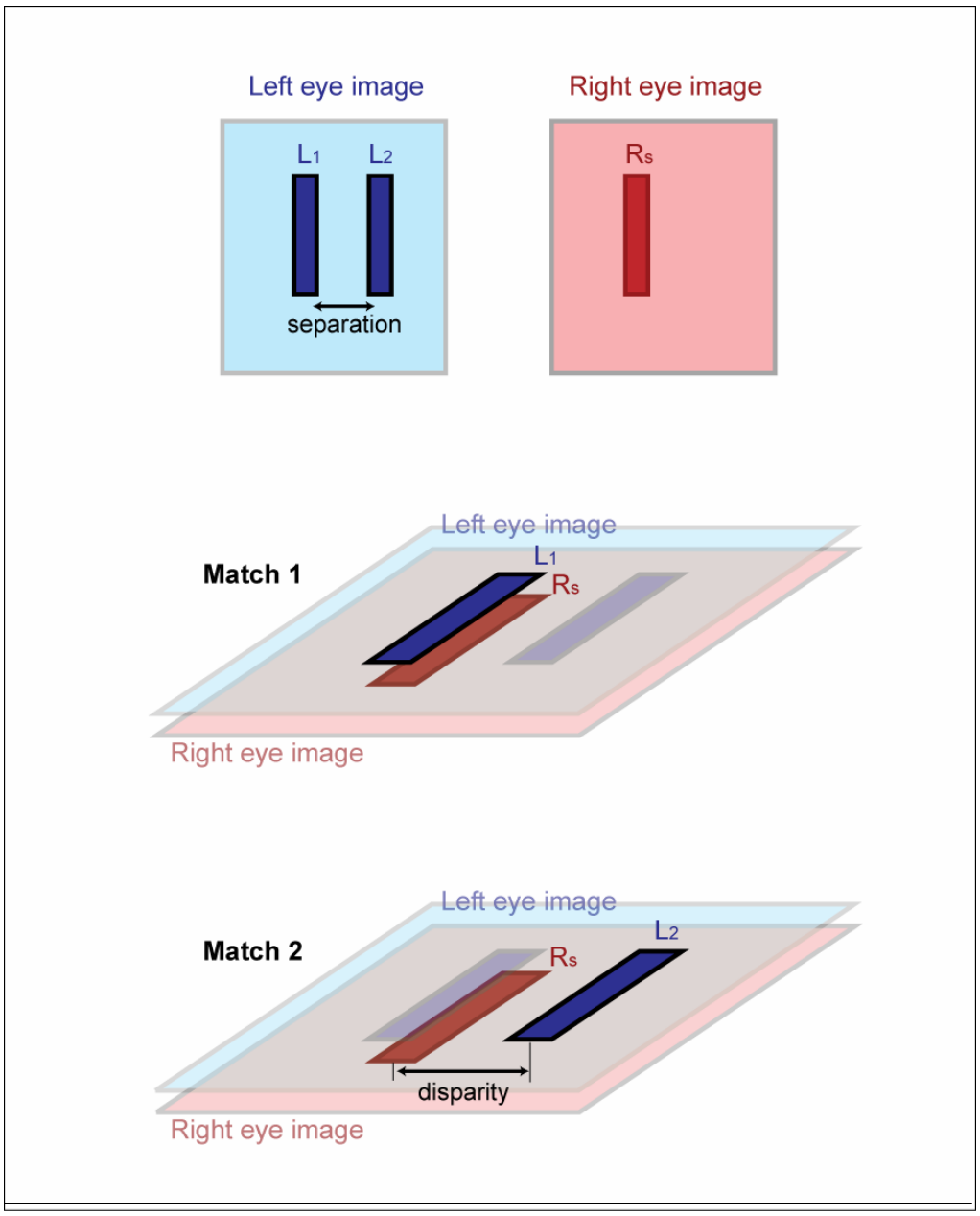


Figure 2.2. The binocular disparities in the double-matching theory. According to the double-matching theory, there are two binocular matches. One is between the lines L_1 and R_s whose disparity is zero and the other is between the lines L_2 and R_s whose disparity is equal to the distance between L_1 and L_2. The distance between the two lines in one eye (degree of separation) is, therefore, equivalent to the relative disparity between the two binocular matches (the difference between the two disparities).

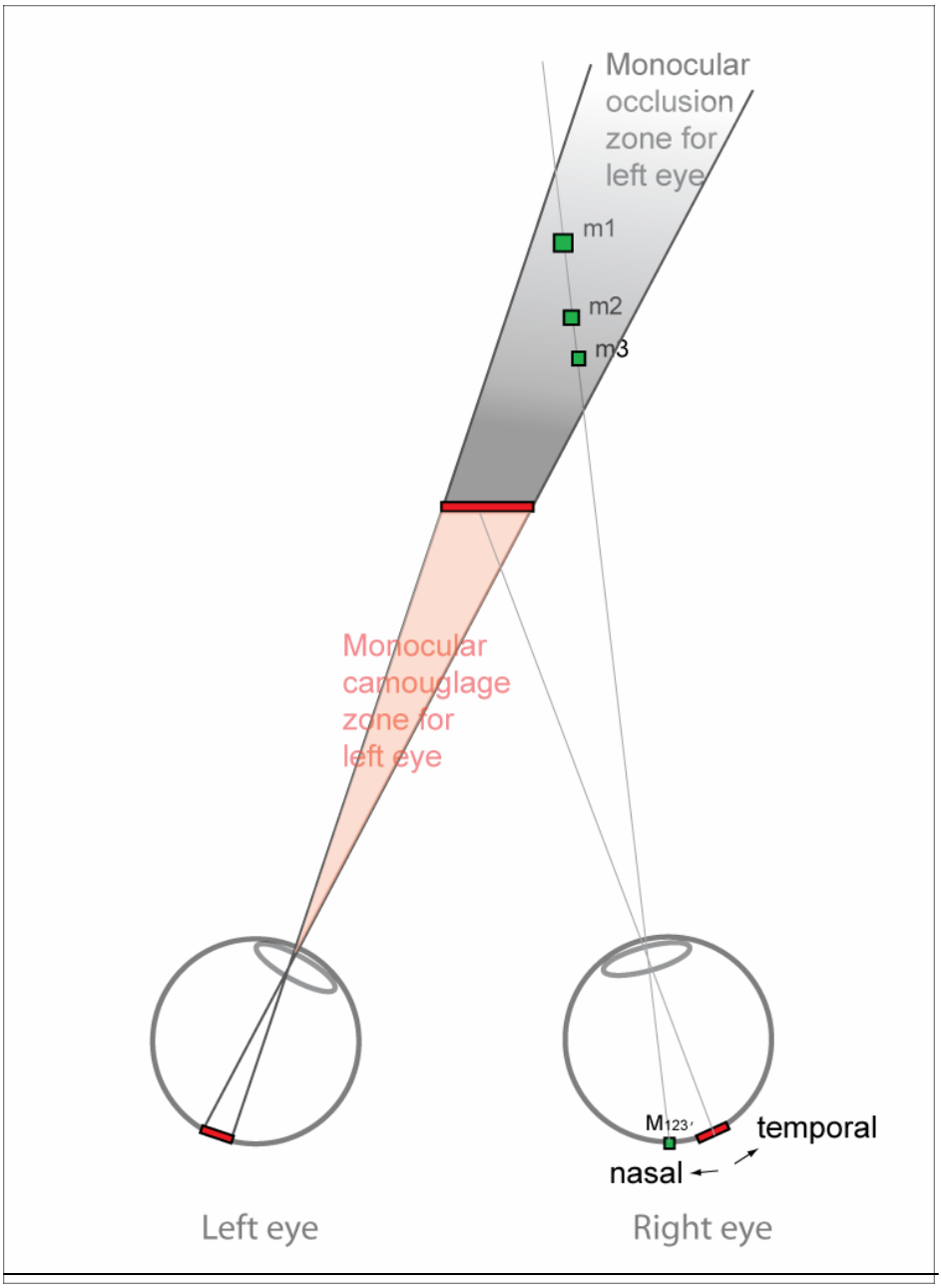


Figure 2.3. Monocular occlusion and camouflage zones in configuration theory. The object in the monocular occlusion zone always projects its image on to the nasal side of the image projected from the occluding surface. The depth between the surface and the object, however, is ambiguous because the monocular image (M_{123}') on the retina could have originated from any depth, for example, m_1, m_2, or m_3, along the line of sight within the monocular occlusion zone.

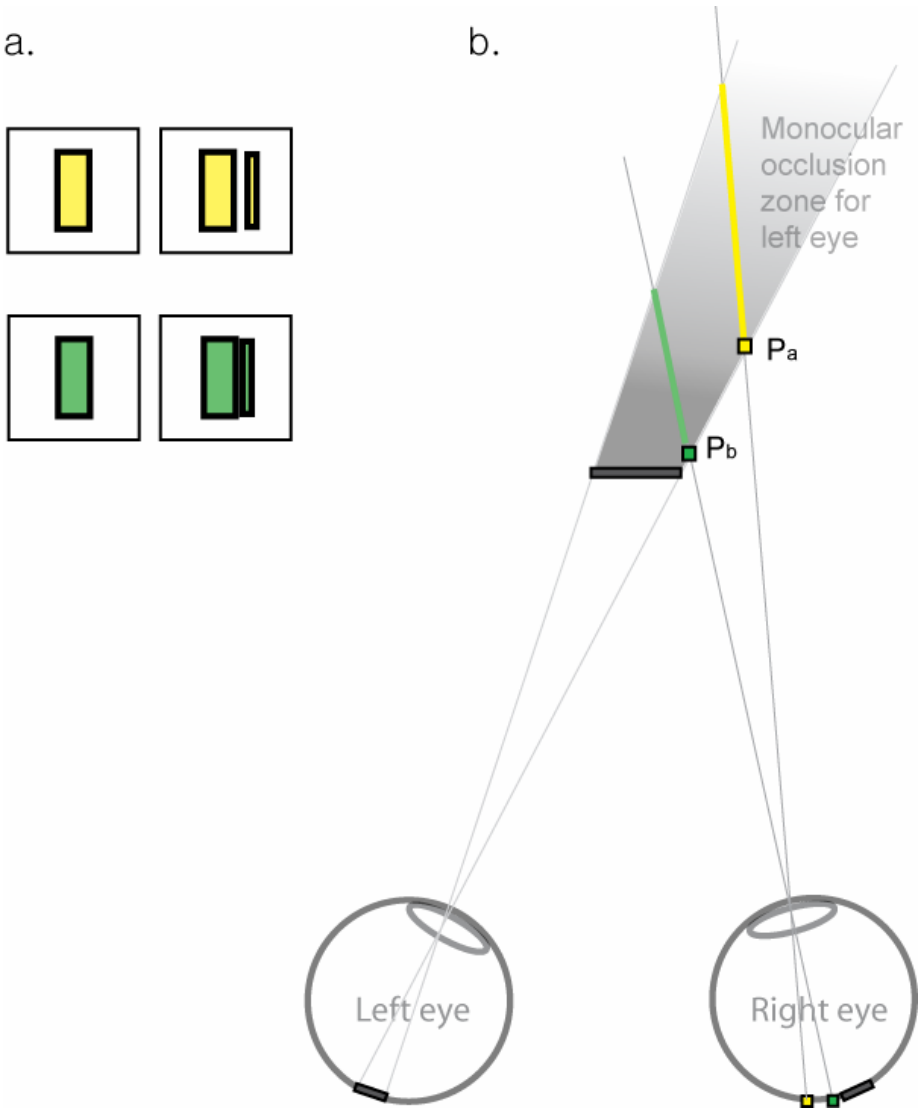


Figure 2.4. Crude quantitative depths in da Vinci stereopsis. a. The color-coded stereograms are examples of typical da Vinci stereopsis stimulus in which the distance between the monocular bar and the binocular surface differ. b. Possible arrangements of the objects in the scene that yield retinal images that are similar to the stimulus in Figure 2.5.a. The green and yellow dots (P_a and P_b) represent the smallest possible depth that can be inferred from the location of the retinal images. The precise depth can not be determined from the information available from the retinal location, but the minimum and maximum depths are defined by the depth constraint lines (green and yellow lines behind the colored dots) in relation to the width of the binocular surface and the distance between the eyes and the surface.

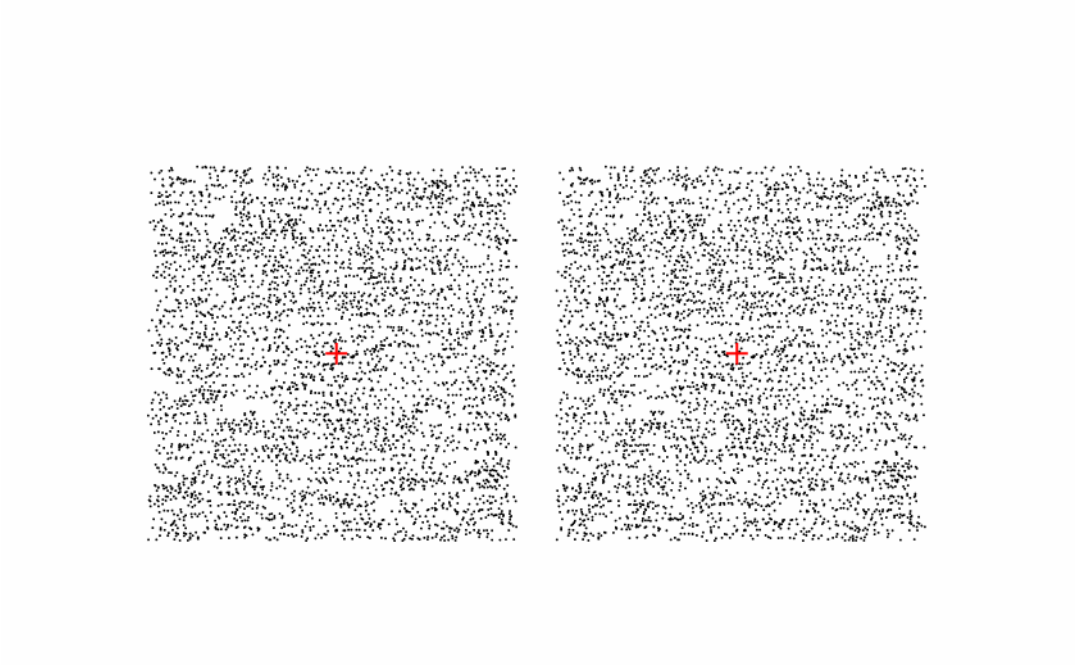


Figure 2.5. Random-dot stereogram adopting PLC configuration. When the stereogram is binocularly combined, one sees two transparent surfaces, one in front of the other. Optimal viewing distance is about 60 cm or greater.

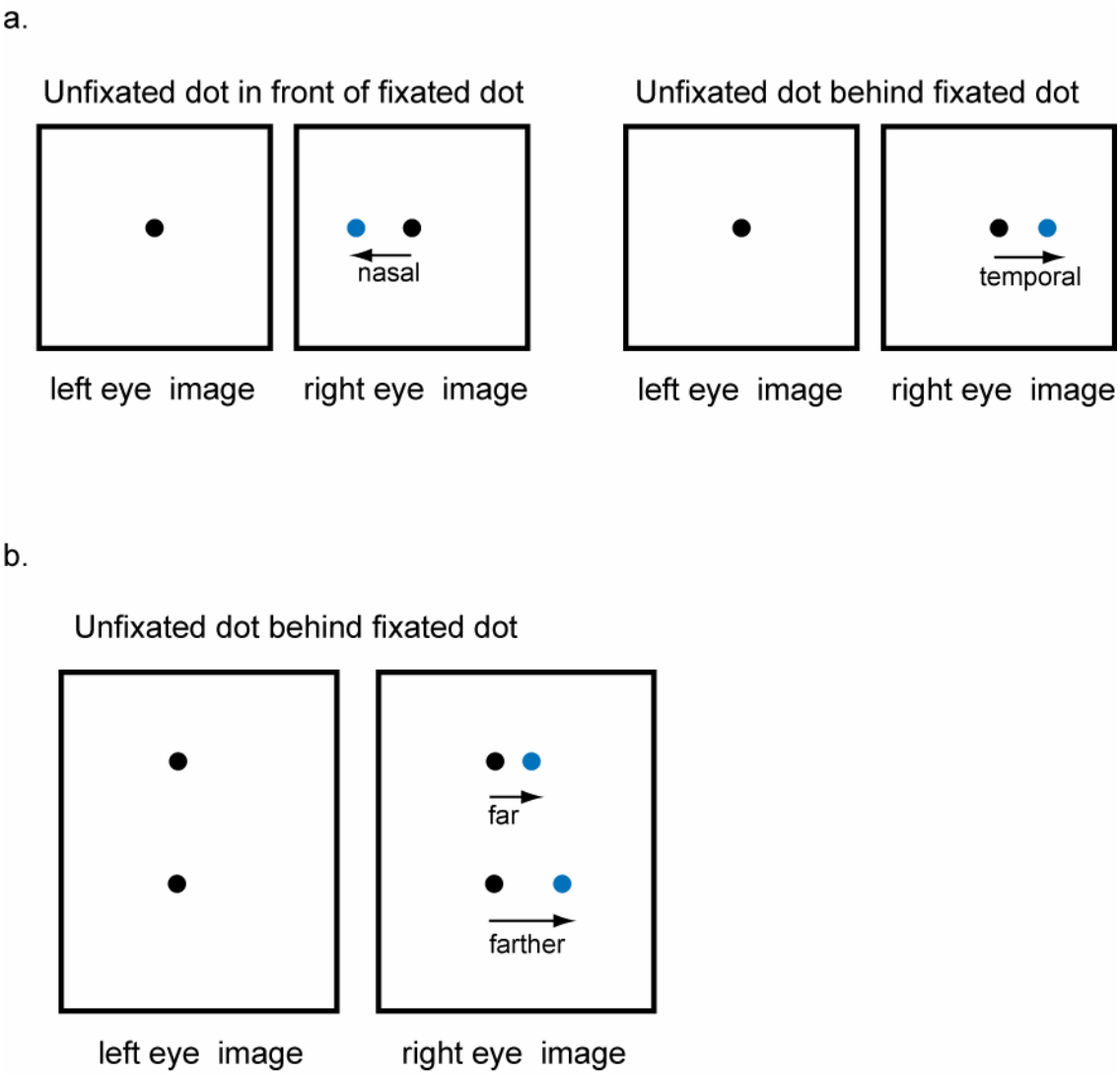


Figure 2.6. Direction and degrees of separation in PLC dot pair in our random-dot stereograms. a. When one fixates on the two black dots in the stereogram, the black dots situates at the corresponding positions on the retinae. The unfixated blue dot on the nasal side of the fixated dots appears in front of the fixated black dot. When it is on the temporal side, it appears behind the fixated black dot. The depth of the transparent surface at off-fixation depth follows this rule. b. The farther away the unfixated dot is the farther off the depth of the off-fixation plane from the on-fixation plane in random-dot stereogram.

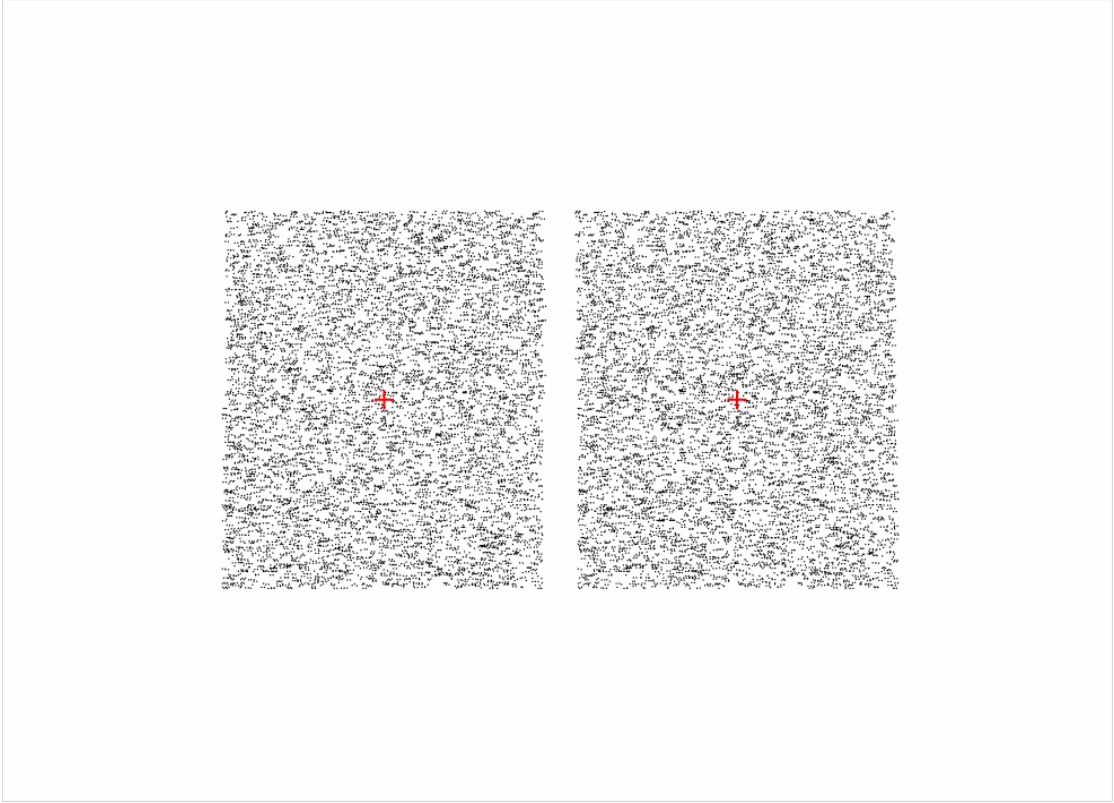


Figure 2.7. The direction of separation in PLC RDS. The degrees of separation in the upper and lower halves of the stereogram are equal, but the directions of the separation are opposite. The viewers see one of the halves in front of and the other behind the common surface at the fixation depth. Optimal viewing distance is about 60 cm.

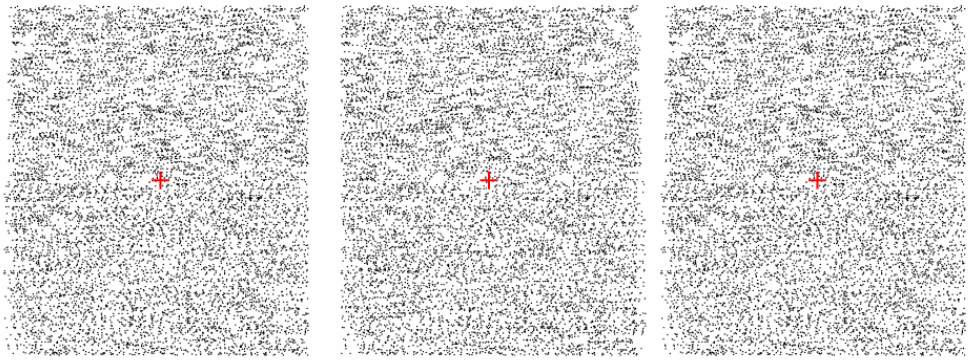


Figure 2.8. The degree of separation in PLC RDS. The direction of separation in the upper and lower halves of the stereogram pair (either the left and middle or the middle and right half images) is identical, but the degree of separation in the lower half is twice that of upper half. The viewer sees two off-fixation planes in different depths but both at the depths either closer or farther than the on-fixation plane. The direction of separation in the two stereogram pairs (the left and middle or the middle and right half images) are opposite so that in one pairing both the off-fixation planes appears in front of the on-fixation plane whereas, in the other pairing both the off-fixation planes appear behind the on-fixation plane. Optimal viewing distance is about 60 cm or greater.

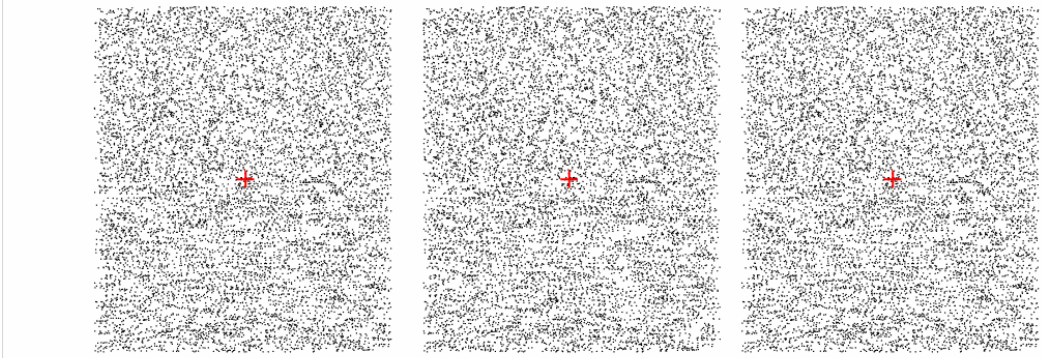


Figure 2.9. Compatibility of the perceived depths in the PLC RDS and double-layered RDS. The stereogram contains a double-layered RDS in the upper half and the PLC RDS in the lower half. The depth impression of the two transparent surfaces in the two configurations appears comparable. One generally distinguishes no depth discontinuity at the border of the two RDSs. Optimal viewing distance is about 40 cm.

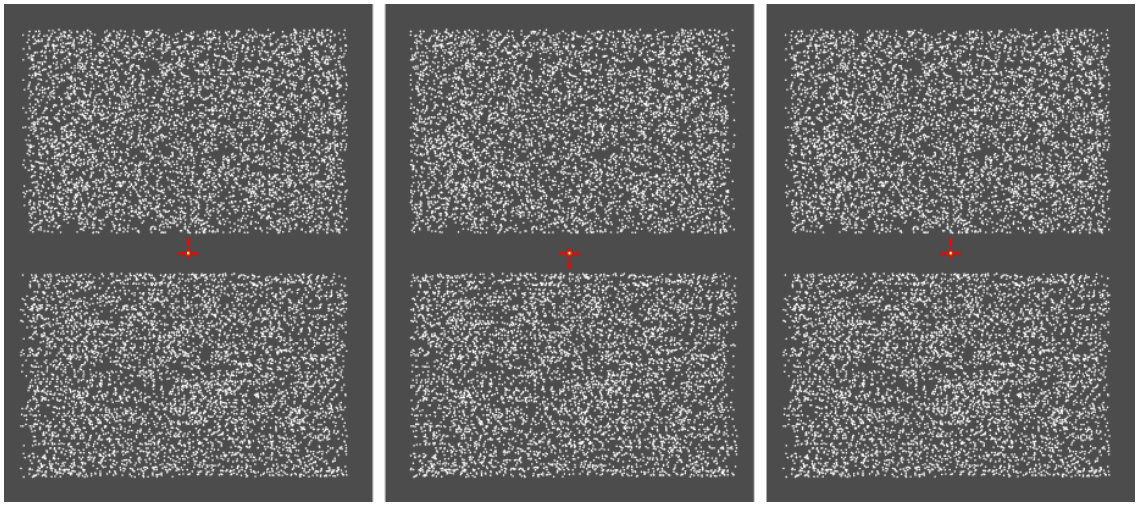


Figure 2.10. Illustration of the stimulus used in the depth matching experiment. The RDS in the upper half is a double-layered RDS and the RDS in the lower half is PLC RDS. The left and the middle image pair is designed for viewers who over-diverge to fuse the stereogram and the middle and the right pair is designed for the viewers who over-converge. The off-fixation surfaces always appeared in front of the zero disparity surface in the experiments. In actual settings, the dots were plotted beyond the width of the lateral boundaries of the screen so that the distinguishable monocular flanks at the lateral edges of the RDSs were masked. The aspect ratio of the stimulus in the experiments was different from this illustration. They were narrow strips as shown in figure 2.11. Optimal viewing distance is about 60 cm.

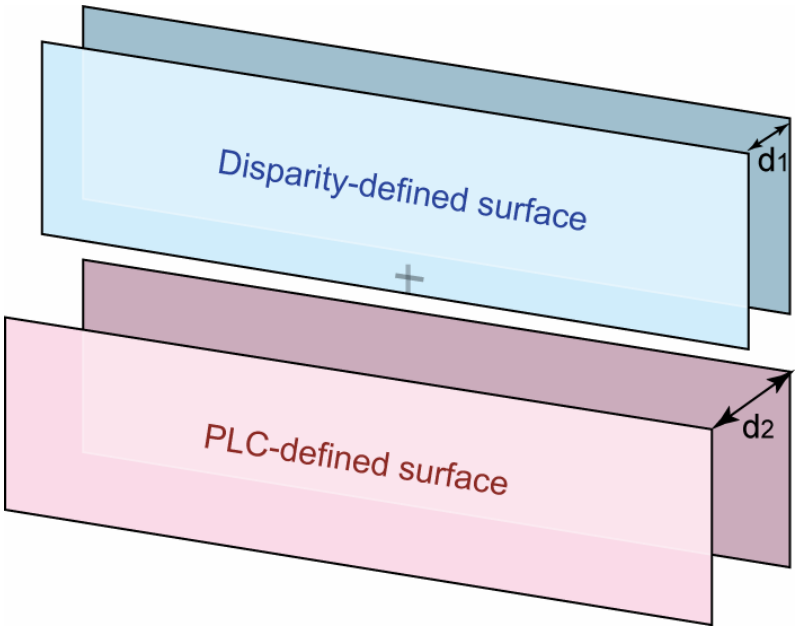


Figure 2.11. A schematic illustration of the surfaces in the RDS stimulus. Participants adjusted the relative disparity of the two random-dot layers (d_1) in the upper half RDS so that the depth difference of these surfaces matched to that of the PLC surfaces (d_2) in the lower half RDS.

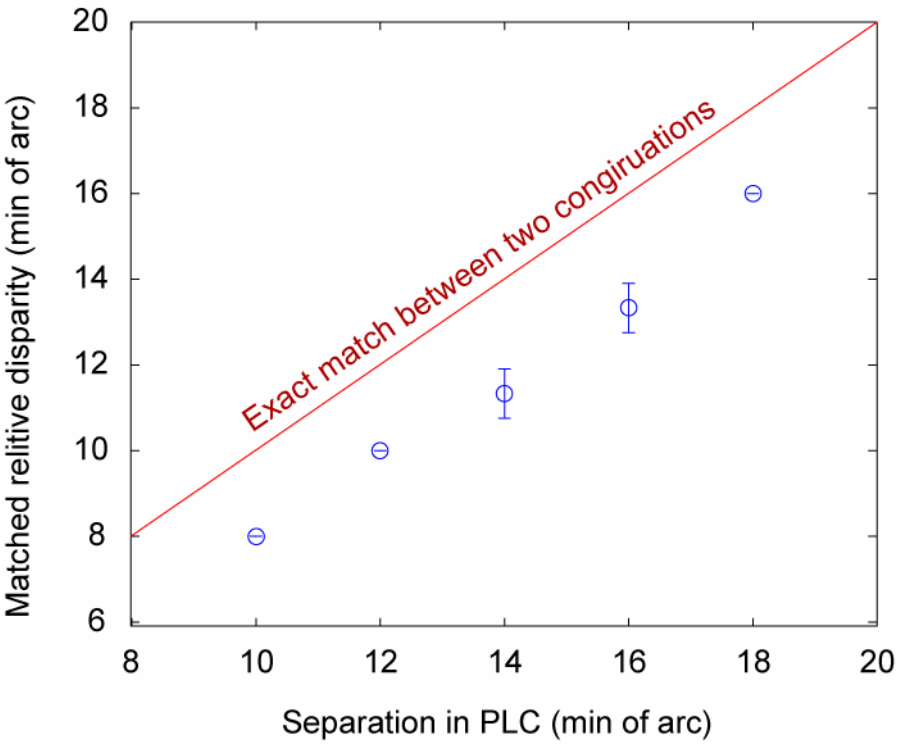


Figure 2.12. Responses of the participants in depth matching experiment. By changing the relative disparity between the two layers of the double-layered RDS, the depth between the two surfaces was matched to that of the PLC RDS. The blue open circles represent the average relative disparities of three participants plotted as a function of the separation in PLC configuration. The red diagonal line denotes the ideal matches when the degree of PLC separation is exactly compatible with the relative disparity. The error bars denote the standard deviation.

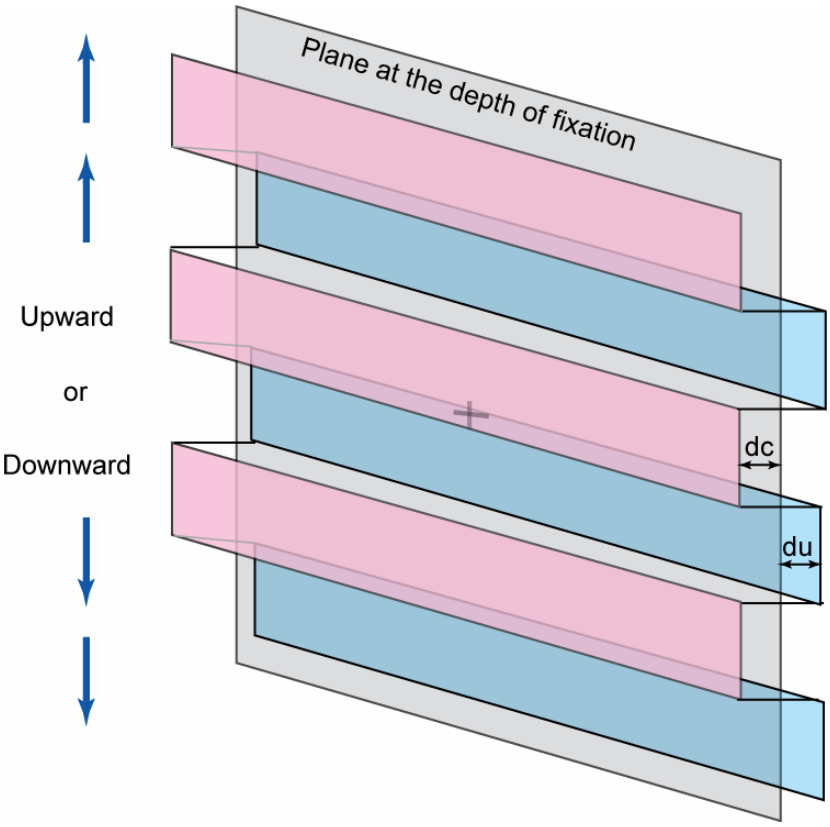
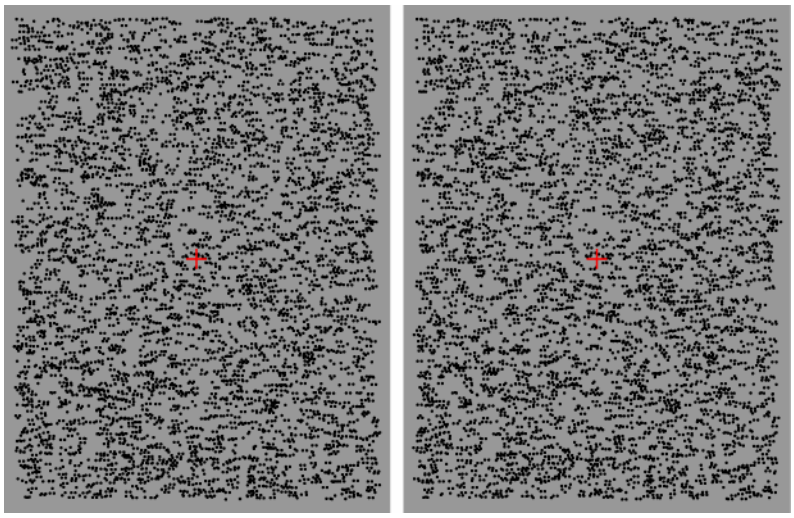


Figure 2.13. A schematic illustration of the depth-defined stripes. Participants reported the direction of moving stripes while increasing the crossed and the uncrossed disparities (d_c and d_u) of the stripes.

a.



b.

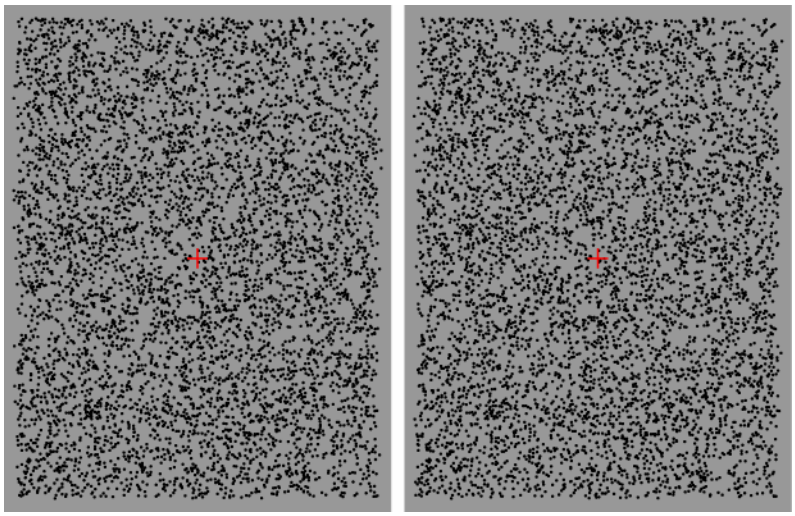


Figure 2.14. Depth-defined gratings in RDS. When binocularly combined, the flat surface at the fixation depth appears sandwiched between the stripes at alternating depths. a. The stripes in alternating depth in PLC RDS are defined with PLC separations in opposite direction in neighboring strips. b. The stripes in the double-layered RDS are defined with disparities in opposite direction in neighboring strips. Optimal viewing distance is about 60 cm or greater.

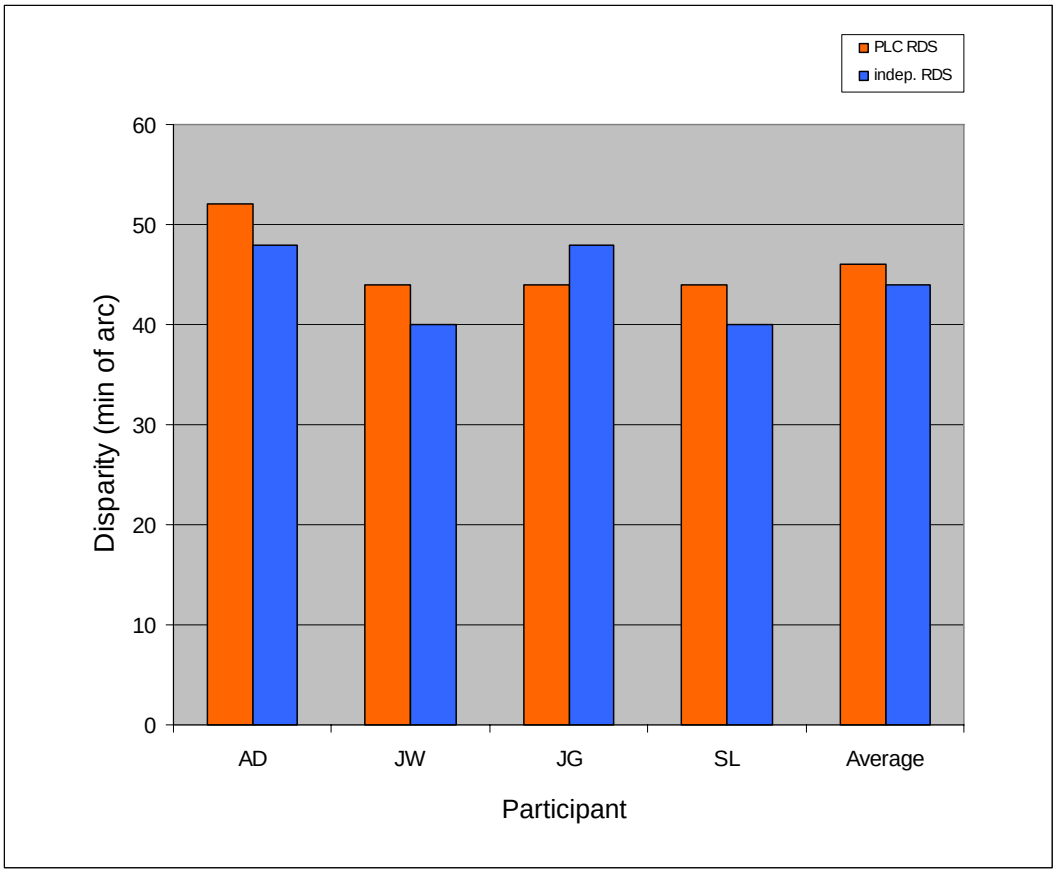


Figure 2.15. The upper depth limit of seeing disparity-defined moving stripes in RDSs. The red bars represent the limit measured with PLC RDS, and the blue bars represent the limit measured with double-layered RDS. The uppers limits in the two configurations are comparable for all participants.

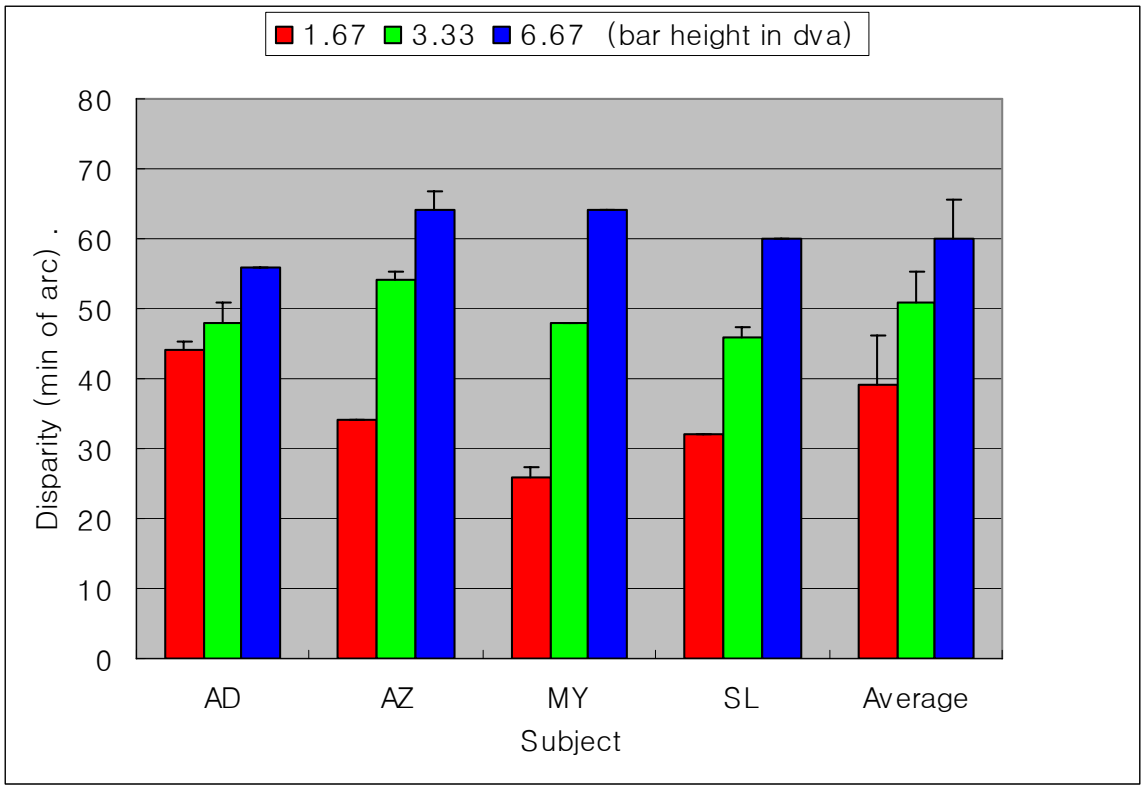


Figure 2.16. The upper depth limit of seeing disparity-defined moving stripes in RDS in relation to the height of the stripes in the double-layered RDS. The red, green, and blue bars represent the detection limits measured with different stripe heights. The upper disparity limit increases as the height of the stripes in the grating increase. Error bars denote the standard deviation.

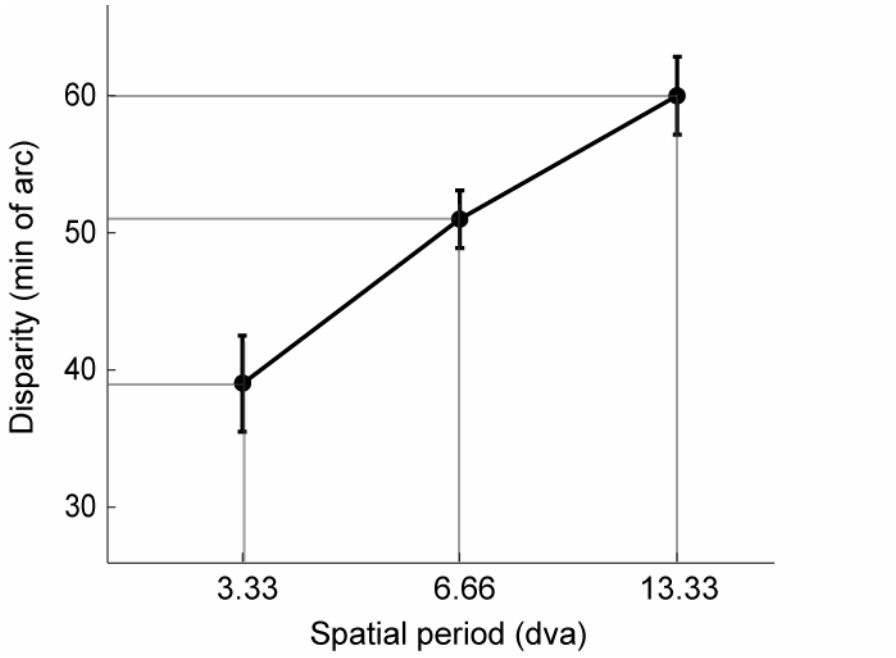


Figure 2.17. Average upper depth limit of seeing disparity-defined moving stripes in RDS. The average in figure 2.17 is redrawn on semi-log scale of spatial period of the stripes. The error bars denote standard deviation.

CHAPTER 3

DOUBLE-MATCHING EXPLAINS THE REVERSAL OF PERCEIVED DEPTH IN PANUM'S LIMITING CASE RANDOM-DOT STEREOGRAM

Images of objects that are not at the depth of fixation fall on different locations on the two retinae. The difference between the two retinal locations is called binocular disparity, which gives rise to a vivid sensation of depth. For a line stimulus, when the binocular disparity is small, generally less than 5 min of arc at the fovea, the two matching images become fused and the objects appears single in depth. Beyond this fusion limit, the line appears double in depth up to about 15 min of arc and beyond that ceases to give rise to an impression of depth (Hampton and Kertesz, 1983; Mitchell, 1966; Ogle, 1952; Palmer, 1961). This is the conventional stereopsis (binocular depth discrimination) that accounts for many ordinary depth impressions.

In Panum's Limiting Case (Panum, 1858 in Ogle, 1962), one views a pair of vertical lines in one eye and a single vertical line in the other eye (figure 1.3). When these images in the two eyes are binocularly combined, the viewer sees two lines at different depths. In Panum's limiting case (PLC) in figure 2.2, there are two possible stereo matches, one match between the left line in the left eye image and the single line in the right eye image, and the other match between the right line in the left eye image and the single line in the right eye image. If there is no additional visual guidance for fixation, viewers show an innate bias to register one of the two lines onto the single line. Researchers have studied whether the depth impression of the unregistered line results

from detecting the binocular disparity between the unregistered line and the single line in the other eye, or requires an alternative explanation that does not concern conventional stereopsis.

3.1 Conflicting Evidence Concerning Double-fusion in PLC

To validate or to invalidate conventional stereopsis in PLC, some researchers hypothesized that the line pair in one eye should appear simultaneously fused with the single line in the other eye when the binocular disparities of the two matches are smaller than the fusional limit. This is referred to as "double-fusion theory," which is a more stringent implication of double-matching in stereopsis (Hering, 1879, in Howard and Rogers, 2002). There are two studies so far that directly support double-fusion in PLC. McKee, Bravo, and Smallman (1995) investigated the binocular masking effect in PLC with a vertical black line flanked by white lines and found that a single visual target in one eye can simultaneously mask two targets presented to the other eye. The other study, by Gillam, Blackburn, and Cook (1995), claimed that a single vertical line can fuse to two oblique lines or two curved lines simultaneously so that they appear slanted or curved in depth.

On the other hand, there are other studies that have disputed double-fusion in PLC. Wang, Wu, Ni, and Wang (2001) reported that the fused appearance was observed in only one of the line pair when orientation disparity was used. Wang et al. (2001) used a leftward-tilted vertical line in both eyes and an additional rightward-tilted vertical line next to this line that was present in only one eye. The fusion was determined by the appearance of the tilt in depth for the fused line. Frisby (2001) supported their finding

with a PLC variant, in which a wavy line in one eye was fused to only one of two straight lines in the other eye, creating an impression of undulation in depth. The disparity range used in these two studies (Wang et al., 2001; Frisby, 2001), however, was much greater than the generally known fusion limit called Panum's fusional area – the sum of the upper fusion limits in the crossed and uncrossed directions.

Other studies that concern the fused state of the PLC stimulus include the study of visual direction of the fused images. When the matching images on the two retinae are fused, the apparent visual direction takes place at the average of the two monocular visual directions. Krol and van de Grind (1980) reported that the positions of the two apparent lines took place at the fused locations, supporting double-fusion in PLC, however, Nakamizo, Shimono, Kondo, and Ono (1994) disputed this by showing that the fused position occurred only for one of the two apparent lines when the disparity was within the Panum's fusional area.

Researchers who have disputed double-fusion in PLC have often inferred an alternative explanation for depth impression in PLC, namely da Vinci stereopsis (Ono, Shimono and Shibuta, 1992; Wang et al., 2001). Nakayama and Shimojo (1990) first described da Vinci stereopsis as a new type of stereopsis based on differential occlusion. In Nakayama and Shimojo's stimulus, a monocular bar presented lateral to a binocular surface gave rise to depth that is quantitatively related to the separation between the bar and the outline surface when they conformed to an occlusion configuration (see figure 2.4). When an opaque surface partially occludes another surface, there is a region on the occluded surface that is visible to only one eye. The width of the occluded region increases as the relative depth between the two surfaces increases. A small object placed

within this monocular occlusion zone is visible to only one eye and the retinal image of the object is always situated on the nasal side of the image of the occluding surface. Although the exact depth cannot be determined because there is a range of depths from which the image of the object can originate, the maximum and minimum depth can be determined. Nakayama and Shimojo (1990) reported that the perceived depth in their da Vinci depth phenomenon approximated the minimum of the possible depth range.

There is evidence that the human visual system utilizes monocular features in perceiving depth. Monocular occlusion facilitates depth perception by reducing the time required to recognize depth discontinuity (Gillam and Borsting, 1988). Monocular occlusion also enhances the impression of depth edges (Shimojo and Nakayama, 1990). Beside those complementary roles, however, there is yet no conclusive evidence that a monocular region is in itself sufficient to give rise to depth sensation in the absence of binocular disparity (Liu, Stevenson and Schor, 1994; Gillam, Blackburn and Cook, 1995).

3.2 Double-fusion May Be Irrelevant to the Depth Phenomenon in PLC

The evidence concerning double-fusion in PLC is diverse and conflicting; therefore inconclusive to validate or invalidate whether the conventional stereopsis is responsible for seeing depth in PLC. However, demonstrating double-fusion in PLC may not be relevant to account for depth impression in PLC for two reasons. First, double-fusion may be inevitably improbable in PLC because the disparity gradient - another governing factor for fusion - exceeds the fusible limit in PLC. The disparity gradient (Burt and Julesz, 1980) is the ratio of the relative disparity and the angular separation for

adjacent features. The magnitude of the disparity gradient in PLC is 2 which is double the maximum disparity gradient supporting fusion (Nakmizo, Kondo and Ono, 1994).

Second, McKee and Verghese (2002) recently showed that elements of the surfaces that are not necessarily fused give rise to a striking appearance of transparent surfaces in depths. In their stimulus, the dot elements were positioned to create a local disparity gradient greater than unity and therefore diplopic, although the binocular disparities by themselves were small enough to be fused otherwise. Nevertheless, the participants correctly matched the perceived depth of the transparent surface to that of disparity-defined opaque surfaces. Furthermore, participants easily detected a test target that appeared sandwiched between the two transparent surfaces in depth.

While many researchers have focused on showing the fusibility/infusibility of the unregistered match in PLC, Gillam, Cook and Blackburn (2003) used an opposite approach. They studied whether the depth phenomenon occurs when the unregistered match is fundamentally unfusible. In their study, Gillam et al. (2003) substituted one of the two lines in one eye with a small disc so that it was not fusible with the single line in the other eye. They found that the disc failed to give rise to quantitative depth related to the different degrees of separation from the line. However, their PLC variant yielded a crude sense of veridical depth when the disc and lines were positioned in stimulus to conform to the occlusion configuration in which the line occluded the disc.

From the study of Gillam et al. (2003), it is clear that the fusibility of the elements disrupts the quantitative depth perception in PLC; however, it is not clear whether the crude depth sensation results from the suboptimal disparity detection between the disparate features or from an alternative mechanism such as da Vinci

stereopsis. Since human observers can acutely label the depth of diplopic images (Blakemore, 1970; Richards and Foley, 1971; Westheimer and Tanzman, 1956) including surfaces (McKee and Verghese, 2002). Furthermore, the random-pattern stereogram is known to create depth impression although the disparate images are not fusible (Kaufman and Pitblado, 1965). Kaufman and Pitblado (1965) created a stereogram with letters, in which the binocularly matching features are the letter O and letter I in the two eyes. The lateral offset assigned to the central region of the stereogram gave rise to depth impression. In this case, the disparity detectors tuned to low spatial frequencies may have detected the disparity of the disparate features as blobs.

3.3 Adoption of Anti-correlation Into PLC

Because fusion is not necessarily required for perceiving depth surfaces even within a disparity range smaller than the fusional area, we suggest that to test double-fusion in PLC, it is necessary to use PLC elements that are not only fusible but also do not support stereopsis. Binocular matches with opposite contrast signs in each eye generally compete producing binocular luster. Such binocular correlation of the image contrast is referred to as anti-correlation.

In random-dot stereograms (RDS), one sees surfaces at different depths as a result of stereopsis. In figure 3.1a, one sees depth-defined stripes when the two half images are properly registered in the two eyes. The depth structure vanishes, however, if the corresponding dots are of opposite contrast (anti-correlated) as one can see in figure 3.1b. The local matches in the anti-correlated RDS appear to be rivalous with no consistent depth (Julesz, 1967; Cumming and Parker, 1997; Masson, Busettini and Miles,

1997; Cumming, Shapiro and Parker, 1998) except at densities lower that 5% (Cogan, Kontsevich, Lomakin, Halpern and Blake, 1995; Cogan, Lomakin and Rossi, 1993; Cumming et al, 1998).

We created RDSs adopting the PLC configuration (PLC RDS), which is similar to the one invented by Braddick (Marr, 1982). In PLC RDS, two dots with a small horizontal separation in one eye geometrically match a single dot in the other eye, giving rise to a pair of transparent surfaces — one in front of the other. The depth difference between the two surfaces is determined by the horizontal separation of the pair of dots. In figure 3.2a, when the two half images are properly registered in the two eyes, one sees PLC-defined stripes that are invisible when either eye is closed. There is one surface at the depth of fixation cross (on-fixation plane) and the stripes altering in depths, which appear in front of and behind the surface at fixation depth (on-fixation planes). The geometrically registered matches create the on-fixation plane and the geometrically unregistered matches create the off-fixation stripes in depths, the alternating depths depend on the direction of the separation from the registered dots.

When the contrast sign of the unregistered dots is inverted, the binocular matches between this dot and the single dot in the other eye become anti-correlated. Figure 3.2b is the PLC RDS that contains anti-correlated matches (anti-correlated PLC RDS). Because anti-correlated matches do not support consistent depth sensation in RDS, if double-fusion occurs in PLC, the stereoscopic stripes in the RDS would disappear in figure 3.2b. On the other hand, if da Vinci stereopsis is responsible for depth sensation in PLC, one would continue to observe the identical stereoscopic stripes in this RDS despite the

contrast inversion. Surprisingly, the depth phenomenon in this stimulus deviates from both of these predictions.

Figure 3.2b shows a RDS that is geometrically identical to the RDS in figure 3.2a except that the unregistered matches are now anti-correlated. In this anti-correlated PLC RDS, one can still distinguish the PLC-defined stripes, however, the direction of the depth for each strip is opposite to that of the RDS in figure 3.2a. The stripes that appeared in front of the on-fixation plane in figure 3.2a now appear behind in figure 3.2b, and the stripes appeared behind in figure 3.2a appears in front of the on-fixation plane.

To examine the reversed depth phenomenon in anti-correlated RDSs, we determined the disparity range of the anti-correlated matches within which the observers saw the reversed depth. We then measured the upper disparity limit up to which one can distinguish the depth structures defined with anti-correlated matches in RDS.

3.4 EXPERIMENT I: Discrimination of the Depth Sign in Anti-correlated PLC RDS

In a pilot study, the reversed depth phenomenon was limited to near the patent stereopsis limit – the limit up to which one can generally distinguish the change of depth magnitudes. We measured the proportion of the trials in which the participants reported the veridical or reversed depth in anti-correlated PLC RDS and control PLC RDS to determine the range within which the depth reversal occurred.

3.4.1 Methods

General settings. Four individuals (including the author) participated in this experiment. The participants had normal or corrected-to-normal vision and had previous experience

viewing RDS. Each participant sat in a dark room at a distance of 65 cm from the display, which was centered at eye level with the aid of a chin-rest.

Displays were generated from custom-made programs by using OpenGL® graphics libraries on a Silicon Graphics Indigo2 workstation and were viewed on a 21" flat-screen color monitor (SGI GDM-5011P, resolution of 1024 x 768, 96 Hz refresh rate; 48 frames/s per eye) through ferroelectric liquid crystal shutter goggles (Crystal Eyes2, Stereographics Corp., 30% open transmittance and open:closed transmittance ratio of 1000:1). These settings were used in all following psychophysics experiments.

RDS stimulus. We created a dynamic RDS composed of two random-dot layers: one layer containing binocular black random-dots whose position in the both eye images were identical (depth reference layer) and the other layer containing monocular white dots whose positions had small lateral shifts respect to the dots in first layer (test layer). These white dots created the anti-correlated matches with the black dots in the other layer. Half the white dots were placed in the left eye image and the other half was placed in the right eye images. Each white dot was visible to one eye only, therefore, the local geometry was of PLC in which a black dot and a white dot in one eye match a single black dot in the other eye. There are two possible binocular matches in this configuration. The match between the white dot in black-white pair in one eye and the black dot in the other eye is an anti-correlated match, whereas the black dot in the black-white pair matching the single black dot in the other eye is a correlated match. We also created control RDS which contained only black dots for the both layers.

The depth planes extended 34 degrees in width and 23 degrees in height and included a red fixation cross (20 min of arc in width and height) at the center of the display. The left and right eye images contained an equal number of random-dots (4 min of arc x 4 min of arc) covering 15% of a median gray background and were replotted every 100 ms. This dynamic RDS produced a greater impression of depth than did a static RDS.

Experimental procedures. In each trial, the participant fixated on a cross in the center of the display and initiated the stimulus with a keypress. In our dynamic RDS, the first correlated binocular layer appeared at the depth of the fixation cross and was displayed for 500 ms. And then the second monocular layer at the test disparity was overlaid on to the first layer and the two layers were displayed for another 2000 ms. After viewing the display, the participant reported whether the second layer appeared in front of (‘near’ response) or behind (‘far’ response) the first layer with a keypress. The degrees of separation that were tested were randomly chosen among 8, 16, 32, 64 and 128 min of arc in both nasal and temporal directions, and each degree of separation was repeated 20 times. The minimum degree of separation was 8 min of arc because at 4 min of arc the dots merged into a single entity.

3.4.2 Results

The first reference layer appeared as a solid surface at the depth of the fixation cross. In the control trials, the second test layer appeared as another solid surface at the small degree of separation. The solidness of this test surface diminished as the degrees of

separation increased, yet the participants were able to distinguish the nearness/farness of the test layer up to about a degree of disparity. Figure 3.3 shows the result the depth discrimination experiment with control correlated PLC RDS. Beyond this range, the depth of the test layer became less clear, often ambiguous. All participants reported ‘near’ for the trials of nasal separation and ‘far’ for the trials of temporal separation in the stimulus unless the amplitude was too great. In correlated PLC RDS, thus, participants had veridical depth experience.

The anti-correlated PLC RDS appeared qualitatively different from correlated control PLC RDS. Unlike the correlated case, the depth layer created with anti-correlated dots failed to cohere into a surface, but rather formed a diffuse cloud in depth away from the on-fixation plane. The depth of the individual dots was not determinable although it was possible to determine whether the dots were in front of or behind the on-fixation plane.

Despite the absence of a definite depth surface, the participants reported the depth of this diffuse cloud reliably at small separations; however, they reported ‘near’ when the direction of separation in PLC configuration would elicit ‘far’ depth in controls, and ‘far’ when the direction of separation in PLC configuration would elicit ‘near’ depth. Figure 3.4 shows the result of the depth discrimination experiment with anti-correlated PLC RDS. At greater separation, the depth became less clear, and the response rate converged to chance. All participants experienced the clear reversal of depth direction near 16 min of arc for both nasal and temporal separation.

3.5. EXPERIMENT II: Discrimination of the Depth Direction in Double-layered RDS

A single anti-correlated layer in a conventional RDS does not support consistent depth impression (figure 3.1b), however, the anti-correlated layer in our PLC RDS whose correlated version produces two-surface perception give rise to a striking impression of depth modulation (figure 3.2b). To determine whether the reversed depth phenomenon is common in RDS with multiple depth surfaces, we measured the proportion of the trials in which the participants reported the veridical or reversed depth in RDSs composed of two layers – one correlated layer and one anti-correlated layer.

3.5.1 Methods

RDS stimulus. The dynamic RDS contained two layers of random-dots whose positions were independent of each other layer (double-layered configuration). One layer had zero disparity with respect to the fixation depth and contained correlated random-dots. The other layer which contained anti-correlated random dots was assigned different test disparities.

The disparities assigned to the test layer was randomly chosen among 4, 8, 16, 32 and 64 min of arc in both crossed and uncrossed directions, and each disparity was repeated 20 times. For crossed disparities the corresponding dots in the two half images were shifted nasal in the stimulus which elicited ‘near’ depth percept in control RDS, whereas for the uncrossed disparities the corresponding dots were shifted temporal in the stimulus which elicits ‘far’ depth perception in control RDS. The dimensions of the stimulus and the procedures in the experiment were identical to the preceding depth sign

experiment of PLC configuration. Control trials contained same contrast black dots in both reference and test layer.

3.5.2 Results

The appearance of the two surfaces in the control trials was almost identical to that of the PLC control trials. This is expected from the close compatibility between disparity-defined surfaces and PLC-defined surfaces, which was the one of the findings in chapter 2. The result of the depth discrimination experiment with correlated double-layered RDS is shown in figure 3.5. All participants reported veridical depth impression for all test disparities although some participants showed a decrease in number of veridical depth judgments around 32 min of arc.

The results from anti-correlated double-layered RDS are shown in figure 3.6. When the dots in the test layer were anti-correlated, the participants' reports showed some degree of reversal in depth perception, especially at small disparities (4 and 8 min of arc). However, it was not as significant as in the anti-correlated PLC RDS experiment (figure 3.4).

3.5.3 Veridical and Reversed Depths in PLC RDS and Double-layered RDS

Figure 3.7 summarizes the results of experiment I and II showing the averages of the participants' report in the PLC RDS and double-layered RDS. The minimum PLC separation was 8 min of arc whereas the minimum disparity was 4 min of arc in double-layered RDS because the 4 x 4 min of arc dots became merged at 2 min of arc into a single entity. The stimulus disparity was up to only 32 min of arc in double layered RDS

because participants' judgment in anti-correlated configuration converged rapidly to the chance probability.

Participants reported veridical depth above 90% up to a degree for both PLC and double-layered RDS in control correlated experiments. The response patterns match closely in the two cases. The similarity in the two cases for correlated RDS is consistent with the findings in chapter 2 that the PLC-defined depths and the disparity-defined depths are almost identical.

In anti-correlated cases, participants often reported reversed depth at some ranges which differ for PLC and double-layered RDS. The depth reversal was much more significant and present up to greater disparities in PLC than double-layered configuration. For example, on average the depth reversal peaked near 100% at 16 min of arc in the anti-correlated PLC configuration, but the reversed depth response dropped significantly at identical disparity in double-layered configuration (almost chance rate in crossed disparity and below 30% in uncrossed disparity). The different degree of depth reversal in the two configurations is again discussed in chapter 4.

3.6. EXPERIMENT III: The Upper Depth Limits of Anti-correlated RDS

In control depth sign experiments using correlated disparity-defined RDS and PLC-defined RDS, the number of veridical responses decreased at similar disparity/separation near 64 min of arc in both RDS types (figure 3.3, figure 3.5 and figure 3.7). It is not surprising because the depth surfaces created with the two configurations appeared to be very similar when compared side by side (figure 2.9) and the depths between the two configurations can be very closely matched (figure 2.12).

Furthermore, the upper depth limits for seeing motion with PLC-defined and disparity-defined stripes are almost identical. For the anti-correlated RDS, however, the reversed depth phenomenon manifests at different disparity/separation range (figure 3.4, figure 3.6 and figure 3.7).

We first created disparity-defined gratings and PLC-defined gratings in RDS, which appeared as horizontal stripes at two alternating depths when binocularly combined, figure 3.2a for example (see figure 2.13 for perceptual appearance of the depth stripes). These gratings elicit perception of vertical motion when the assigned depths are shifted either upward or downward. We then adopted anti-correlated layer into these two types of RDSs. We asked the subjects to determine the direction of motion while increasing the disparity/separation of the depth stripes. Perception of motion in these RDSs requires stereopsis since there was no monocular cue for motion; therefore, the disparity threshold for motion detection reflects the upper depth limit at which one can discriminate the depth -defined regions in RDS.

3.6.1 Methods

The local geometry and the dimensions of the PLC-defined stripes and disparity-defined stripes are identical to the preceding experiments described in the methods section in 2.5.1. For the anti-correlated PLC RDS, the dots in the second layer that created apparent depth stripes were altered from black to white. For anti-correlated double-layered RDS, the random-dots in the apparent depth stripes had corresponding dots of opposite contrast in the contralateral eye. The experimental procedures were also identical to the preceding depth limit experiments described in section 2.5.1 except that

the initial disparity/degree of separation of the stripes were 8 min of arc and the incremental was 2 min of arc with a keypress.

3.6.2 Results

At small disparities, the stripes were clearly discernable in anti-correlated PLC RDS; therefore the direction of motion was easily judged. The appearance of the stripes, however, was different from that of the same contrast depth gratings. At the smallest separation, the stripes appeared to be merged with the on-fixation layer and formed a single surface with smooth in-depth modulation. Beyond this separation, the dots in the stripes still appeared to be in alternating depths either closer or further from the on-fixation plane, but the distance from the on-fixation plane was not determinable. This amorphous appearance persisted up to a half a degree (28 min of arc) in PLC RDS, but at the greater degrees of separation the boundaries of the stripes became indistinguishable, disabling the judgment of the motion direction. The results of the upper depth limit of detecting the motion of anti-correlated PLC-defined stripes are shown in figure 3.8 along with that of anti-correlated disparity-defined stripes.

In double-layered RDS, the anti-correlated layer and the correlated on-fixation layer appeared as a single merged surface modulating in depth, an observation similar to that of anti-correlated PLC RDS. However, the diffuse cloudy appearance of the anti-correlated matches was not observed in double-layered RDS at any disparity. The average upper depth limit in the anti-correlated double-layered RDS was 16 min of arc which was about a half (57%) of the average upper depth limit of the anti-correlated PLC RDS. Note that these two upper depth limits are related to the height of the depth-defined stripes (see

section 2.6) so that these limits only provide the relative measures between the anti-correlated matches in PLC-defined and disparity-defined depth structures.

3.7 Seeing Veridical Depth in Anti-correlated RDS

Sparse anti-correlated RDS can provide weak impression of depth at density lower than 5% (Cogan, Kontsevich, Lomakin, Halpern and Blake, 1995; Cogan, Lomakin and Rossi, 1993). But when stereopsis occurs, it is at a depth consistent with the binocular disparity of the anti-correlated matches and not reversed in depth (Cumming, Shapiro and Parker, 1998). In contrast, our 15% anti-correlated RDS is considered dense and the depth reversal occurred up to 50% density when the contrast of random-dots were half black and half white against gray background in a static RDS. The hardware that we used to display our dynamic RDS had limitations concerned with the maximum number of dots that can plot within the 100 ms interval. We chose the 15% density to eliminate any possible delay due to the hardware.

The results in the depth sign experiment summarized in figure 3.7 shows that the perceived depth sign from our anti-correlated RDS is inverted at small disparities and the reversal peaked at near 16 min of arc. Beyond these peaks, both in nasal and temporal directions, the participants' response appeared to converge to the chance rate at greater test disparities. Although there is no further data point beyond two degrees, we do not think the depth sign would cross over from the reversed to the veridical beyond two degrees. Because if so, the participants would have detected the motion of anti-correlated PLC-defined stripes up to greater depth limit than half a degree in the motion detection experiment (see figure 3.7).

3.8 Occlusion Relationship in PLC RDS

In our anti-correlated PLC RDS, the layer that contains the unregistered opposite contrast dot in PLC pair appears in the depth direction that is opposite to the assigned geometry. If the visual system uses the configural relationship between the registered and unregistered images, as suggested in occlusion theory or in da Vinci stereopsis, in our PLC RDS, the perceived depths should be conserved qualitatively and quantitatively. Therefore, an occlusion relationship cannot be responsible for the depth impression in anti-correlated PLC RDS.

Some researchers claimed (Ono, Shimono and Shibuta, 1992) that the PLC is a type of da Vinci stereopsis stimulus in which the width of the occluding surface equals the width of the monocular feature on the temporal side in the stimulus. Gillam et al. (2003), on the other hand, made a claim that is reversed. In Gillam et al.'s view, da Vinci stereopsis is a type of PLC in which the boundary of the binocular surface is matched to the monocular line and the quantitative depth perception is possible through double-matching.

3.9 Reversed Depth Phenomenon in Anti-correlated PLC

At first thought the conventional stereopsis also has difficulties to reconcile the depth phenomenon in our stimulus, because anti-correlated matches in dense RDS do not support consistent depth perception. There is, however, other evidence that we think relevant to the depth reversal in our anti-correlated RDS.

One class of evidence is related to the neuronal response to the anti-correlated stimulus. In primates, the neural signals from the two retinae converge on binocular cells

in the primary visual cortex (V1), the earliest cortical area where disparity encoding may occur. These V1 cells modulate their firing rate in response to binocular disparity (Poggio and Fischer, 1977; Barlow, Blakemore and Pettigrew, 1967 in Area 17 in cats). A subset of these cells – binocular complex cells – is also sensitive to the depth changes of disparity-defined structures within RDS (Poggio, Motter, Squatrito and Trotter, 1985; Poggio, 1995). In response to anti-correlated RDS, the disparity-tuning curves of these cells are often inverted (Cumming and Parker, 1997). For example, response of the cells that are tuned for the disparities in front of the plane of the fixation ("near neurons") appears as if these cells are tuned for the disparities behind the plane of the fixation ("far neurons"). More recently, farther in the dorsal stream of visual pathway, other areas including MT (Krug, Cumming and Parker, 2004) and MST (Takemura, Inoue, Kawano, Quaia and Miles, 2001) are reported to contain cells that shows disparity tuning curve inversion.

Other evidence is related the visuomotor response to the anti-correlated RDS. To align both eyes on the same object and to facilitate the binocular fusion of visual images, our eyes converge or diverge as we look at a nearby or distant object. Disparity-selective cells in cortical and subcortical areas are often suggested to support disparity-induced vergence eye movement by sensing the misalignment of the two eyes with respect to the object of interest (Poggio and Fisher, 1977). When anti-correlated RDS was presented briefly at some disparity, vergence eye movement with ultra-short latencies was evoked in the direction opposite to that evoked by same contrast RDS in both human and monkeys (Masson et al. 1997).

Based on this physiological evidence related to the anti-correlated stimulus, we suppose the reversed depth phenomenon in our anti-correlated RDS is an exceptional case in which the underlying neuronal signal is effectively contributing to our conscious depth perception. We pursue further explanation in the following chapter by extending the well-established disparity detection model to simulate the disparity maps of the correlated and anti-correlated stimulus.

In this chapter, we introduced a new type of anti-correlated RDS that gives rise to a depth impression which is opposite to the assigned disparity but consistent with the underlying neuronal signals. Although Read and Eagle (2000) also showed weak depth reversal in anti-correlated band-pass gratings, to our knowledge, our stimulus is the first demonstration of the reversed depth effect related to the anti-correlated RDS.

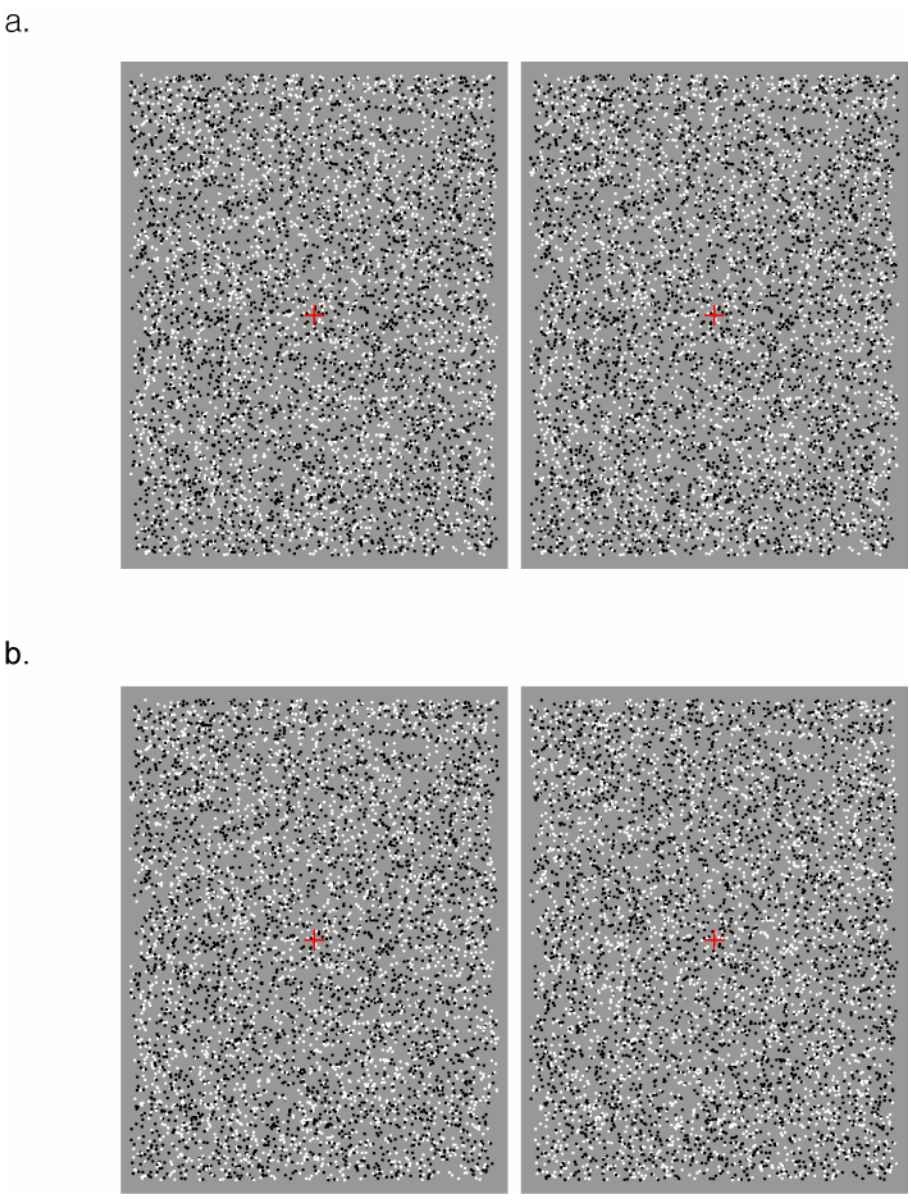


Figure 3.1. Disparity-defined gratings in correlated and anti-correlated RDSs. a. A correlated RDS. The binocularly matching dots are of same contrast. One sees disparity-defined gratings in alternating depths when the stereogram is binocularly combined. b. An anti-correlated RDS. The binocularly matching dots are of opposite contrast. One sees no depth structure although the geometry of the random-dots is identical to that of the RDS in a.

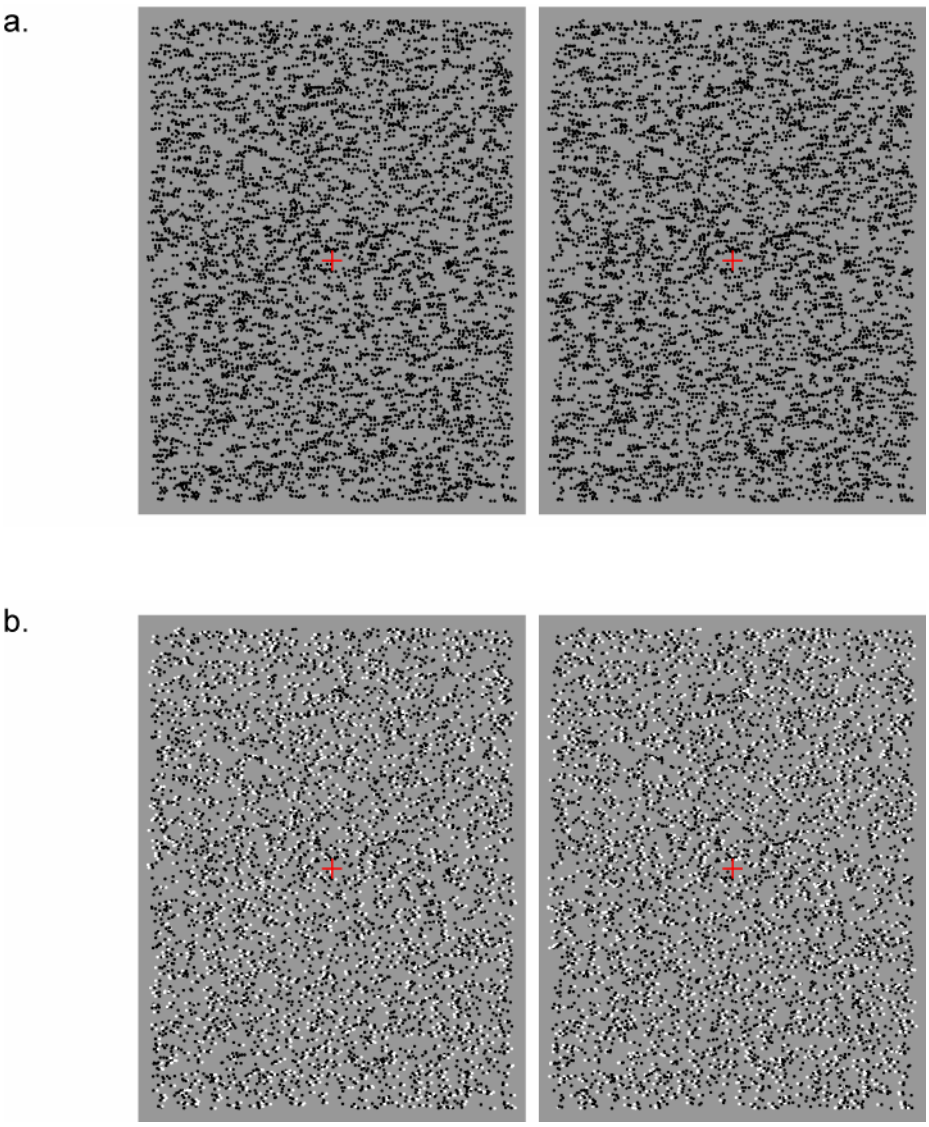


Figure 3.2. Disparity-defined gratings in correlated and anti-correlated PLC RDSs. a. A correlated PLC RDS. One sees disparity-defined grating in alternating depths sandwiching another flat surface at fixation depth. b. A PLC RDS that contains anti-correlated layer. The depth-defined grating is still observed when binocularly combined, however, the depths of the stripes are at opposite to the depth perceived with the RDS in a although the geometry of the random-dots is identical.

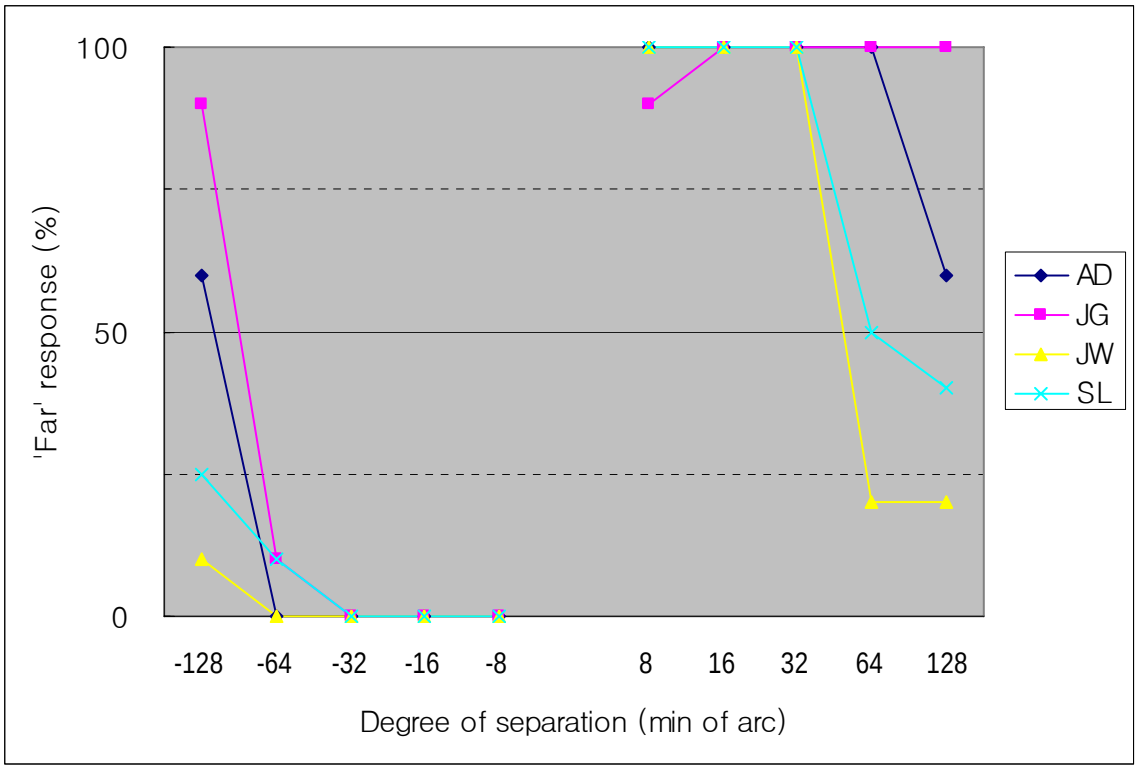


Figure 3.3. Control depth direction experiment. Proportion of trials in which participants judged the test layer to be behind the reference fixation plane ('far' response). The test layer was composed of dots with same contrast as the random-dots in reference layer. The participants' responses concur with the veridical depth that is expected with the PLC configuration unless the degree of separation was too great.

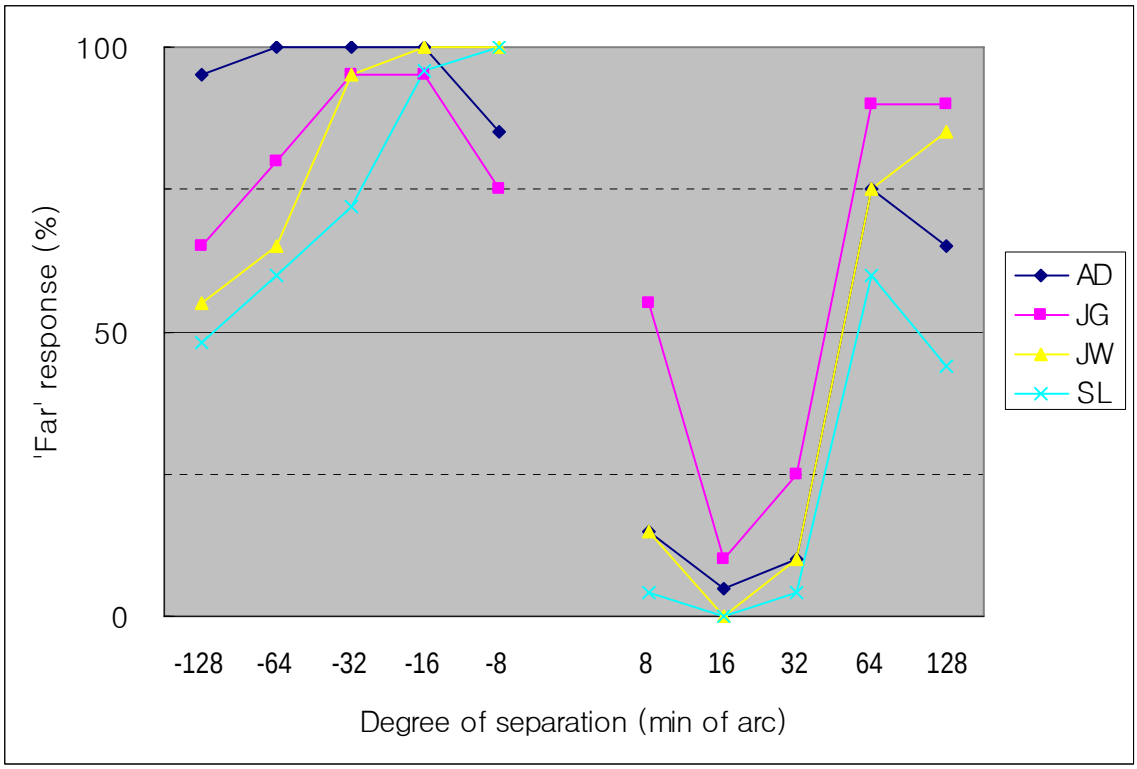


Figure 3.4. Anti-correlated PLC depth direction experiment. Proportion of trials in which participants judged the test layer to be behind the reference fixation plane ('far' response). The test layer was composed of dots with opposite contrast to the random-dots in reference layer. The participants' reports switched to the opposite direction at small disparities.

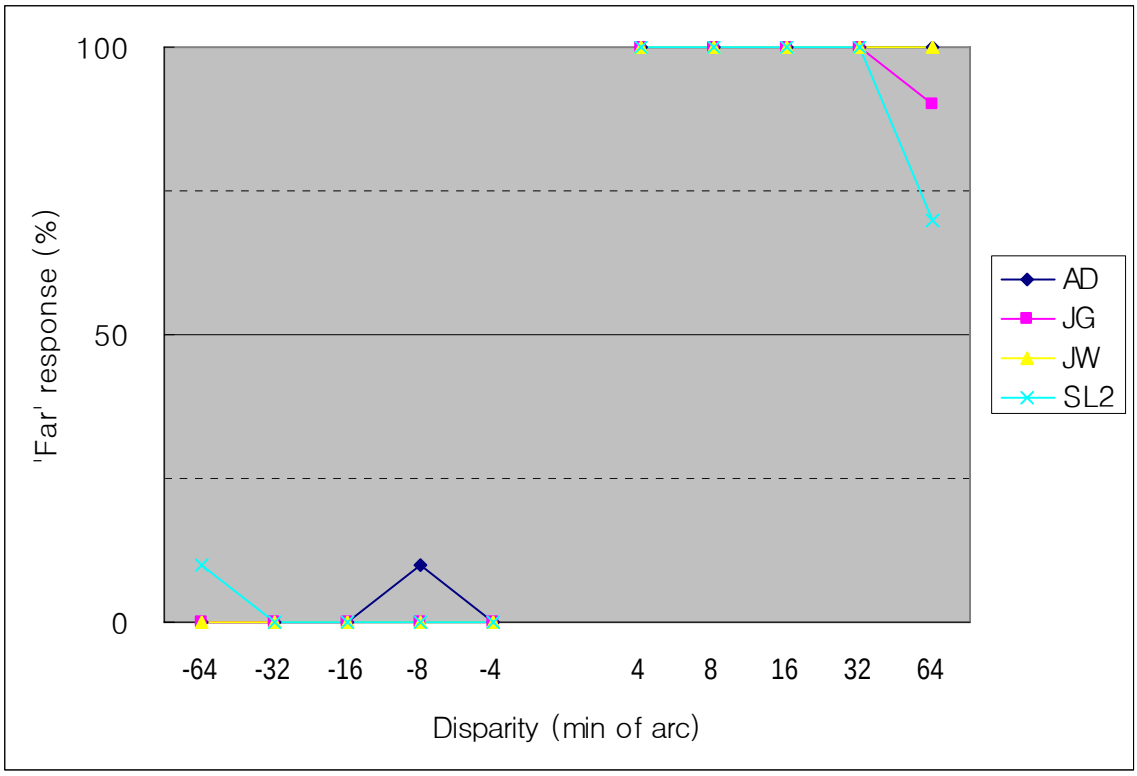


Figure 3.5. Control depth direction experiment using independent-layer RDS. Proportion of trials in which participants judged the test layer to be behind the fixation plane ('far' response). The test layer was composed of dots with same contrast as the random-dots in reference layer. The participants' responses concur with the veridical depth of the testing disparities.

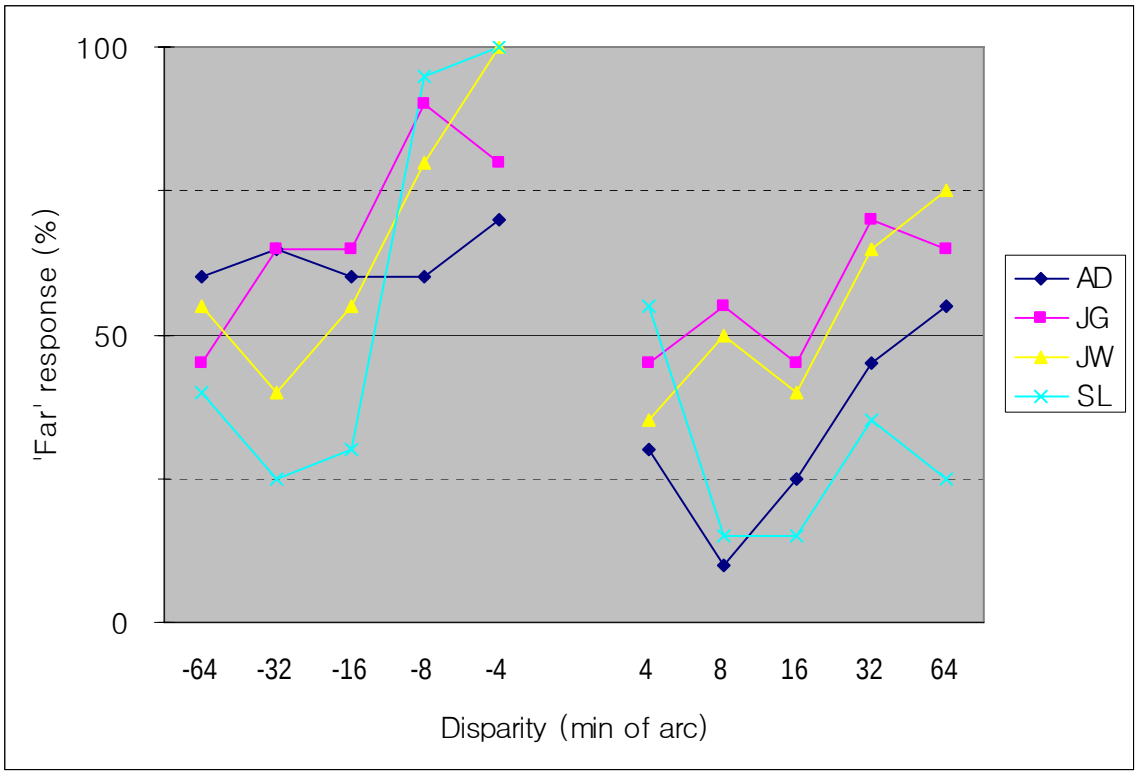


Figure 3.6. Depth direction experiment using opposite contrast independent layer RDS. Proportion of trials in which participants judged the test layer to be behind the fixation plane ('far' response). The test layer was composed of dots whose contrast was opposite to the corresponding dots in the other eye image. There was some degree of reversed depth effect but not as significant as in the anti-correlated PLC RDS.

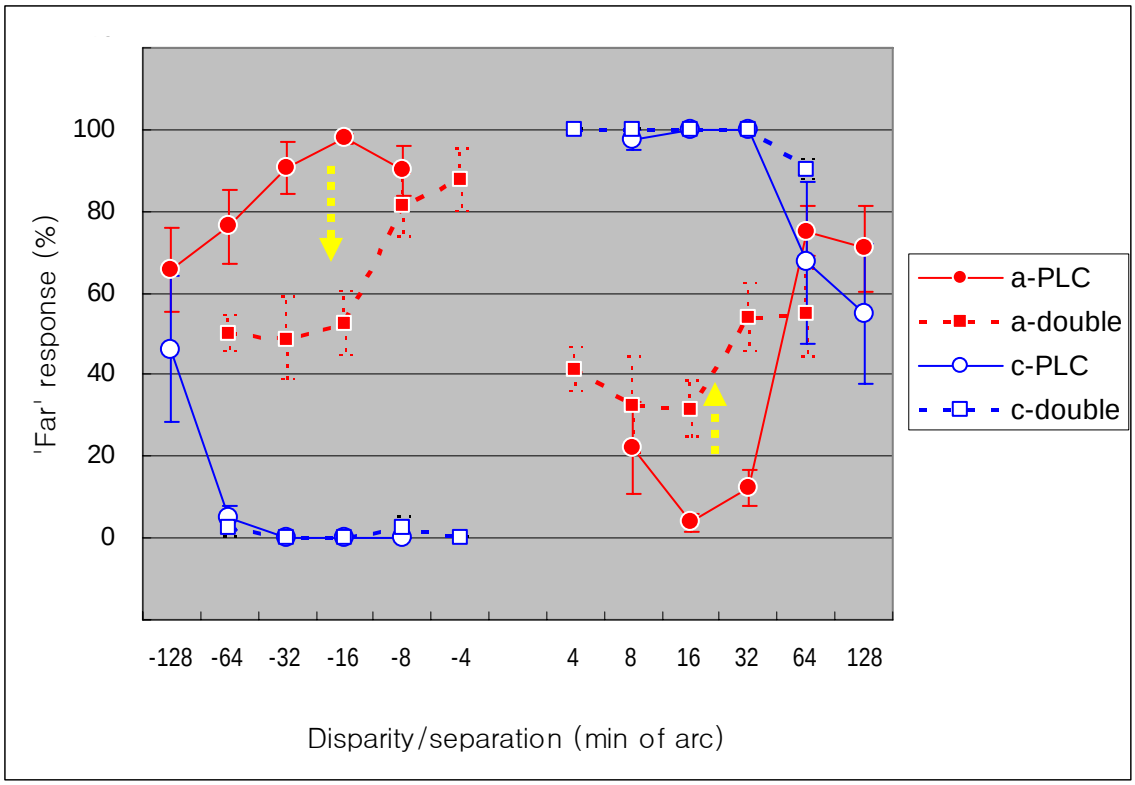


Figure 3.7. Averages of four participants from the depth direction experiment with PLC RDS and double-layered RDS. The effect of depth reversal is much reduced in anti-correlated double-layered RDS than anti-correlated PLC RDS (yellow dotted arrows). Filled markers represent the anti-correlated configuration and the open markers represent the correlated configuration in the RDSs. Solid lines represent the PLC configuration and dotted lines represent the double-layered configuration in RDSs. (a-PLC: anti-correlated PLC RDS, c-PLC: control correlated PLC RDS, a-double: anti-correlated double layered RDS, c-double: control correlated double-layered RDS). Error bars denote standard deviation.

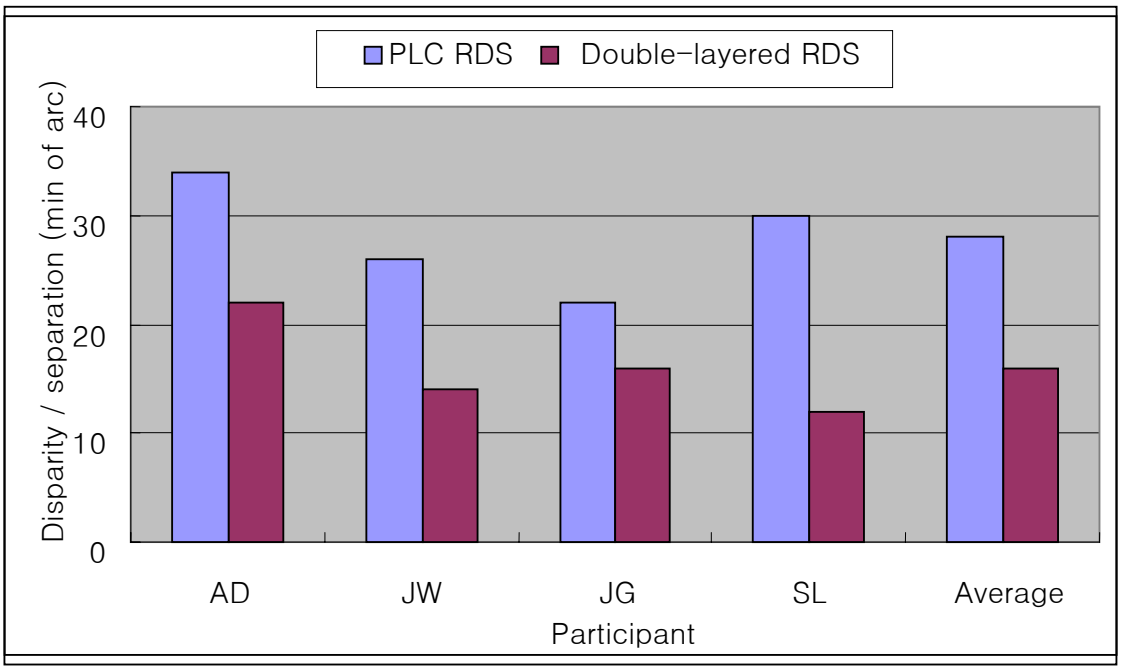


Figure 3.8. The upper depth limit of seeing PLC-defined and disparity-defined moving stripes in opposite contrast RDS. The blue bars represent the limit measured with PLC RDS, and the purple bars represent the limit measured with double-layered RDS.

CHAPTER 4

SIMULATION OF STEREO MODEL THAT RESOLVES TRANSPARENCY AND PREDICTS DEPTH REVERSAL

Many experiments have demonstrated the neural substrates for disparity coding at the earliest stage of binocular convergence (Barlow, Blakemore and Pettigrew, 1967; Poggio and Poggio, 1984). However, the question of how the population of disparity selective cells could be used to compute disparity maps and further represent / assign depth is left open. Ohzawa, DeAngelis and Freeman (1990, 1997) proposed a model based on their quantitative physiological studies, the binocular energy model whose response is simulated by summing up the squared responses of quadrature pair of binocular simple cells (figure 4.1). Although there is no direct anatomical evidence supporting the quadrature pair method for constructing binocular complex cells from binocular simple cells yet, the method is valid at least as a phenomenological description for the response of a subset of real complex cells. The model captures many aspects of the behavior of disparity-tuned V1 cells (Barlow, Blakemore and Pettigrew, 1967; Poggio and Fischer, 1977).

Implementing the binocular energy model of Ohzawa-DeAnglelis-Freeman (ODF binocular model), Qian (1994) demonstrated that the disparity map could be computed from RDS. Furthermore, Tsai and Victor's model (Tsai and Victor, 2003) which adopted an energy model in the front end and uses disparity templates to represent the

disparity map has also been shown to resolve transparency in RDS (McKee and Verghese, 2002).

In the case of anti-correlated images in which the features in one eye are geometrically matched to the features of reversed contrast in the other eye, the binocular energy model predicts the inversion of the disparity-tuning curve (Ohzawa et al., 1990) as shown in figure 4.2 (a, b, c and d). In addition to Ohzawa et al.'s (1990) original report which used anti-correlated bar in cats, Cumming and Parker (1997) reported that the V1 disparity selective cells often inverted their disparity tuning when tested with anti-correlated RDS. Other studies showed inversion of the disparity tuning curves farther in the dorsal stream of visual pathway including area MT (Krug, Cumming and Parker, 2004) and MST (Takemura, Inoue, Kawano, Quaia and Miles, 2001).

In spite of the physiological evidence that the neuronal signals related to the anti-correlated stimulus persist up to some range of visual processing, the anti-correlated RDS does not produce consistent depth sensation (Cumming, Shapiro and Parker 1998). The interpretation of the inverted signal and the mechanism that hinders these binocular signals from affecting our conscious perception is not addressed clearly yet. One suggestion is that with anti-correlated features, the peaks of the disparity-tuning curves would shift different degrees depending on the cell's periodic nature of the spatial sensitivity in the receptive region on the retinae (Read and Eagle 2000).

In chapter 3, we introduced a new type of RDS that contains an anti-correlated layer which gives rise to the impression of depth that is opposite to the depth expected with the assigned disparities. We simulated the response of ODF binocular model units across scale with each of correlated, anti-correlated and PLC bar stimuli to explore the

explanation of why our anti-correlated PLC RDS yields depth perception while the standard anti-correlated RDS does not. Then, based on the observation from the model simulation, we tested whether limiting the range of scales within a population of disparity detectors affected the depth discrimination in RDS composed of bandpass elements. Last, we adopt a simple rule that integrates the response of ODF binocular model units at different scales and show that the combined responses across scale discriminate transparent surfaces and predict the reversed depth phenomenon in anti-correlated PLC RDS.

4.1 Simulation of ODF Binocular model Across Scale

Ohzawa et al. (1990) proposed a disparity detection mechanism that behaves like a binocular complex cell in the visual cortex, which is composed of 4 binocular simple-cell-like subunits (ODF binocular model). The receptive retinal regions of early cortical cells (receptive field, RF) called simple cells are composed of elongated subregions that respond to either onset or offset of a flashed bar of light, whose profiles are outlined in figure 4.3a. The sensitivity profile to the luminance contrast is modeled with Gabor function - Gaussian weighted (envelope) sine function (carrier) – which is shown in figure 4.3b.

One simple cell that receives input from left eye visual field and the other simple cell that receives input from the geometrically corresponding right eye visual field are assumed to be paired to compose a simple binocular subunit. These two simple cells in the two eyes are also assumed to have similar spatial frequency in their RFs. Depending on the phase difference between the carrier functions (sine) in the simple cell pair, the

integration of the two cell responses behaves as a simple disparity detector (figure 4.4). When the RFs in the two eyes have identical phase shift, the binocular simple cell is tuned for zero disparity (figure 4.4a). When the phase of the one eye RF is shifted in relation to the other eye, the given binocular simple cell responds best to the disparity equal to that shifted phase (figure 4.4b). This type of disparity detection is referred to as phase-disparity mechanism. Phase-disparity detection units are tuned to a disparity approximated to the interocular RF phase shift (detector disparity).

The disparity detection of a single binocular simple cell, however, is sensitive to the contrast polarity and also dependent on the stimulus position within the RFs (figure 4.5), which is not a suitable characteristic for true binocular detectors. In figure 4.5, the given binocular simple cell model is tuned for zero disparity (figure 4.5a). This model unit responds well when the binocular dark bars are positioned in the off subregions (figure 4.5b). However, when the contrast polarity of the bars at the same position are inverted to white (figure 4.5c), or when the positions of the black bar stimuli shift simultaneously within the left and right RFs (figure 4.5d), the model unit changes its response dramatically.

A full description of the disparity-tuning of the binocular simple model unit is plotted in figure 4.6. The activations of the left and right monocular simple cells are added and plotted in color scale at the corresponding coordinate in the binocular RF map. For example, the peak response occurs in the center of the RF map because the activations in the left and right monocular simple subunits have the maximum activation at the given coordinates (figure 4.6a). The actual response of the given binocular simple

model unit is shown in figure 4.6b, which was stimulated with a binocular white bar stimulus.

On the other hand, complex cells are insensitive to contrast polarity and broadly selective to the stimulus position within their RFs. These two characteristics can be achieved when the responses of multiple simple cells are combined, whose periodic component of the RFs vary in phase. In figure 4.7, four Gabor functions (the model function of the simple cell RF) have multiples of 90 degree phase shifts in relation to each other. The squared sum of half-rectified Gabor functions is uni-polar and much broader than the individual functions.

In the ODF binocular energy model, the RF positions and RF periods of the four binocular simple subunits are identical but the RF phases differ in multiples of 90 degrees (figure 4.1). The responses of four subunits (figure 4.8a) are halfwave-rectified (figure 4.8b) and then squared (figure 4.8c). The summed response of these four subunits (figure 4.9) shows the disparity-tuning map similar to that of the complex cells in cats (Ohzawa et al., 1990) and monkeys (Cumming and Parker, 1997). We simulated the response of ODF binocular energy model at different scale with correlated, anti-correlated and PLC bar stimuli to compare the distribution of the peak response across scale.

4.1.1 Methods

We used the following 1D Gabor filters to describe the RF profile of a binocular simple cell:

$$f_L(x) = 1/(\sqrt{2\pi}\,\sigma)\exp(-x^2/2\sigma^2)\cos(\omega_0 x + \phi_L),$$

$$f_R(x) = 1/(\sqrt{2\pi}\,\sigma)\exp(-x^2/2\sigma^2)\cos(\omega_0 x + \phi_R), \qquad (1)$$

where ω_0 is the peak preferred frequency, σ is the horizontal Gaussian width that vary with $\omega_{0,}$ ϕ_L and ϕ_R are the left and right phase parameters respectively. Simple binocular cells compute the sum of the left and right images filtered by the respective RF (Anzai, Ohzawa, and Freeman, 1999; Qian, 1994):

$$r_s = L + R = f_L \bullet I_L + f_R \bullet I_R, \quad (2)$$

where L and R are the left and right eye contributions to the linear response, I_L and I_R are the right and left images, and the symbol '$\bullet$' denotes the inner product operator. The response of the ODF model is the sum of squared response of two pairs of simple cells in quadrature phase:

$$r_q = (L_1 + R_1)^2 + (L_2 + R_2)^2, \quad (3)$$

where the phase parameters of the underlying simple cells differ by 90^o:

$$\phi_{2,L} = \phi_{1,L} + \pi/2 \text{ and } \phi_{2,R} = \phi_{1,R} + \pi/2 \, . \quad (4)$$

The period of the carrier function in RFs (cosine) ranged from a minimum of 9 pixels scaled up repeatedly by a factor of 1.25 up to 105 pixels. The detection range was limited to one period of the carrier function in the Gabor filter because the disparities beyond multiples of a cycle become indistinguishable form one another.

Figure 4.10 illustrates one of the scaled RFs used in the simulation. The RF shifts between the two eyes in ODF binocular energy model units covered disparities from $-\pi$ to π at different scale as shown in figure 4.10a. Disparity-detectors that incorporate the phase-disparity mechanism do not distinguish the phase shift beyond its RF period. The receptive fields of these units are centered at the same location. One of these units responds maximally when the stimulus within the RF matches the phase disparity of the detector. In figure 4.10b, the stimulus disparity is d which matches a quarter wavelength

shift in one of the detectors (captured in the red rectangle) in figure 4.10a. This unit shows the peak response (denoted with red bar) within this spatial frequency of RF. The overall model units that included all test RF frequencies can be imagined as a composite of scaled detectors that are stacked onto each other at the same location as a “wedding cake”.

4.1.2 Results

Figure 4.11a shows the response of the scaled ODF binocular energy model units to the correlated bar stimulus. We used twelve spatial frequency channels in the model simulation. The bar stimulus (15 pixels) had crossed disparity of 5 pixels. The disparity estimates (peaks denoted with x in the figure) of the detectors at each scale were clustered near the tested disparity (denoted with red line). Beyond the peak activation at the RF period of 27 pixels, the estimates were very consistent.

Figure 4.11b shows the activation pattern of the model units to the anti-correlated bar stimulus. The peak activation at each scale occurred on the upper half of the plot which is opposite to the expected depth. Note that there is another peak within a scale on the other side of the activation map. The degrees of phase difference (disparity estimates) at different scales were dispersed, so that the coarser the scale (the lower the spatial frequency) the more erroneous the disparity estimate became. This is consistent with the suggestion by Read and Eagle (2000) that the peaks would shift in relation to the detector’s RF spatial frequency.

The peaks at different scales are clustered in the correlated bar stimulus, whereas the peaks were spreading out across scale in the anti-correlated bar stimulus. The model

response to the anti-correlated PLC stimulus was intermediate between those of correlated and anti-correlated bar stimuli. In figure 4.11c, the peaks are on the opposite side of the expected depth but are not as dispersed as the responses to the anti-correlated bar stimulus.

4.1.3 Depth Impression in Anti-correlated PLC RDS

The anti-correlation in the bar and PLC stimuli resulted in inversion of the response pattern so that each disparity estimate (the peak at each scale) occurs on the opposite side of the disparity estimate of the correlated bar stimulus. Another important observation is how the pattern of the disparity estimates is distributed across scale. With the stimuli that give rise to consistent depth impression – the correlated bar and the anti-correlated PLC, the peaks across scale are clustered near the maximum response and the difference of disparity estimates among the more activated peaks are small (compare the peaks at RF period of 22, 27 and 34 pixels).

In contrast, in the anti-correlated bar stimulus, which does not give rise to consistent depth impression, the difference among the more activated peaks are great and even greater at larger scales. In figure 4.11, the gray line that estimates the linear change in the more activated peaks has much steeper slope for the anti-correlated bar stimulus than for the other two. The magnitude of the slopes in the correlated bar and anti-correlated PLC stimuli appears compatible.

The limited dispersion of disparity estimates in anti-correlated PLC is due to the black and white contrast alternation in the one eye - effectively a bandpass stimulus - which reduced the activated range of RF frequency channels. We think such effect creates

the disparity estimate pattern that is similar to that of the correlated stimulus giving rise to the depth impression. However, the disparity estimates still contain some degree of dispersed estimates at greater scales, and evidently fails to support the surface integration in our anti-correlated PLC RDS.

4.2 Depth Discrimination in Low-passed and Band-passed Stimuli

The model simulation of binocular correlation in the preceding section showed that the disparity estimates are clustered in response to stimuli that give rise to consistent depth sensation, whereas the disparity estimates are dispersed across scale for stimuli that do not. We hypothesized that the coherency of the disparity estimates across scale is important for depth discrimination and tested if limiting the activation of detectors to particular spatial frequency channels would affect the performance of the depth discrimination task.

It is possible to decompose an image as a summation of a series of sine and cosine terms of increasing spatial frequency. Because complex cells and the ODF model are tuned to a specific spatial frequency, by manipulating the frequency content of the stimulus, we can limit the population of the scaled disparity detectors. We created two variants of random-dot stereogram (RDS): one containing vertical bars and the other containing small gratings composed of black and white bars (shown in figure 4.13). Figure 4.12 shows the power spectrum of these two element types. The passband in the grating is focused around the fundamental frequency of the grating (reciprocal of the period of black and white contrast alternation) whereas the passband of the bar is much broader. Our prediction is that the depth discrimination would be more accurate in RDS

composed of bar elements than the grating elements because the bar elements would have a broader range of detectors activated, whose activation are coherent.

We tested this hypothesis in a psychophysical experiment in which the participants determined the depth order of the two adjacent surfaces. If an increasing portion of the random-elements of the surfaces is substituted with uncorrelated elements, one has increasing difficulty to discriminate the depths of the two neighboring surfaces. We measured the percent correct of the participant's depth judgments in the two RDS types.

4.2.1 Methods

RDS stimulus design. The RDS stimulus contained two copies of identical random-element pattern in the upper and lower halves of the display. The two copies were separated by 1 dva and had a fixation cross with nonius lines in between them. Each copy of the pattern extended 25.6 degrees in width and 7.5 degrees in height. One of the two RDS surfaces had a small crossed or uncrossed disparity of 1.5 min of arc so that one of the copies appeared closer than the other. The random-elements of these RDS were either black and white bars or a small grating composed of black-white-black-white bars. The width of the bars was 3 min of arc and the height of the bars was 12 min of arc. When the grating elements in the RDS were blurred or viewed at a far distance, they appeared gray and were indistinguishable from the gray background. Figure 4.13 shows the two RDS variants used in these experiments. In both type of RDS variants, the elements filled 5% of the median gray background. We also tested 20% RDS for grating elements, because

the number of grating elements were a quarter of the bar elements when the density was equal.

Experimental procedures. Four individuals (including the authors) participated in the experiments and two were naïve observers. Each participant sat in a dark room at a distance of 98 cm from the display, which was centered at eye level with the aid of a chin-rest.

The experiments had a two stage procedure. The initial stage was used to estimate the proportion of uncorrelated bar elements at which the participants could discriminate the depth order of the two neighboring surfaces correctly 75% of the time. In the second stage, we used this estimated threshold to measure the percentage of correct responses in the following procedures using two element types. The task of the participants and the procedures in each trial were identical in the two stages. Only the method of varying the ratio of the correlated and uncorrelated elements was different; namely, the staircase method was used in the first stage and the method of constant stimuli in the second stage.

In each trial, the participant fixated on the cross in the center of the display and then initiated the stimulus presentation with a keypress. After 500 ms delay, the upper and lower halves of the RDS appeared for 100 ms. The fixation cross disappeared during the RDS presentation and then reappeared afterward. One of the two RDS was randomly assigned a small depth offset (1.5 min of arc). The participant reported which half appeared closer than the other with a keypress. A new set of RDS was created for every trial. The RDS contained a mixture of correlated and uncorrelated elements whose ratio

affected the judgment of the participants. The participants had no feedback about the correctness of their responses.

In the first of the two stages in the experiment, we estimated the depth discrimination threshold while increasing the proportion of the uncorrelated elements. The participant started with 100% correlated RDS. The correlated elements then were substituted with uncorrelated elements in response to the participant's discrimination performance. The number of uncorrelated substitutes increased with two consecutive correct responses and decreased with one incorrect response by 5% ('two-up/one-down' staircase procedure). This procedure was completed with 15 reversals between correct and incorrect judgments. The estimated proportions of the correlated elements were measured with 5% density bar RDS.

In the second stage, we used the estimated correlated:uncorrelated ratio from the preceding stage to measure the percentage of correct depth discrimination in the two RDS types. Each individual participated in two sessions on different days and we added the results from the two sessions. Each session consisted of two repetitions of three 50-trial sets. Those three sets were trials using bar elements covering 5% (5% bar RDS), small grating elements covering 5% and 20% (5% grating RDS and 20% grating RDS). In each set of 50 trials, 25 trials had the estimated correlated:uncorrelated ratio from the preceding procedure and the other 25 trials had a control correlated: uncorrelated ratio in random order. In control trials, the correlated elements ranged from 70% to 95%, a range in which the participants generally had confidence in their judgment. Six sets in a session were presented in a random order. The control trials helped sustain the confidence of the

participant's judgment and reduced fatiguing from these relatively difficult trials due to a very brief presentation.

4.2.2 Results

The stimulus appeared as two surfaces at different depths and the depth order was readily resolved in both component types when the elements were 100% correlated. As the number of the uncorrelated substitutes increased, the uncorrelated components appeared to be hazy and the correlated components appeared to stand out in depth among uncorrelated elements. Near the discrimination threshold, participants experienced very weak impression of depth difference and became less confident about their judgments. Figure 4.14 shows typical data from a staircase procedure of the first stage in this experiment.

When the estimated correlated:uncorrelated ratio was adopted to the three RDS configurations, the depth impression was marginal and the participants reported to feel they were mostly guessing the depth order in the stimulus. However, the bar RDS had much higher correct response compared to the equivalent density grating RDS (5% gratings) and even to the stimulus with an equivalent number of grating elements (20% grating). Varying the number of elements (or the density) affected the performance to some degree but in every case, participants had higher correct response to the RDS composed of bar elements than the RDS composed of grating elements (19% higher on average).

The result is consistent with our prediction that the depth discrimination would be more accurate in the relatively broad band low-passed stimulus (bar) than the narrow band band-passed stimulus (grating).

4.3 Model Simulation for Transparent Surfaces

The disparity map of RDS produced with ODF model units generally shows broad disparity estimates that appear to have significant thickness; therefore, the depth edges are not sharp. The deterioration becomes worse as the spatial scale increases. Both of these are in clear distinction to human perception, suggesting the necessity of additional stereo processing. The following sections demonstrate this problem and suggest improvements.

4.3.1 Simple Simulation of ODF on Non-transparent RDS

The original ODF binocular energy model described in section 4.2 has been demonstrated to compute the disparity map from RDS (Qian, 1994). A single ODF model unit is tuned to a specific disparity but a set of ODF model units that varies interocular RF phase shifts can produce disparity estimates, once the phase of the maximally activated unit is identified. We reproduced the disparity mapping of RDS with one-dimensional input pattern for simplicity, which does not concerns the effect of different orientations and the vertical disparities. The model network is, therefore, 1-D array of ODF model units. There were 200 model units laterally positioned to cover 200 pixel width. The spatial frequency of RF is tuned for 0.024 cycle/pixel. The RDS in figure 4.16a is the 200 x 100 RDS used in the simulation, whose central region has uncrossed

relative disparity. Because the ODF model units were implemented in a 1-D array, only one line of random-dots was used at a time and the activations were averaged over 100 repetitions.

Figure 4.16b shows the disparity map produced with the single scale model units using the RDS shown in figure 4.16a. A few observations include the fact that the transition from one depth to another is smooth not sharp, and the disparity estimate for each surface is very broad. These observations deviate from our perceptual experience with the RDS used in this simulation. To improve the model response, spatial pooling was incorporated by averaging several quadrature pairs of simple cells with nearby and overlapping receptive fields (Qian and Zhu, 1997; Zhu and Qian, 1996). Although weighted pooling sharpens the depth edges, the estimates are still too broad to discriminate another surface at a nearby depth, for example a surface overlaid on another surface. Tsai and Victor (2003) resolved transparency by matching the population activity of the ODF energy model to a set of pre-established pattern of templates that are the expected response of the ODF model. The templates were created by calculating the averaged response of the model to white noise stimuli. The template matching makes greater use of the information in the population activity by keeping the pattern, than by choosing the identity of the maximally responsive neurons in the population (Qian, 1994) or by weighted averaging of depths across scale (Mikaelian and Qian, 2000). However, it is not known that the cells in early visual cortex do such calculation for comparison.

We propose a simple summation mechanism across scale that is more biologically plausible and solves transparencies in RDSs. Our model incorporates a front

end that encodes disparity by a family of ODF binocular cells and a second stage that integrates the population activity.

4.3.2 Summation of Population Response Across Scale

In the ODF binocular energy model, the units are tuned to a preferred frequency of ω_0 with a Gaussian width σ that determines the width of the RFs (see Eqs (1) in section 4.2.1). The cells in the visual cortex cover a wide range of preferred spatial frequency and bandwidth (determined by ω_0 and σ) (DeValois et al., 1982; Shapley and Lennie, 1985), and the visual system is known to analyze the stimuli through multiple frequency channels (Campbell and Robson, 1968; Graham and Nachmias, 1971). Qian and Zhu (1997) compared the disparity maps simulated with ODF model at different scales with non-transparent RDS. As the spatial scale increases, the sharpness of transition at disparity boundaries gradually deteriorates.

Psychophysical evidence indicates that disparity signals from different frequency channels interact with each other (Wilson, Blake and Halpern, 1991; Mallot, Gillner and Arnd, 1996). Although the exact mechanism used by the brain for combining scales remains unknown, computational studies have suggested ways of pooling across different scales (Marr and Poggio, 1979; Fleet, Wagner and Heeger, 1996). The simplest method is to average across the disparity maps computed by different scales (Sanger, 1988). The sharpness of disparity boundaries in the averaged map depends on the range of spatial scales that are included in averaging.

Figure 4.17a shows the disparity map produced with the single array of ODF binocular energy model at different scales. The parameters of the model simulation is

identical to those describe in the section 4.1. The period (λ) of the RFs in this simulation was 8, 12, 18, 28 and 43 pixels. Phases are generally expressed in degrees (or radians) and the phase shifts can be converted into measures of length in relation to the spatial frequency. However, we used phase-shift to refer to the disparity detection mechanism involved in the ODF model and the disparities are expressed in pixels in our experiments. Therefore, for clarification, we use the term "detector disparity" to refer to the interocular RF phase shift in the model units.

This demonstration is a 1-D version (or a cross section of the 2-D disparity map) of the Qian and Zhu (1997) simulation. One can hand-pick a RF frequency that results in a better description of the depth surfaces than other spatial frequencies. Figure 4.17c shows the average of the disparity maps from five different scales. Besides that the dimensions we used in our simulation are different from that of Zhu and Qian (1996), another important difference is that we included model units whose RF frequency is too high to detect the test disparities. Without identifying the maximally responsive units within a given scale (Qian, 1994), and with the highly activated model units whose RF periods are too small to discriminate the test disparities, it is clear that simple averaging across scale does not work better than a single scale estimate.

To investigate the problems with simple averaging, we produced the activation histograms of individual ODF units at each location across scale. Prior to the investigation across scale, we first describe the convention of the plot. Three of the activation histograms within a scale are sampled from the 6 x 17 matrix in figure 4.19 and shown in figure 4.18. Each histogram represents the activation of the ODF unit at a given location for a given detector disparity, which were cumulated for 100 repetitions.

The ODF units had RF frequency of 0.036 cycle/pixel and the detector disparities were -5, 0 and 5 pixels. The detector disparities of the two ODF units (5 and -5 pixels) were closely matched to the binocular disparity assigned to the central region and the surrounding of the input RDS (see figure 4.16a). Sets of ODF units that had matching detector disparity to the stimulus disparity showed increased activation at the corresponding locations.

Figure 4.19 shows a 6 by 17 matrix of activation histograms covering 5 scales and the simple sum, and detector disparity ranged from -8 to 8 pixels. In actual simulation, the activation was computed from -20 to 21 pixels of detector disparity producing 6 by 42 matrix, however, only 16 of them are shown in figure 4.16 due to the space limit. The RDS in figure 4.16a were used to stimulate the model units.

Among the five RF frequencies that were subsumed in the simulation, the two RF frequencies 0.056 and 0.036 cycle/pixel appears to describe the two different depths in RDS input (-6 and 6 pixels) optimally. However, when the responses of units having equal detector disparity are summed across scale at each location (the bottom row in figure 4.18), the elevated activation of the high RF frequency units overwhelmed the discrete responses of the two optimal frequencies. In the literature (Qian, and Zhu, 1997, Micaelian and Qian, 2000), when the averaging or weighted summation was implemented, the fine scales that were too small to detect the testing disparities were often excluded to avoid this undesired effect. Therefore, our first goal in the model simulation was to develop a second stage mechanism that sharpens the distribution of disparity estimates, which also affects the sharpness of the disparity boundaries without excluding units with high RF frequencies.

4.3.3 Improving the Model by Removing the Monocular Component in the Responses

We propose a simple mechanism that removes the monocular component in the response of ODF model unit. In this mechanism, the monocular components are removed by inhibitory connections between the two model units within the same scale, whose relative detector disparity is equal to half the period. Figure 4.18 shows the activation histograms of ODF model units at different detector disparities for two different RF frequencies. The random-dot stimulus had an uncrossed disparity of 5 pixels. In the first row, the maximum activation occurs at a detector disparity of 5 pixels as expected. The minimum activation occurs at a detector disparity of -1 pixel, a half period (12 pixels) away from the maximum activation. In the second row, the RF period is 18 pixels and the minimum activation occurs at 9 pixels which is also a half of the period away from the maximum activation.

These minimum activations do not result from the binocular interaction but from pure monocular responses and their contribution to the overall response is further aggravated due to the squaring in the ODF model. Once subtracted by means of inhibitory connections ("π-inhibition"), the resulting response is free of monocular components. The core mechanism in the second stage is applying π-inhibition within a scale before summing across scale. The summation of these π-inhibited responses produces much narrower disparity estimates. Figure 4.16d shows the sum of the disparity estimates across scale whose monocular components are removed by the π-inhibition.

. This method works for broad range of arbitrary disparities and the result is not disturbed by including high frequency RFs which is to high detect the stimulus disparity. In figure 4.21, the simulation of the model network incorporating the π–inhibition and

inter-scale summation shows significantly improved disparity-mapping with both minimum stimulus disparity of 1 pixel and another larger arbitrary stimulus disparity of 10 pixels. The result would be further enhanced if a more number of RF frequencies were employed in the averaged pool.

4.3.4 Inter-scale Summation With π-inhibition Discriminates Transparent Surfaces

The ODF model by itself does not work with transparent surfaces because of the broad disparity estimates for a given stimulus disparity. With the second stage that sharpens the disparity estimates by removing the monocular component from the energy model response, we tested whether our improved model discriminates transparent surfaces at two different depths. The model network consisted of an array of composite models described in section 4.4.3. The composite model cell is a collection of several scaled model units. Each member of a composite unit has an inhibitory connection with another member whose detector disparity differs by is one half the RF period or π radians.

We created two types of RDSs that elicited perception of transparent surfaces: 1) two layers of random-dots at different depths (double-layered RDS); 2) Panum's limiting case RDS (PLC RDS), which are shown in figure 4.22. A line of random-dot array from each 200 x 200 RDS was used to calculate the model response and repeated 200 times.

Figure 4.23 shows the results of the model network with non-transparent and transparent surfaces. Figure 4.23a is a disparity map of a non-transparent single layer provided for comparison. The disparity map of double-layered RDS is shown in figure 4.23b. The five rows from top to bottom show the scaled raw model responses prior to applying π-inhibition and the bottom row shows our improved model response. The

disparity map shows clear dissociation of the two surfaces in the improved model simulation. Figure 4.23c shows the results with a transparent PLC RDS. The pattern of responses of scaled raw response prior to applying π-inhibition and our improved model are almost identical to those in the transparent double-layered RDS.

Investigating each scaled model responses in transparent RDSs, one notices that there is one scale that captures the separation between the two surfaces (second and the third column in the gray rectangle), the disparity map in the second row (RF period = 12). This response map appears to contribute significantly to the final summed response. However, a very similar response pattern is observed in the response to non-transparent single layer RDS at same scale (first column in the gray rectangle). Similarity in the response pattern despite the different depth structures accentuates the fact that the disparity tuning is often ambiguous within a single scale. The disparity maps produced with our model show almost identical results for the two transparent RDS types, which is consistent with our perception with these stimuli. The minimum separation between the two layers that was identified with the model network was 5 pixels apart, which is a pixel larger than half the minimum RF period of the composite model.

We ran additional simulations using anti-correlated RDSs. Figure 4.24a and figure 4.24b show the results from the single layer anti-correlated RDSs whose stimulus disparities were 0 pixel and 6 pixels, respectively. The spreading of the disparity estimates was common in both disparity maps. With larger RF dimensions and smaller RF period incremental, the spreading of the disparity estimate would appear ideally smooth. The response of zero disparity also would appear symmetric across horizontal axis.

Figure 4.24c and figure 4.24d show the disparity maps of the anti-correlated PLC RDSs. When the separation in PLC configuration was small (figure 4.24b), the disparity estimates formed a single layer on the opposite side of the expected depth. The disparity estimates were as well-confined as those of the correlated non-transparent layer and the transparent layers shown in figure 4.21 and figure 4.23. However, spreading of the disparity estimates increased as the separation became greater (figure 4.23b), yet not as continuous as the anti-correlated RDS in figure 4.24b. These two results show an approximation of our perceptual experiences with these RDSs, because we see a single surface at opposite depth at small separation whereas we see an amorphous cloud of dots at greater separations in anti-correlated PLC RDS.

4.3.5 Dependence of Depth Reversal on the Phase-disparity Mechanism

Our model simulations incorporated the well-established binocular model of Ohzawa, DeAngelis and Freeman. This model subsumes phase-disparity detection mechanism for their model complex cell. In real complex cells, however, the disparity detection mechanism also includes position-disparity. In position-disparity mechanism, the locations of the complex cell RFs (Gaussian envelopes) do not correspond in the two eyes. Instead, the difference of the two positions provides the disparity information. Recordings in monkeys and cats show that both mechanisms are responsible for disparity-tuning in early visual cortical area (Prince, Cumming and Parker, 2002). Subsumption of the position-disparity instead of phase-disparity in our model can lead to different results when the position shift of the RF envelope (Gaussian function that determines the width and the position of the RF) exceeds half a RF period. The peak of

the disparity-tuning curve for an anti-correlated stimulus does not occur on the opposite side and the reversal of the depth may not occur.

One of the undesired effects in phase-disparity detection is the wrap-around of the disparity estimates. This occurs because the phase-disparity detector does not distinguish the two different phases of $\pi+\alpha$ and $-\pi+\alpha$. This is noticeable in several simulations, figure 4.17a (the first second and third rows) and figure 4.25 for example. This makes the single scale model response ambiguous from the transparent surface cases (see examples captured in gray rectangle in figure 4.23). Note that the roll-over of disparity estimate is also eliminated in our model responses besides the sharpening of the disparity estimates in figure 4.24.

The RF sensitivity profile of the simple cells which are the basis of the image detection in the model varies widely in real simple cells as well (Anzai, Ohzawa and Freeman, 1999). The specific function that ODF energy model chose is Gabor. The parameters of Gabor function that describes RF of the simple monocular cell in our simulation are adopted from the commonly occurring values in the literatures (Ohzawa et al. 1990; Anzai et al., 1999; Qian, 1994; Zhu and Qian, 1996; Read and Eagle 2000; Hayashi et al., 2003), which affect the response of the ODF model. However, the specific numbers and parameters of filters mentioned above are not critical in the results. The inter-scale summation and the π-inhibition are the two key processes that led to the distinction of the two depths in transparent surfaces.

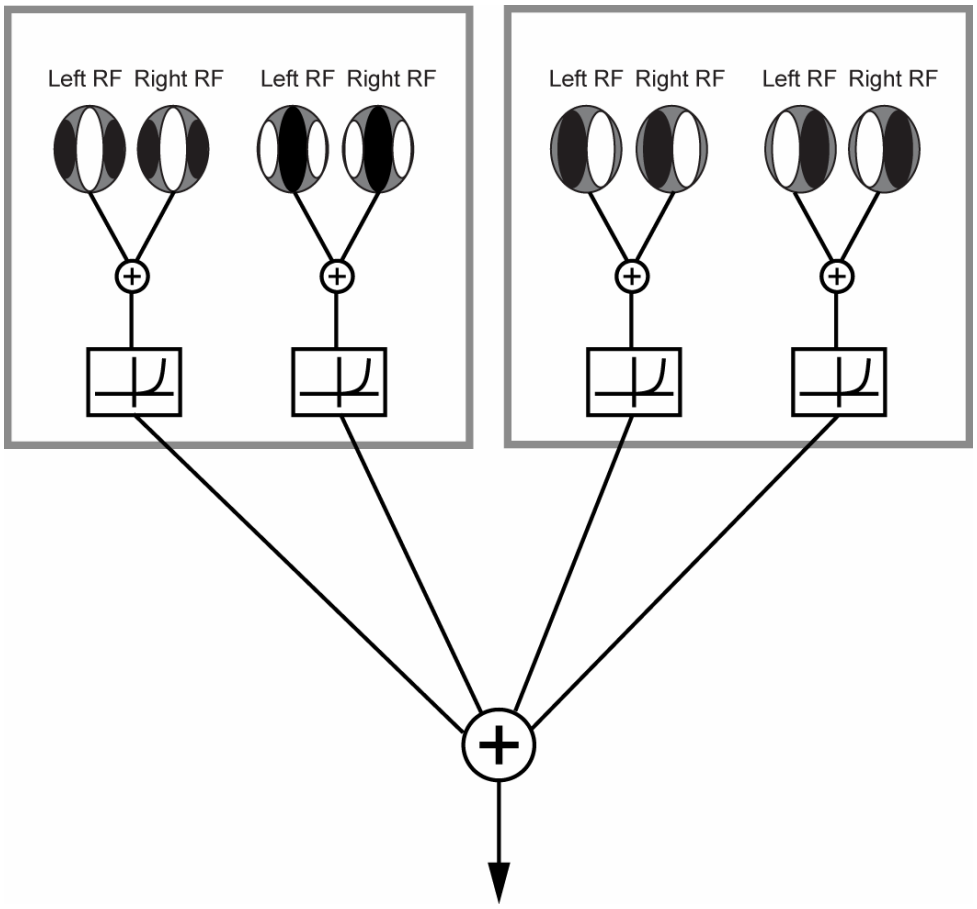


Figure 4.1. Ohzawa-DeAngelis-Freeman binocular energy model. The model consists of 2 major units (each of which is enclosed by a gray rectangle) that are in quadrature, i.e., spatial phases of monocular RFs for 1 unit and those for the other are 90° apart. The sum of the two monocular units are halfwave-rectified and then squared. The final output is the sum of the outputs of the four pairs.

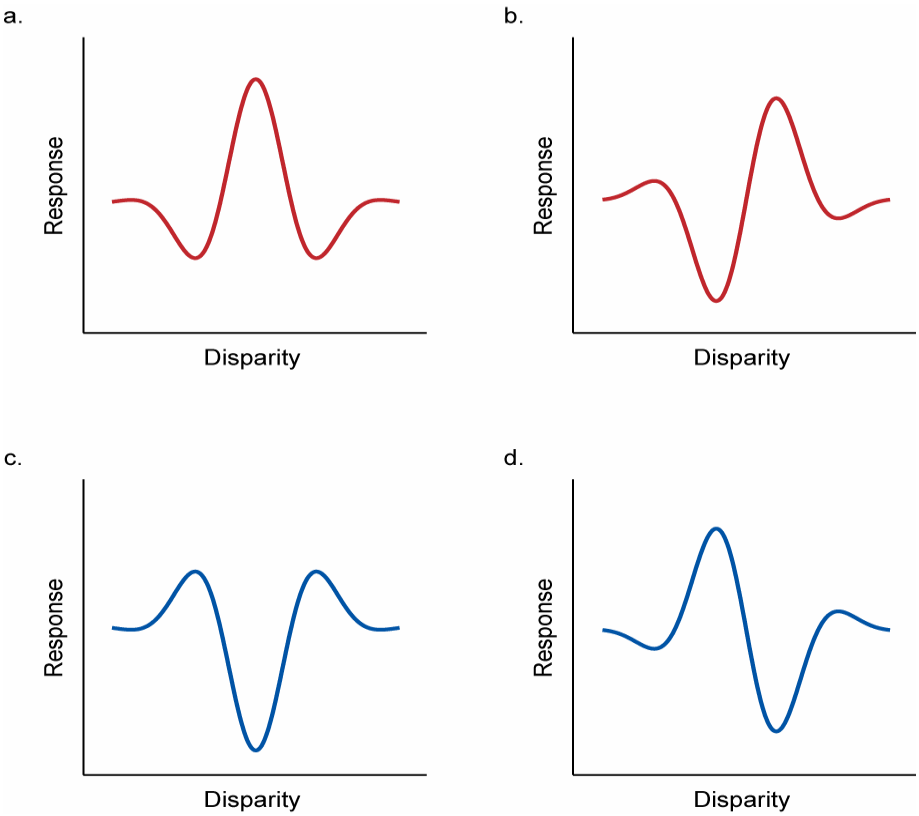


Figure 4.2. Disparity tuning curves obtained with ODF model cells. a. “Tuned excitatory” type cell. b. “Far” type cell. c and d. Disparity tuning curves of the same model cells, respectively, in response to the anti-correlated stimulus.

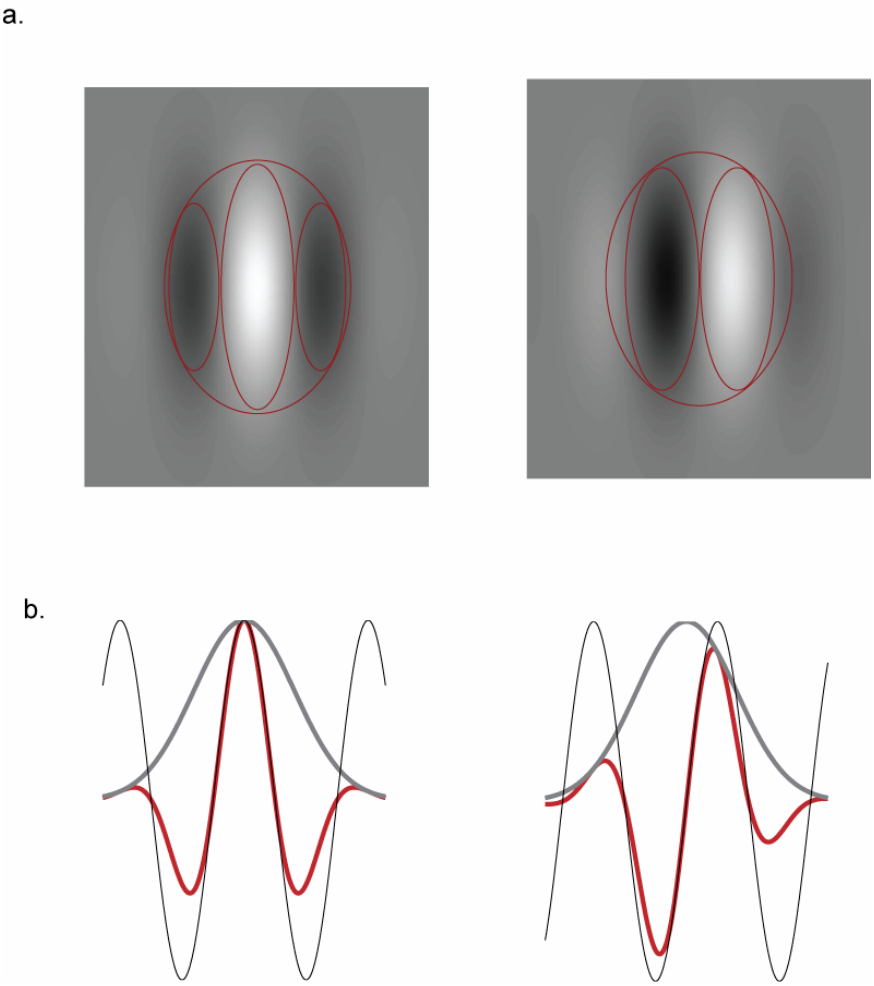


Figure 4.3. Modeling of simple cell receptive fields. a. A simple cell receptive field includes elongated subregions which respond to a bar stimulus that is brighter or darker than the background. The light regions are sensitive to light bar stimulation and the dark regions are sensitive to dark bar stimulation. b. The sensitivity profiles are modeled with Gabor functions (red line). The thin black line is the carrier function (sine) and the thick gray line is the envelope function (Gaussian).

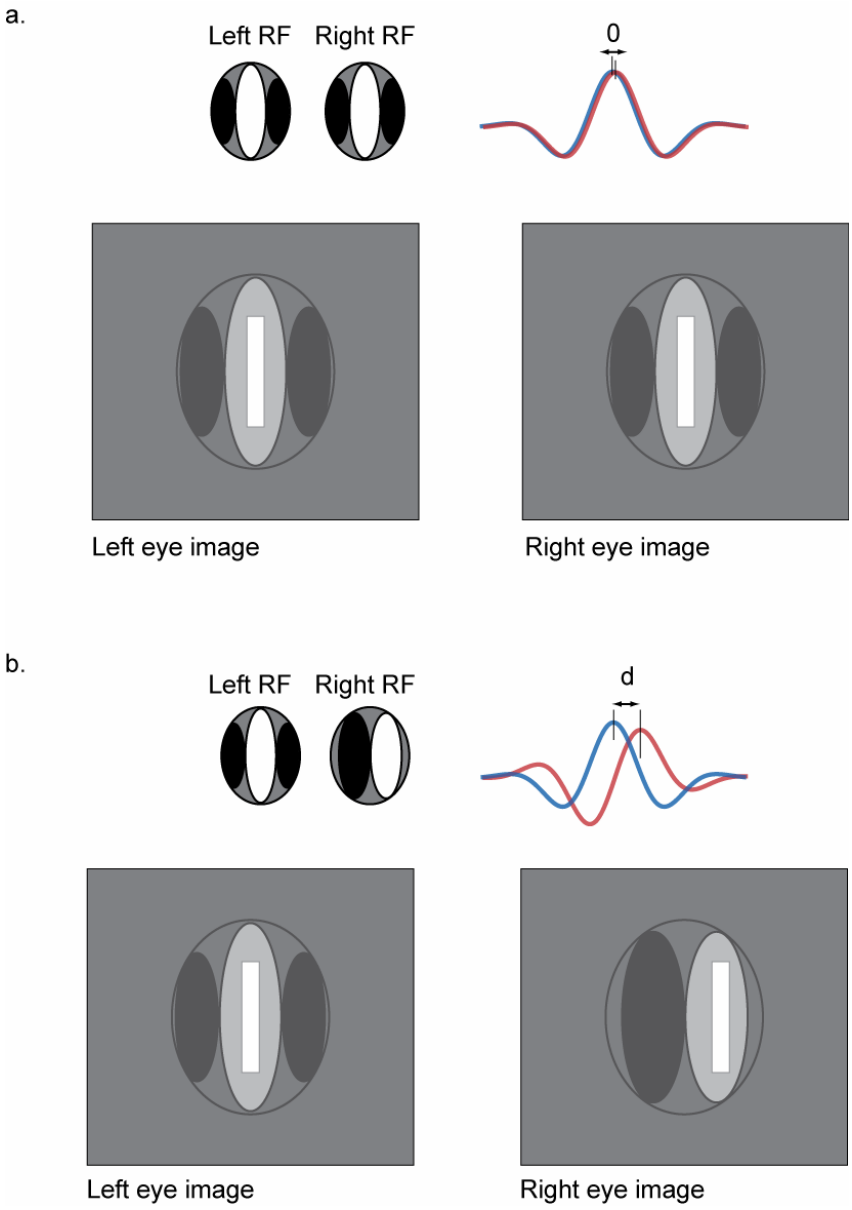


Figure 4.4. Schematic diagrams of disparity detection in the binocular simple cells. The RF profiles are denoted in the upper left corner. The RFs are overlaid on to the white bar stimulus. a. The simple cells that have identical RF profiles can detect zero disparity. b. The simple cells that have their RF phase shifted in respect to the other eye can detect the disparity equal to that phase shift. (blue: left RF, red: right RF, d: interocular phase difference)

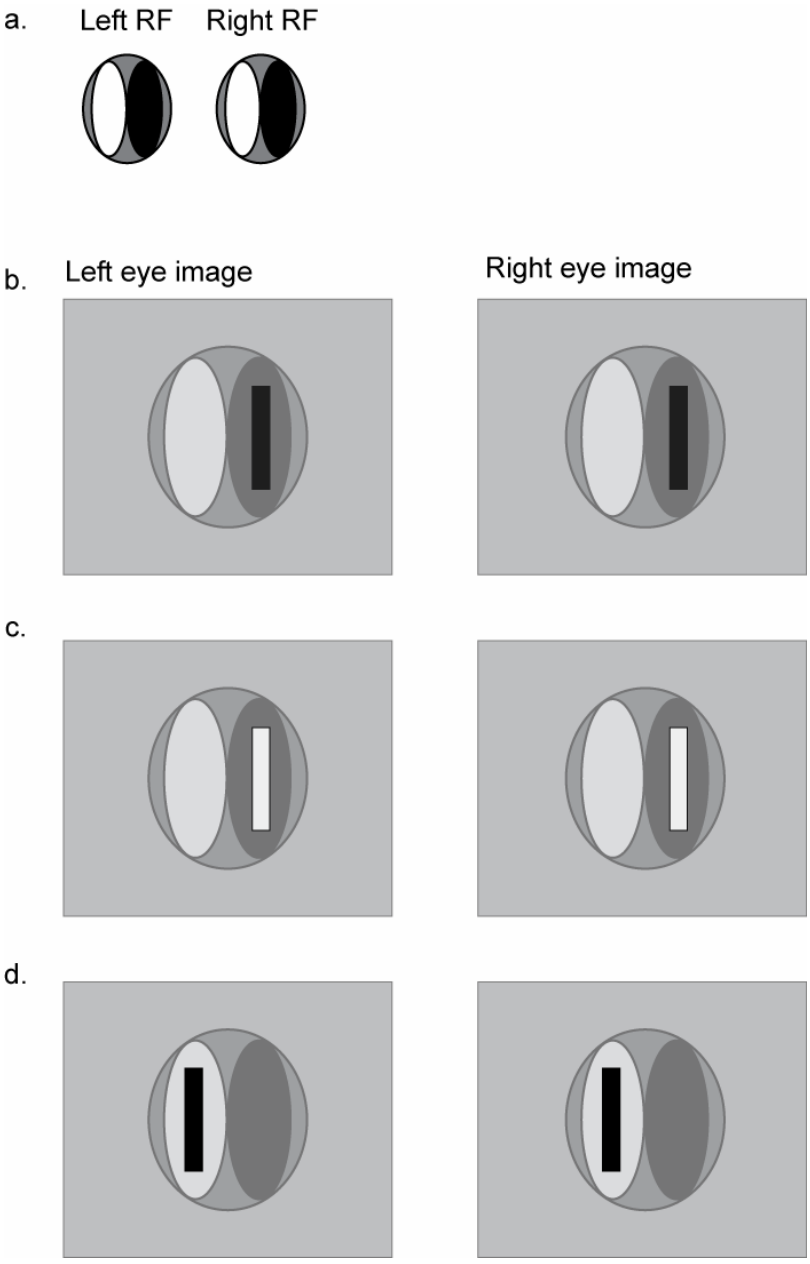


Figure 4.5. The response of the binocular simple cell is dependent on the position and contrast polarity of the stimulus within its RF. a. A binocular simple cell that is tuned at zero disparity. b. The optimal position of the dark bar stimulus. c and d. The two zero-disparity stimulus sets that are within the RFs do not induce good response.

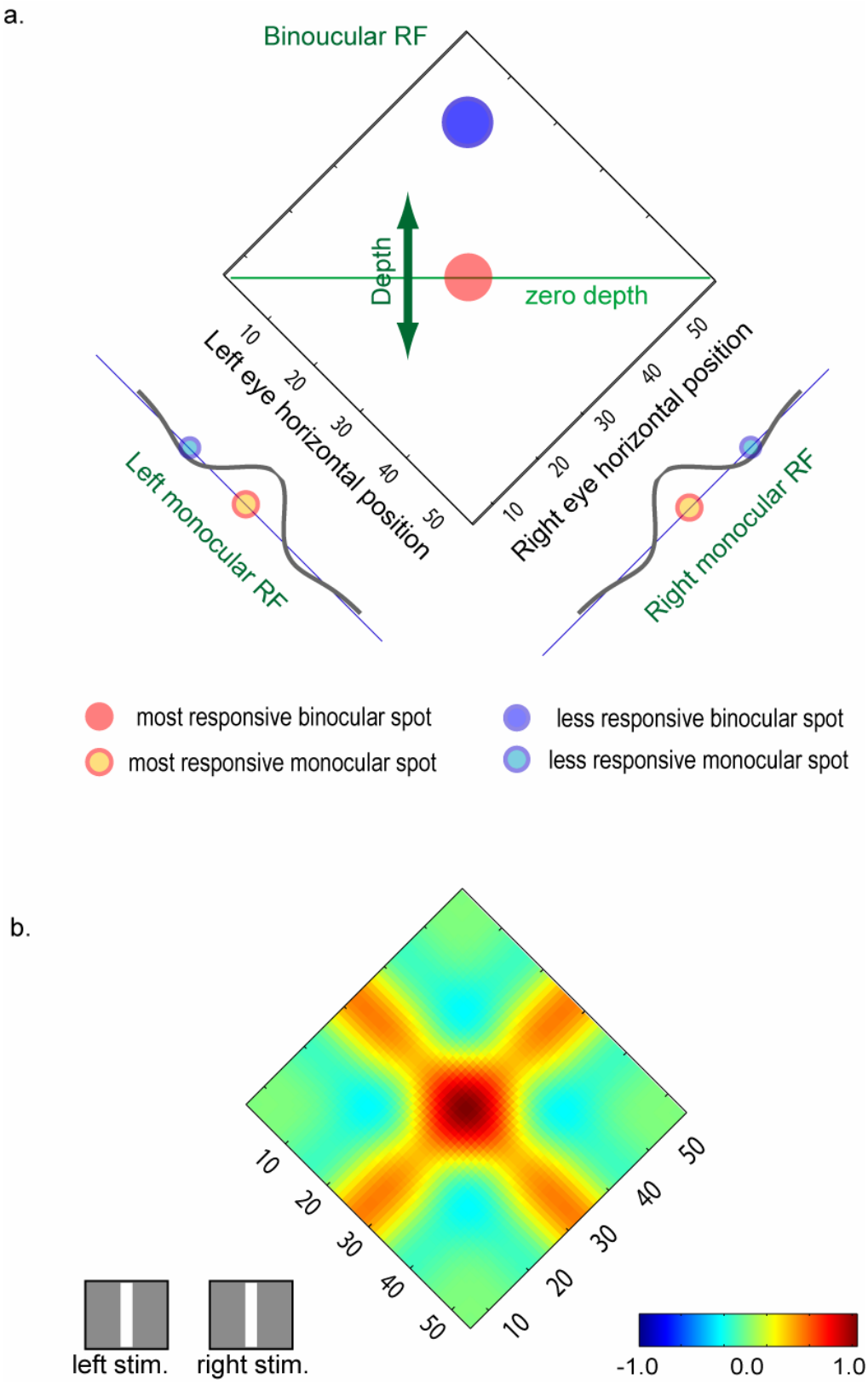

a.
Binoucular RF
Depth
zero depth
10
20
30
40
50
Left eye horizontal position
Right eye horizontal position
Left monocular RF
Right monocular RF
most responsive binocular spot
less responsive binocular spot
most responsive monocular spot
less responsive monocular spot
b.
left stim.
right stim.
-1.0
0.0
1.0

Figure 4.6. Disparity-tuning map of a model binocular simple cell. a. Convention of the plot. The plot is rotated for -45 degrees for easy indication that the left and the right bottom axis correspond to the horizontal locations of left and right eye. This simple cell unit responds maximally to the bright bar at the center of both monocular RFs (red spot). b. Full description of the disparity-tuning. The RF profiles are identical to the two Gabor functions shown in the bottom of b. The response map is produced with binocular white a bar at all locations in the RFs.

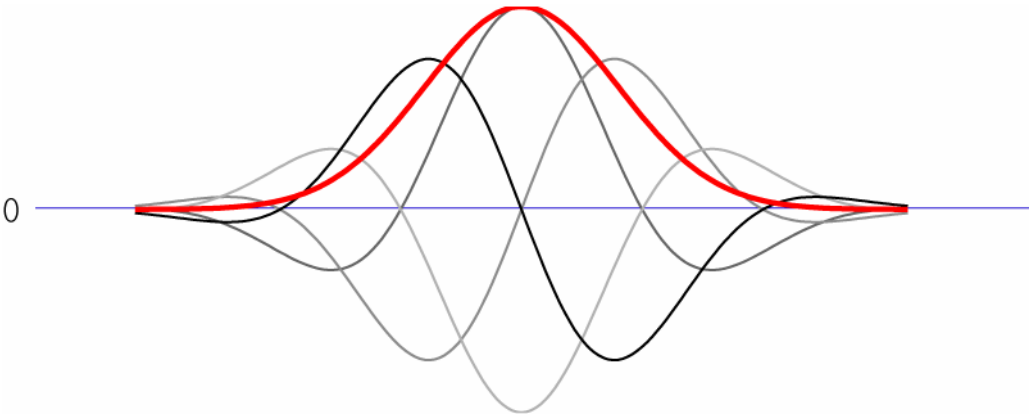


Figure 4.7. Combination of four Gabor functions. The sum of four squared Gabor functions is broader than the individual Gabor function.

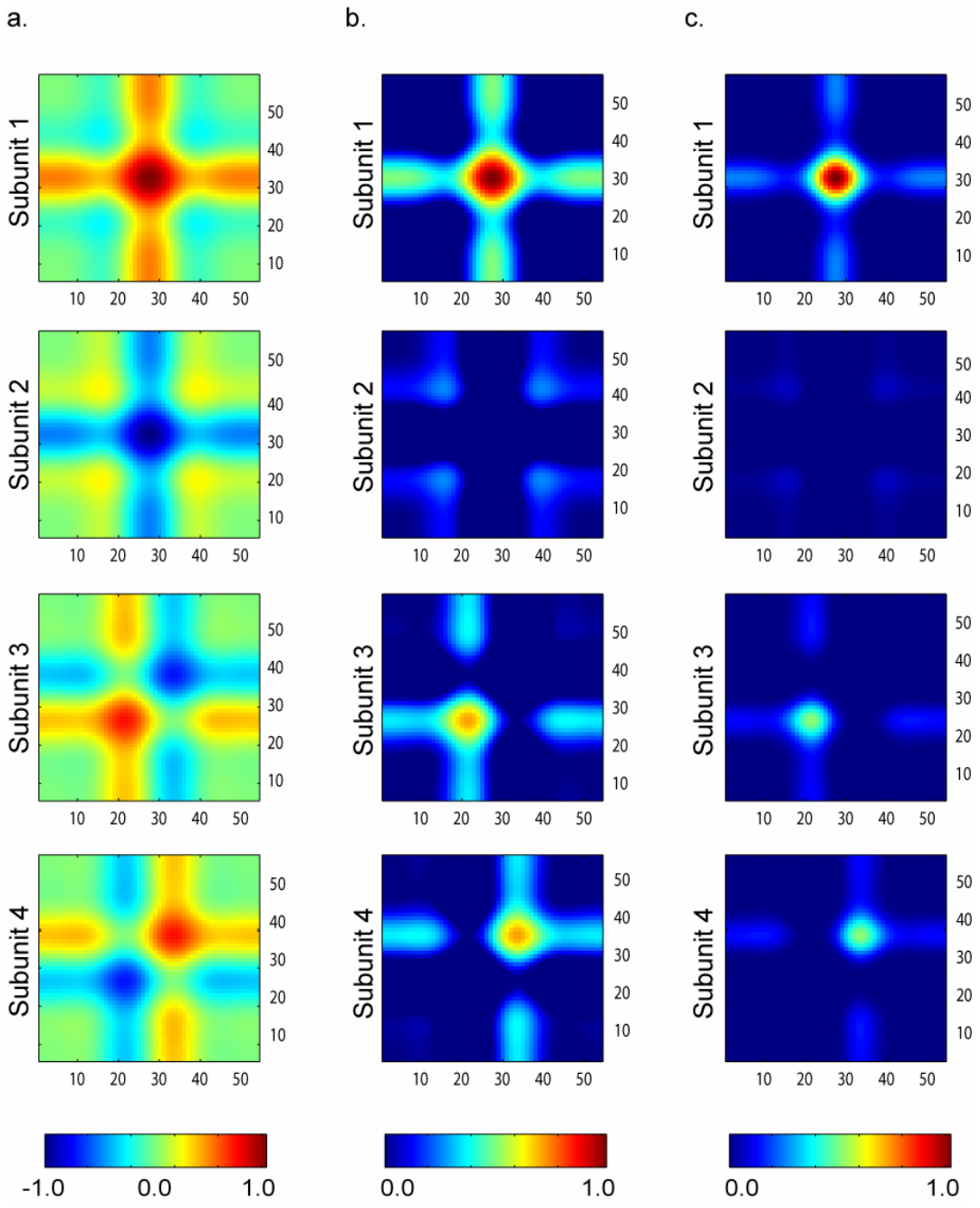


Figure 4.8. Disparity-tuning maps of four subunits in the ODF energy model. The ODF unit is tuned for zero disparity. a. Responses calculated from the simple sum of left and right monocular simple cell activation with a bright bar. The activation is both negative and positive. b. The simple sum is halfwave-rectified. c. The halfwave-rectified sum is squared. The monocular activations become less significant after half squaring.

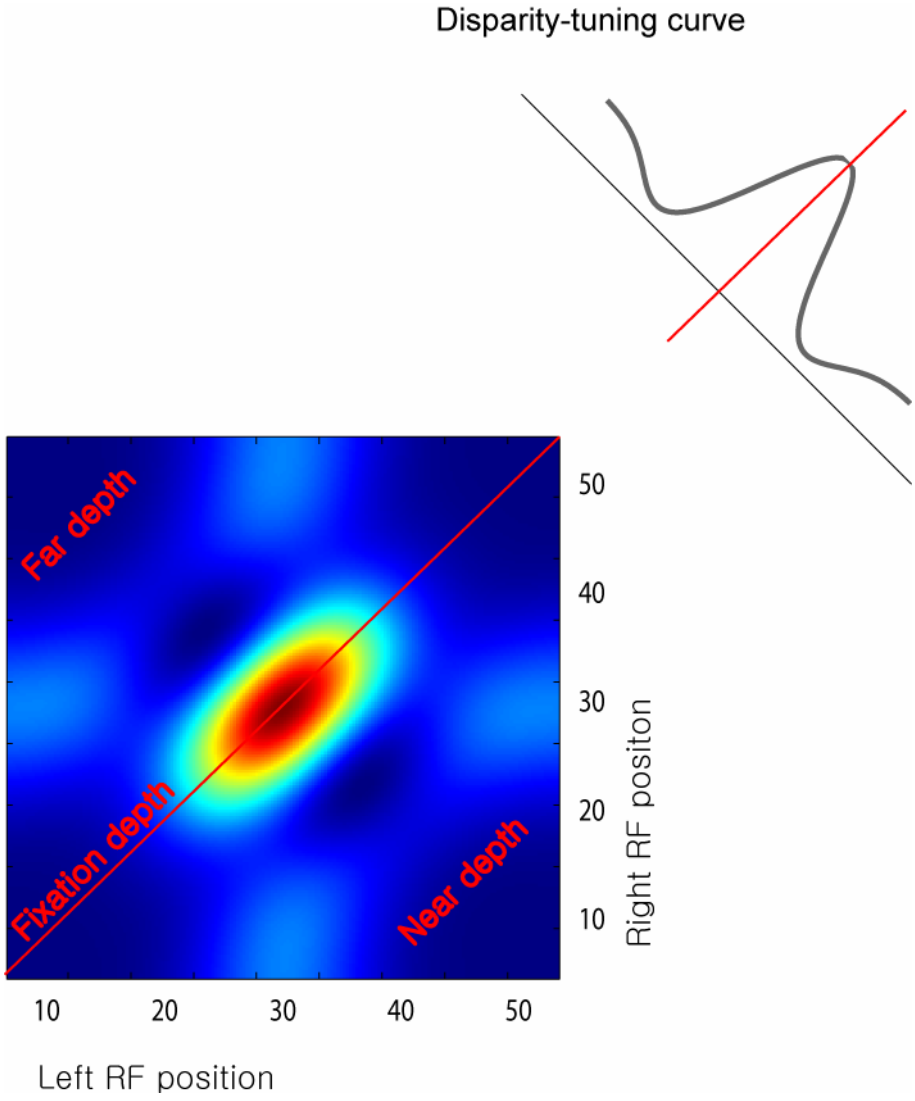


Figure 4.9. Typical ODF model response. The model unit is tuned at zero disparity. The summed response of four subunits that are half-wave rectified and squared (see figure 4.8c) is broadly tuned within the RF.

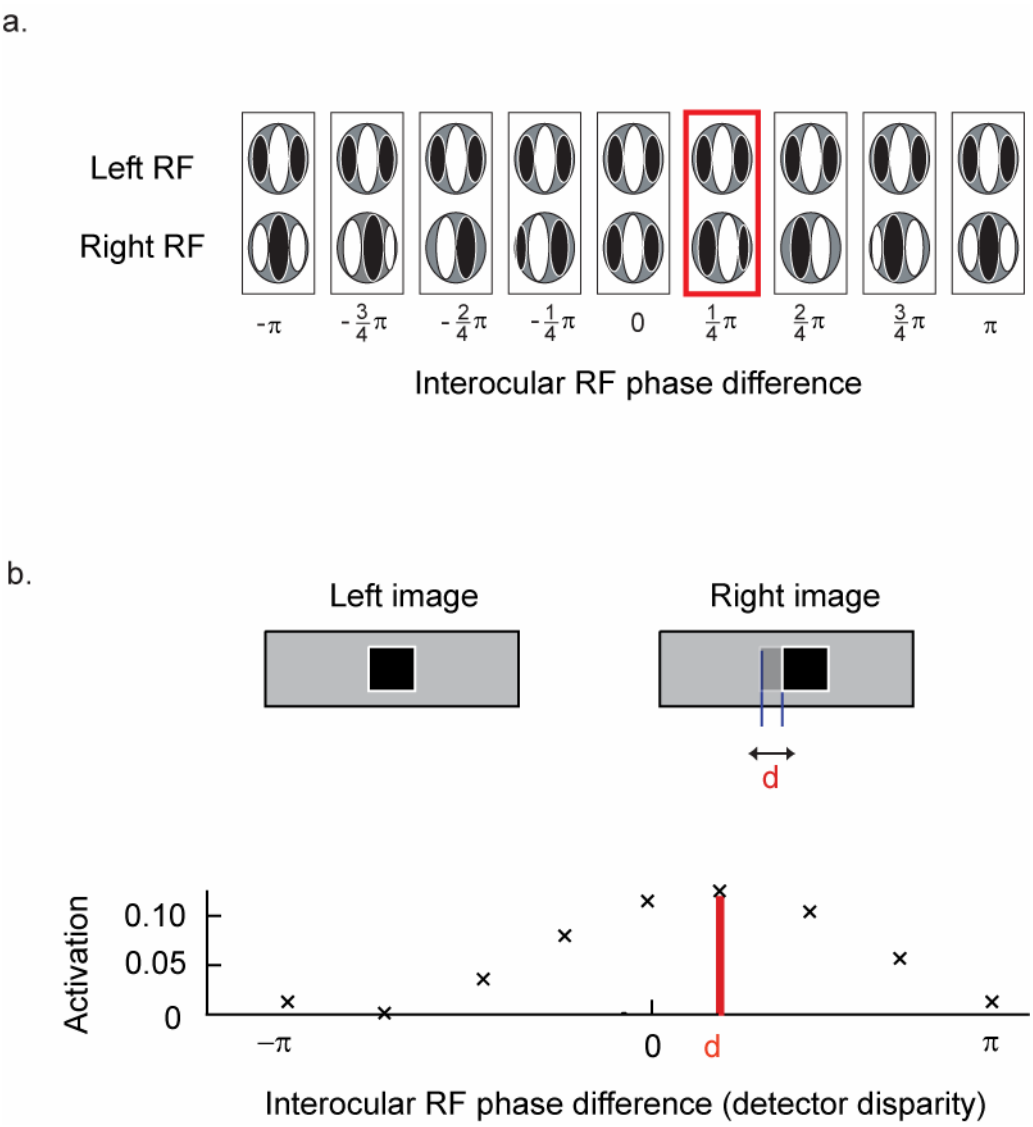


Figure 4.10. Interocular phase shift in detectors and the magnitudes of activation in response to the binocular stimulation. a. Schematic cartoons of the interocular phase difference in the left and right RFs. The interocular RF phase difference determines the binocular disparity of the detector unit. The nine ODF units are stacked at the same location. b. The disparity in the binocular stimulus (d) maximally activates the detector that has the identical degree of phase shift in RFs. The red bar denotes the magnitude of the maximum activation in the detector whose phase difference is captured in the red rectangle in a.

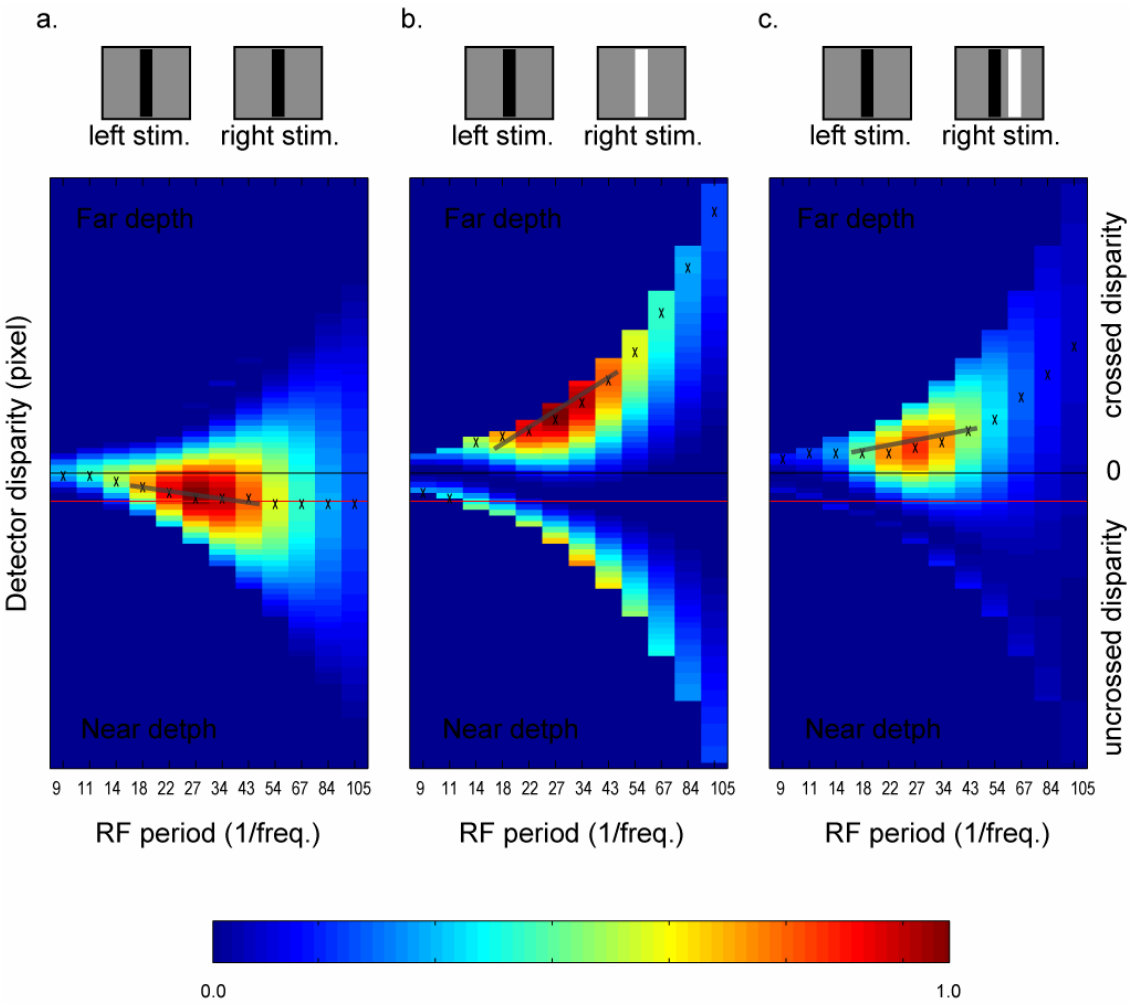


Figure 4.11. Distribution of the peak response across scale in ODF binocular energy model. The pair of panels above each figure is the stimulus used in each model simulation. bar width 15 pixels and disparity was 5 pixels. a, b,and c. Activation map of the model units tested with same contrast bar stimulus, opposite contrast bar stimulus and PLC stimulus, respectively. The peak activation of the stimulus with opposite contrast component takes place in the opposite side of the expected depth with same contrast stimulus (b and c). The peaks at different scales are clustered in the stimulus configurations that give rise to consistent depth impression, either veridical or reversed depth impression. (red line: the disparity in the stimulus, x: the peak response at the particular spatial frequency)

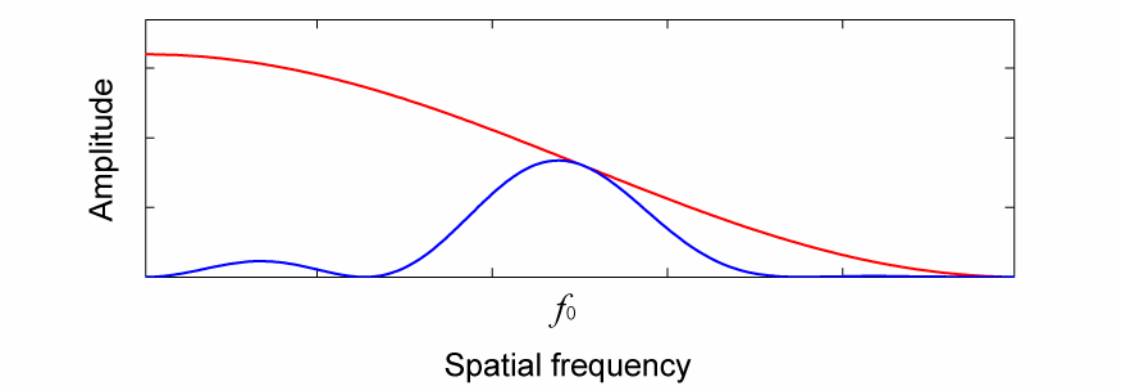


Figure 4.12. Power spectrum of a bar and a grating element of the RDS. The distribution of spatial frequency components in the bar covers wide range whereas that in the grating is focused around the fundamental frequency (f_0 = 1/(bar width*2)).

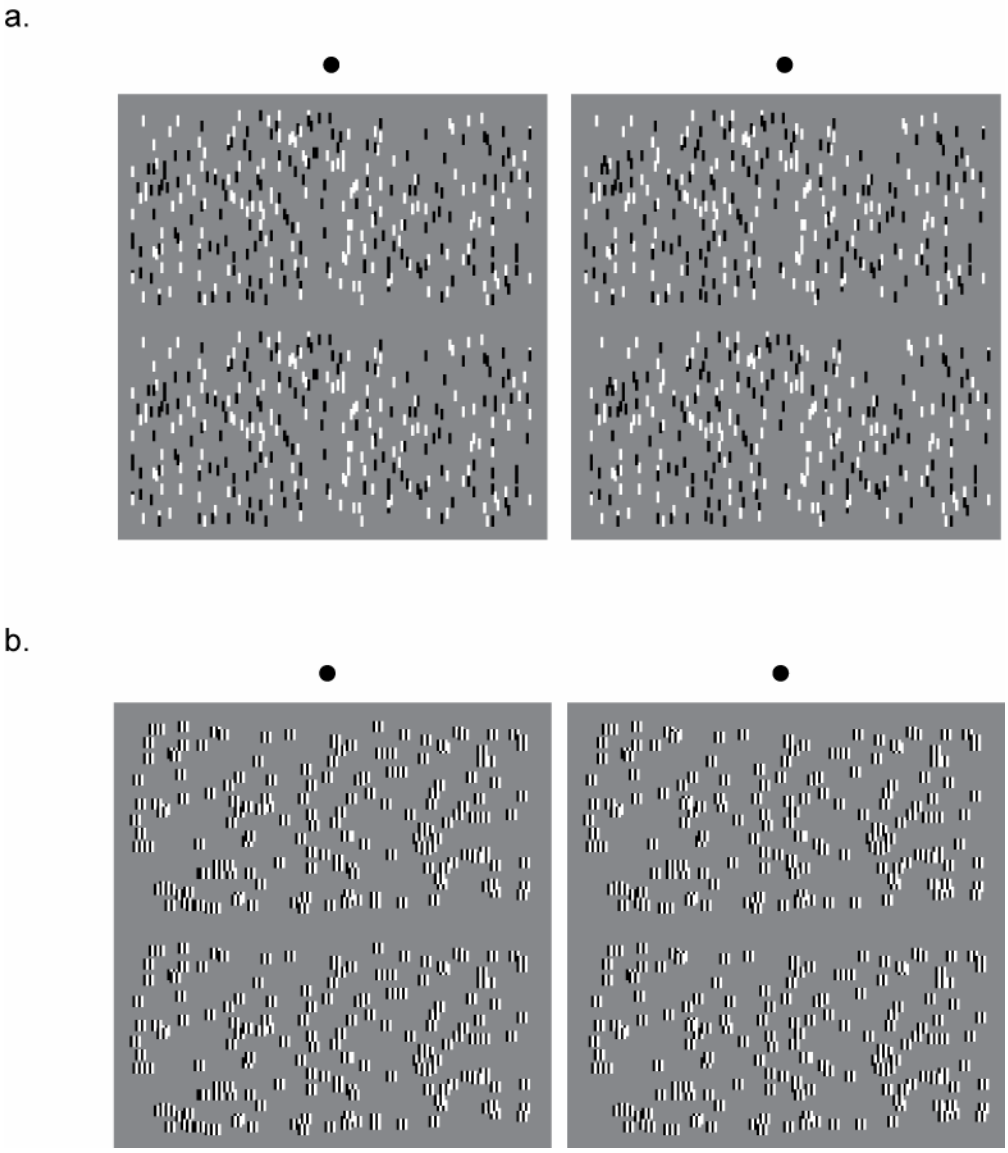


Figure 4.13. Typical RDS variants used in the experiment. The upper and lower halves of the stereograms are identical copies that are assigned different disparities. The participants determined which half appeared closer. Both RDSs contained varying amount of uncorrelated elements that made the task difficult when presentation was very brief. a. RDS composed of black and white bars. b. RDS composed of black-white-black-white gratings.

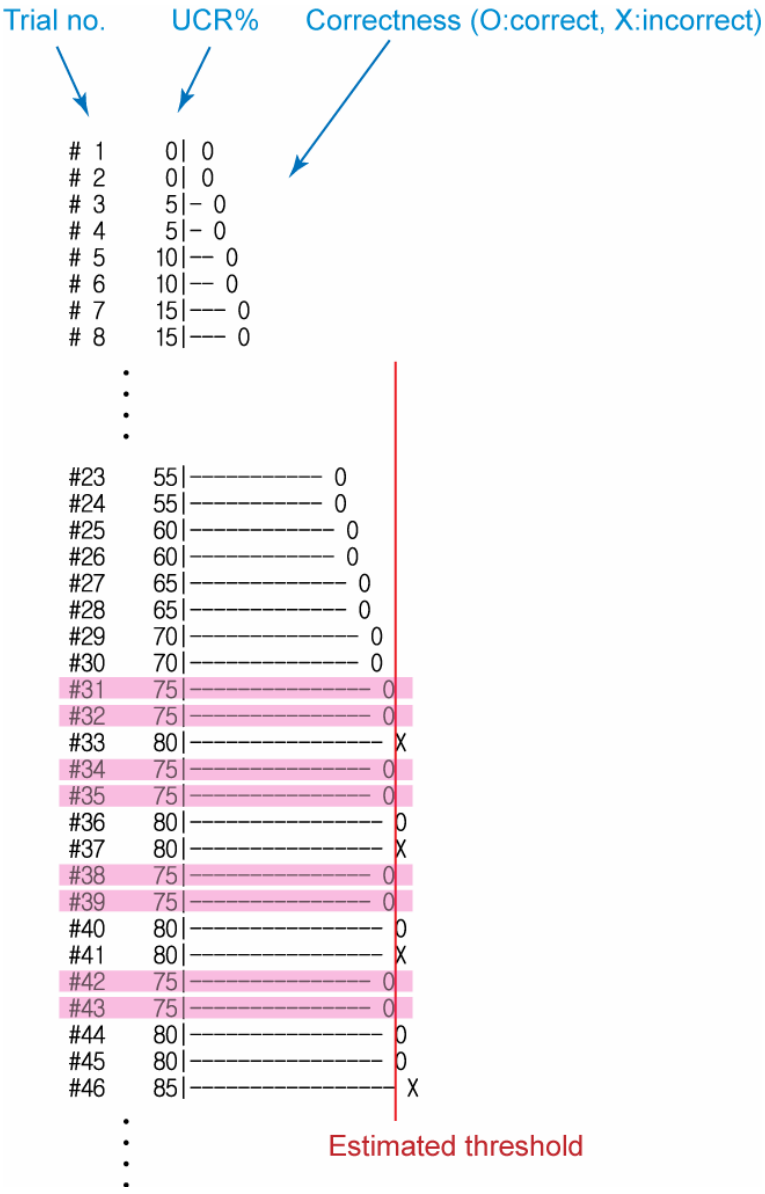


Figure 4.14. Typical data from a staircase procedure to determine the threshold in depth discrimination tasks. The participant reported which one of the two copies of neighboring RDS appeared to be closer than the other. The correctness of the response increase/decrease the proportion of the uncorrelated substitutes for the correlated elements. The shaded rows are the trials of maximum uncorrelated elements up to which the participant generally reported correct response.

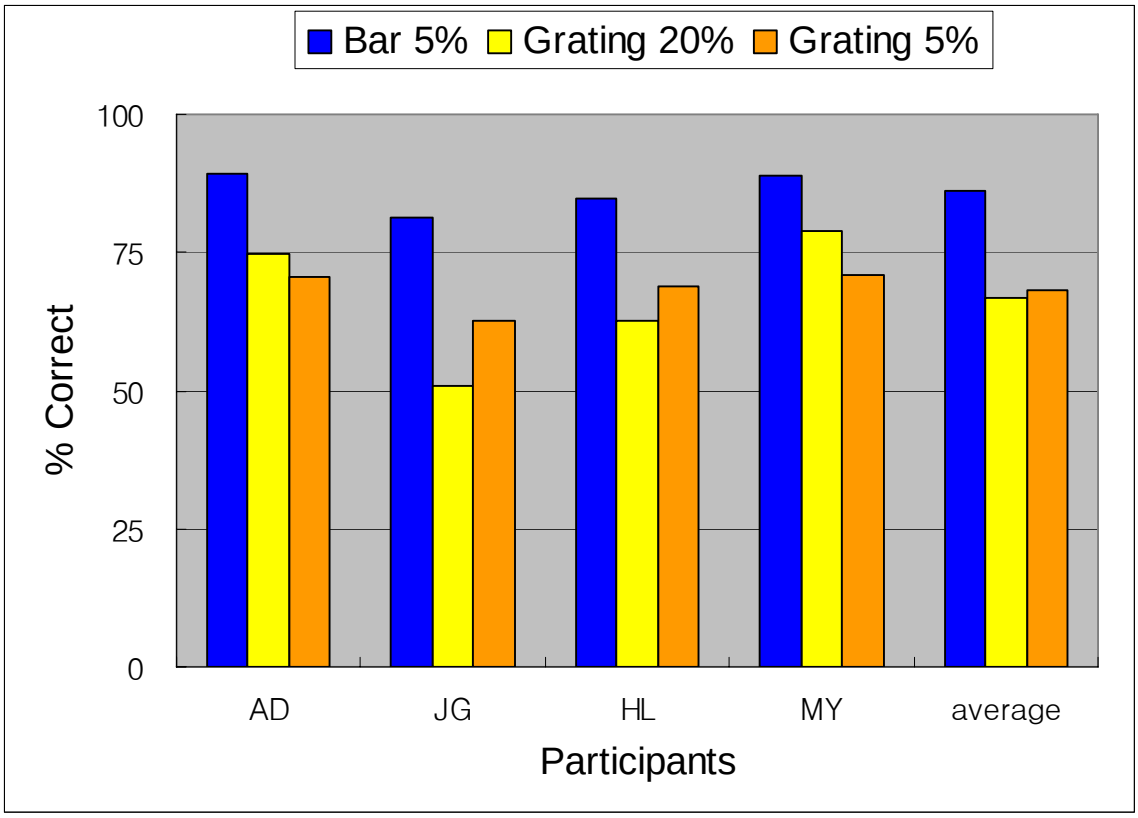


Figure 4.15. Proportion of trials in which participants made correct judgments of the depth order of the two neighboring RDSs. The RDSs contained a mix of correlated and uncorrelated elements. The RDS was composed of black and white bars covering 5% and small black-white-black-white gratings covering 5% and 20%. The discrimination accuracy was significantly higher for the bar component that the other two configurations for all participants.

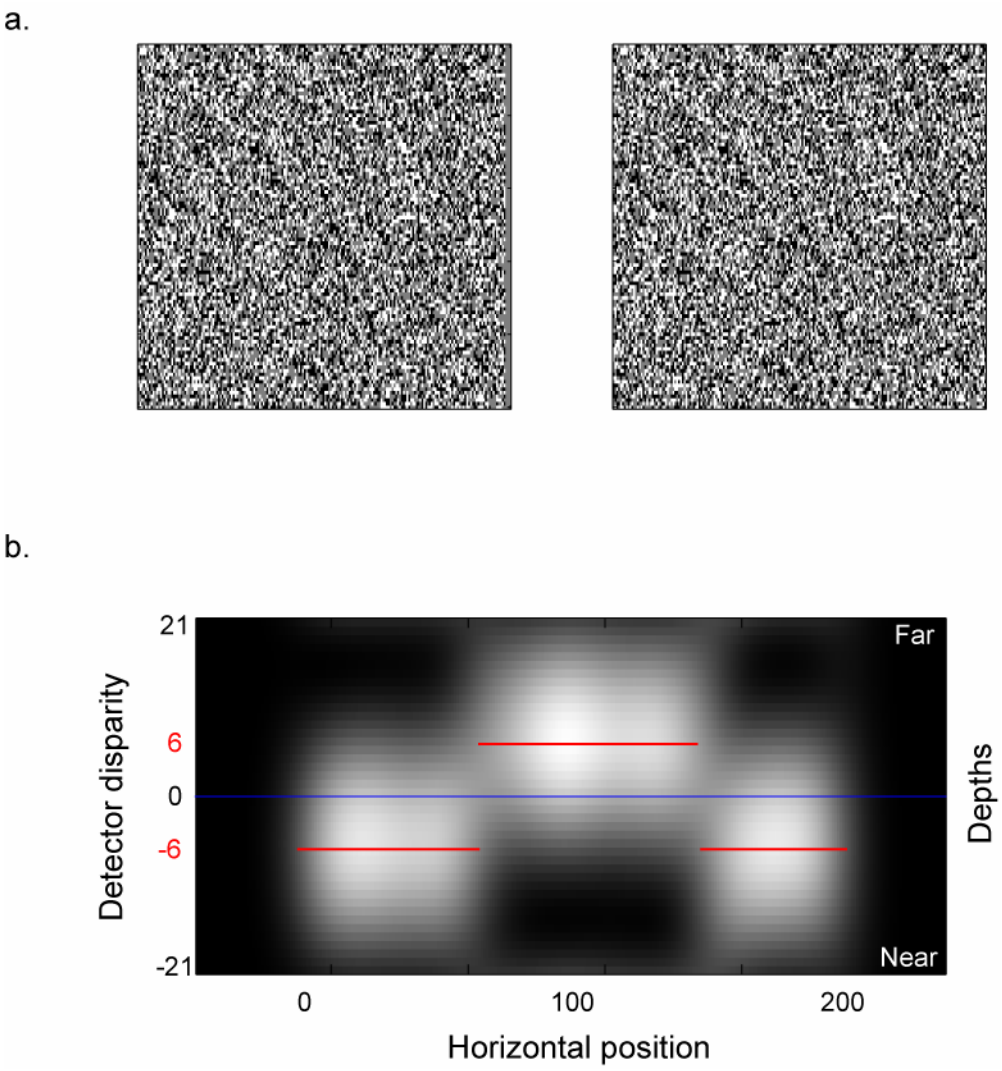


Figure 4.16. Disparity map of RDS. a. A 200 x 100 RDS with the central region laterally shifted. b. A 1-D array of the ODF binocular energy model cells is simulated with each line of random-dot in the stereogram above a. and the responses are averaged over 100 repetitions. (red line: stimulus disparity)

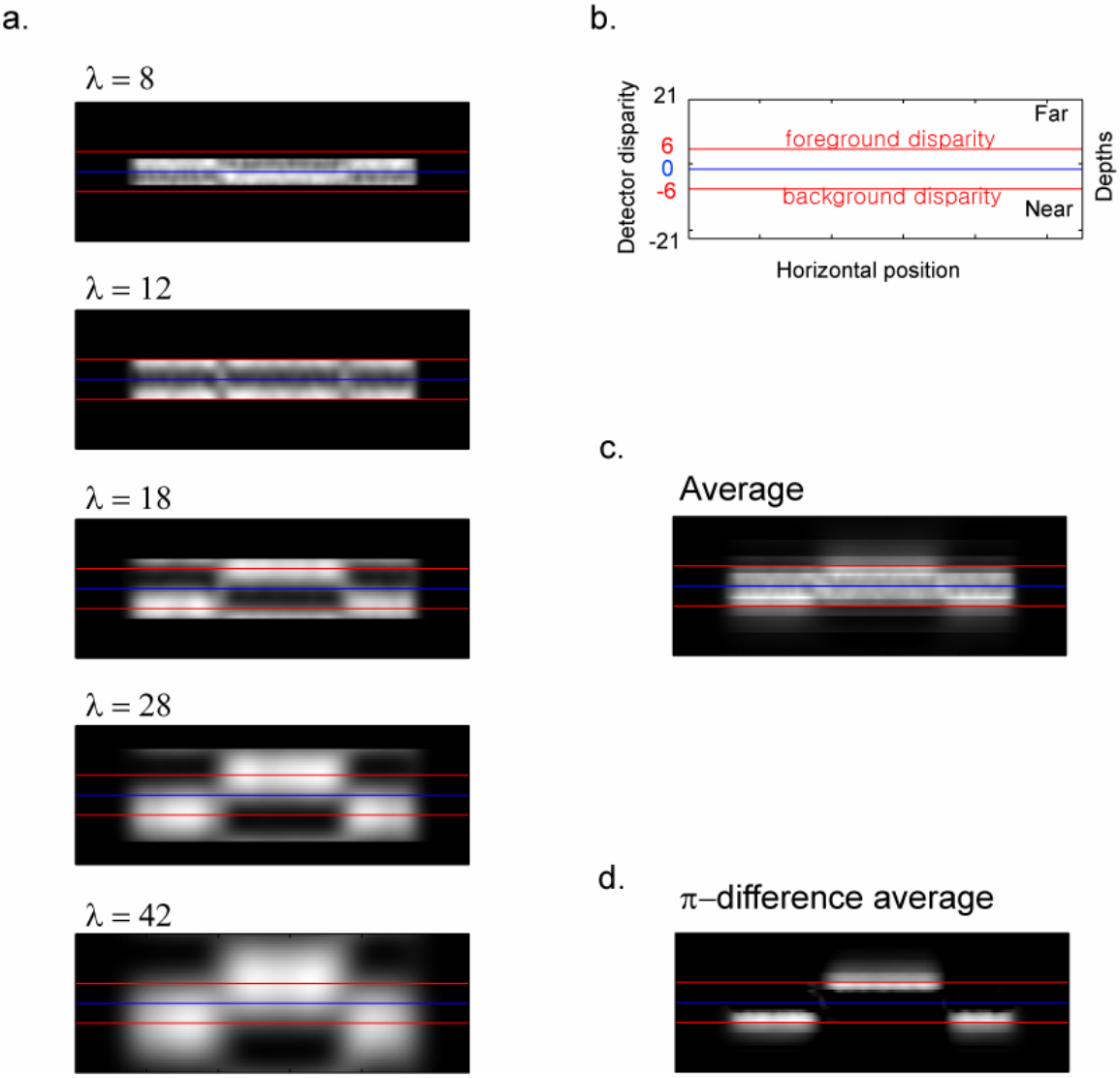


Figure 4.17. Disparity maps produced with suggested models. (red line: stimulus disparity) a. Disparity maps produced with ODF model at different scales. The RDS used in the simulation is similar to the one used in Figure 4.14.a. (λ = RF period). b. The conventions of the graph which is similar to the one used in Figure 4.14.b. c. The average of the five responses in a. d. The average of the improved model suggested in the text (see section 4.4.3).

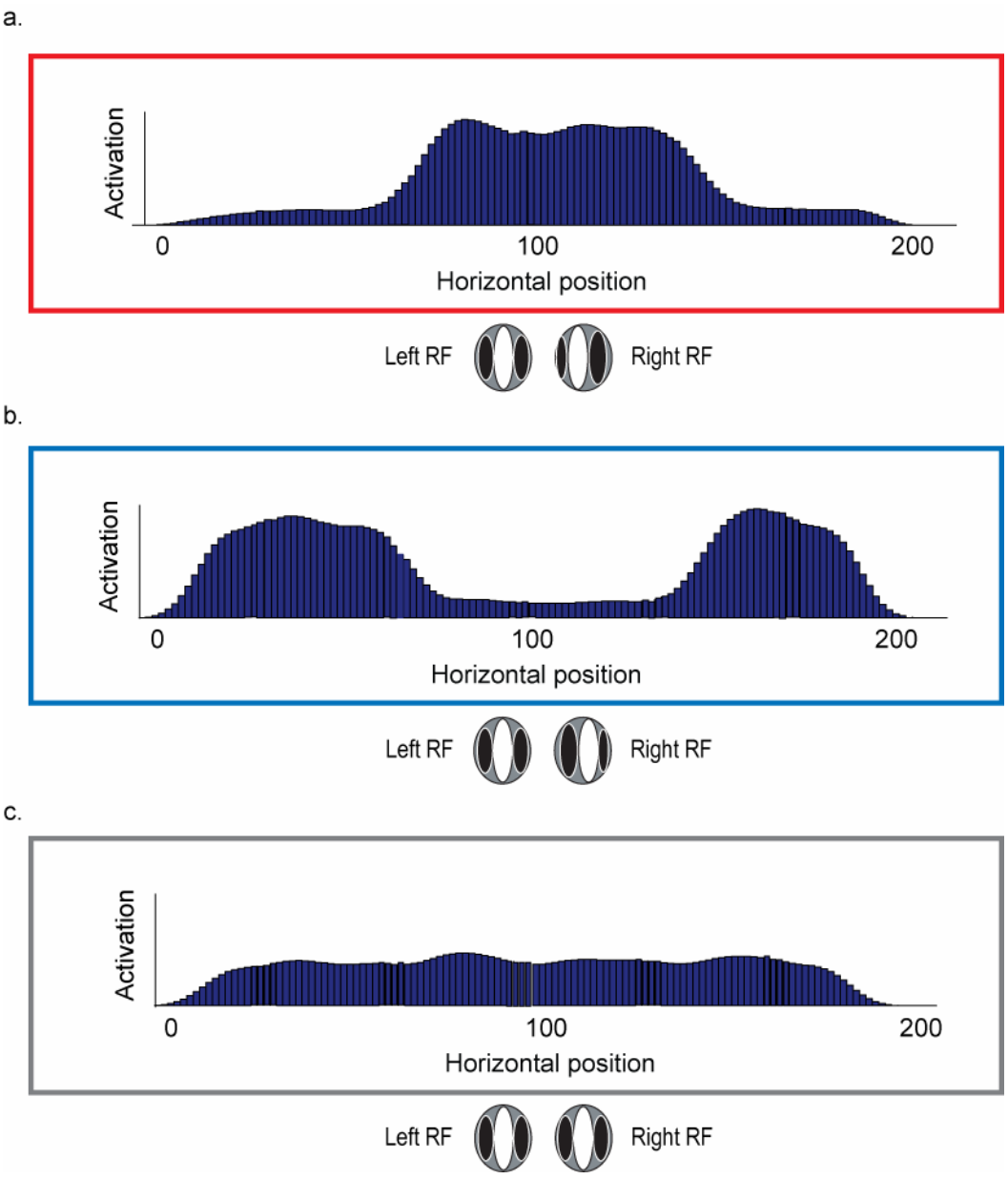


Figure 4.18. The response across the RDS by ODF units tuned to three different disparities. The ellipse pair at the bottom of the each histogram denotes the interocular RF phase shifts (detector disparity). 200 ODF units of the denoted detector disparity are positioned to cover each horizontal location. The activation at the given location is cumulated over 100 trials. The red, blue and gray rectangle match the color-code in the figure 4.19 from which the histograms are sampled (RF frequency: 0.036 cycle/pixel, detector disparity: a. -5 pixels, b. 5 pixels and c. 0 pixel).

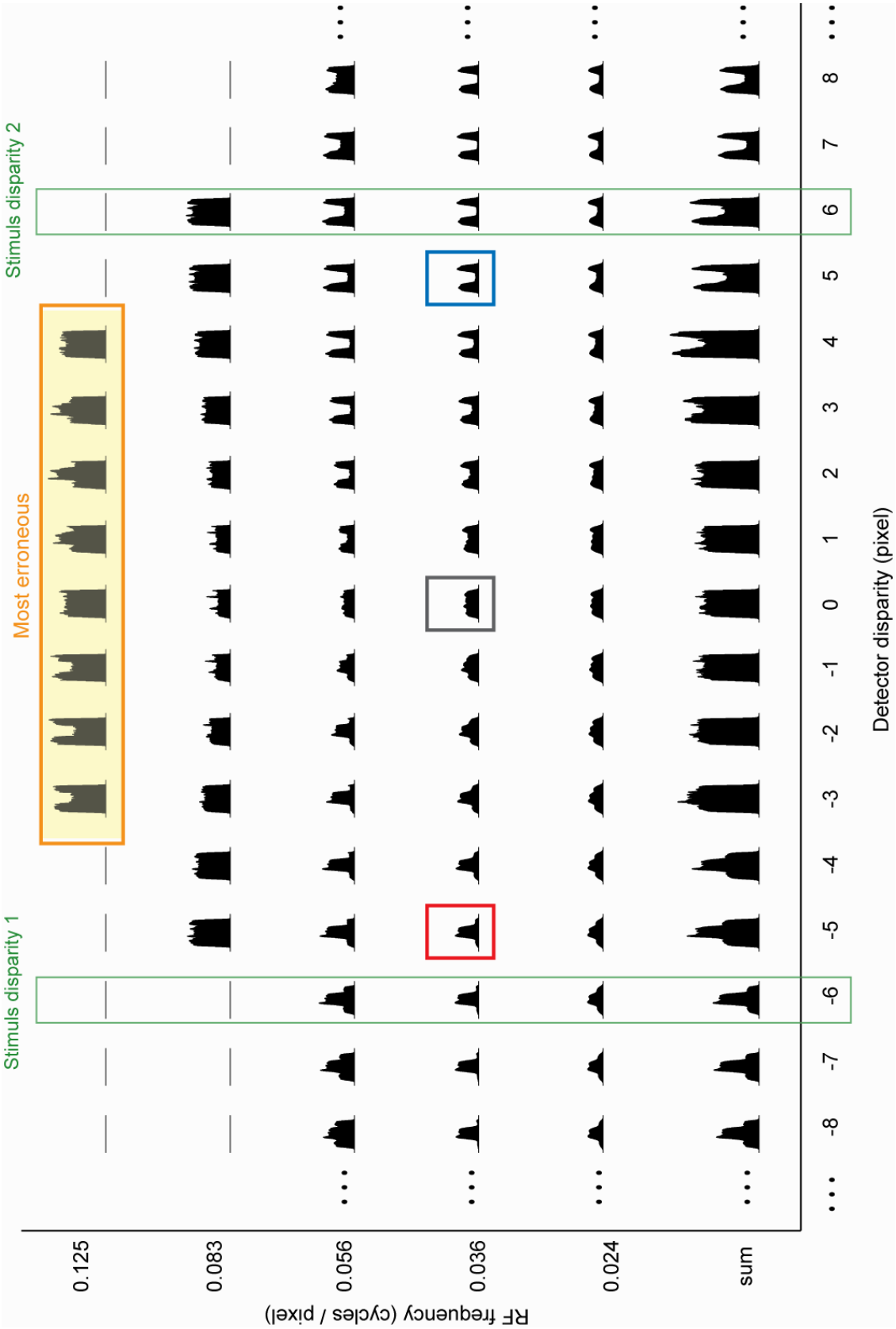

Stimuls disparity 1
Most erroneous
Stimuls disparity 2
0.125
0.083
0.056
0.036
0.024
sum
RF frequency (cycles / pixel)
-8
-7
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6
7
8
Detector disparity (pixel)

Figure 4.19. Activation histograms of ODF unit array. Each histogram shows the amplitude of activation at the given horizontal location (x-axis of each histogram, see figure 4.18 for the convention) in response to the RDS shown in figure 4.16.a. The RF parameters of each histogram is indicated by x-axis (detector disparity) and y-axis (RF frequency). The third and the fourth row closely describe the stimulus disparities whereas the first row shows the most erroneous activation. The histograms in the bottom row show the sum along the columns, the sum across scale for the unit activations at the equal detector disparity. Histograms of simple summation provide deteriorated description of the stimulus depths. The detector disparity in actual simulation ranged from -20 to 21 pixels. The histograms in the red, gray and blue rectangles are magnified in figure 4.18.

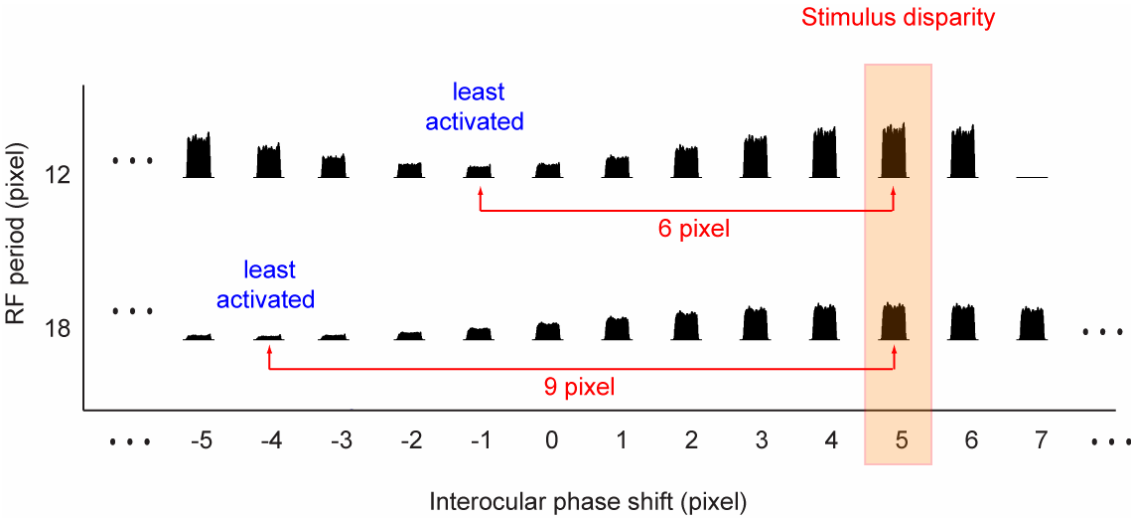


Figure 4.20. The activation histograms of ODF model units at different interocular phase shifts for two different RF frequencies. The minimum activation occurs at a half period away from the maximum activation.

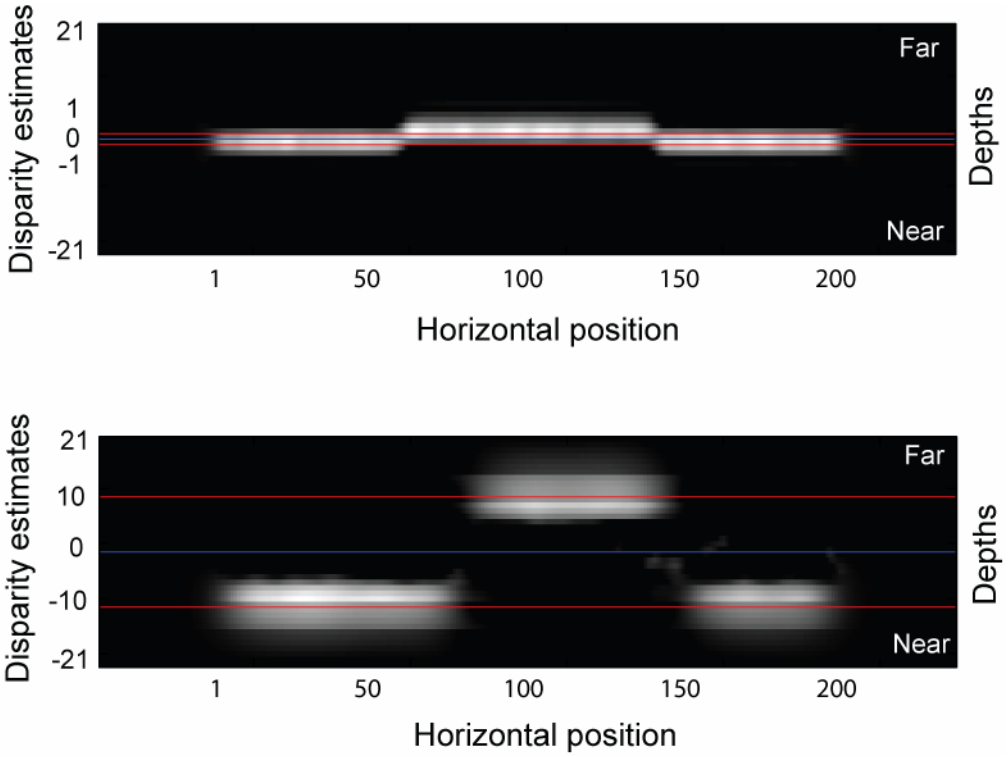


Figure 4.21. Disparity maps of the RDSs. Our model works with arbitrary disparities within the range of largest RF period. The upper disparity map is produced with -1 and +1 stimulus disparities and the bottom disparity map is produced with -10 and +10 stimulus disparities (see figure 4.16a). red line: stimulus disparity.

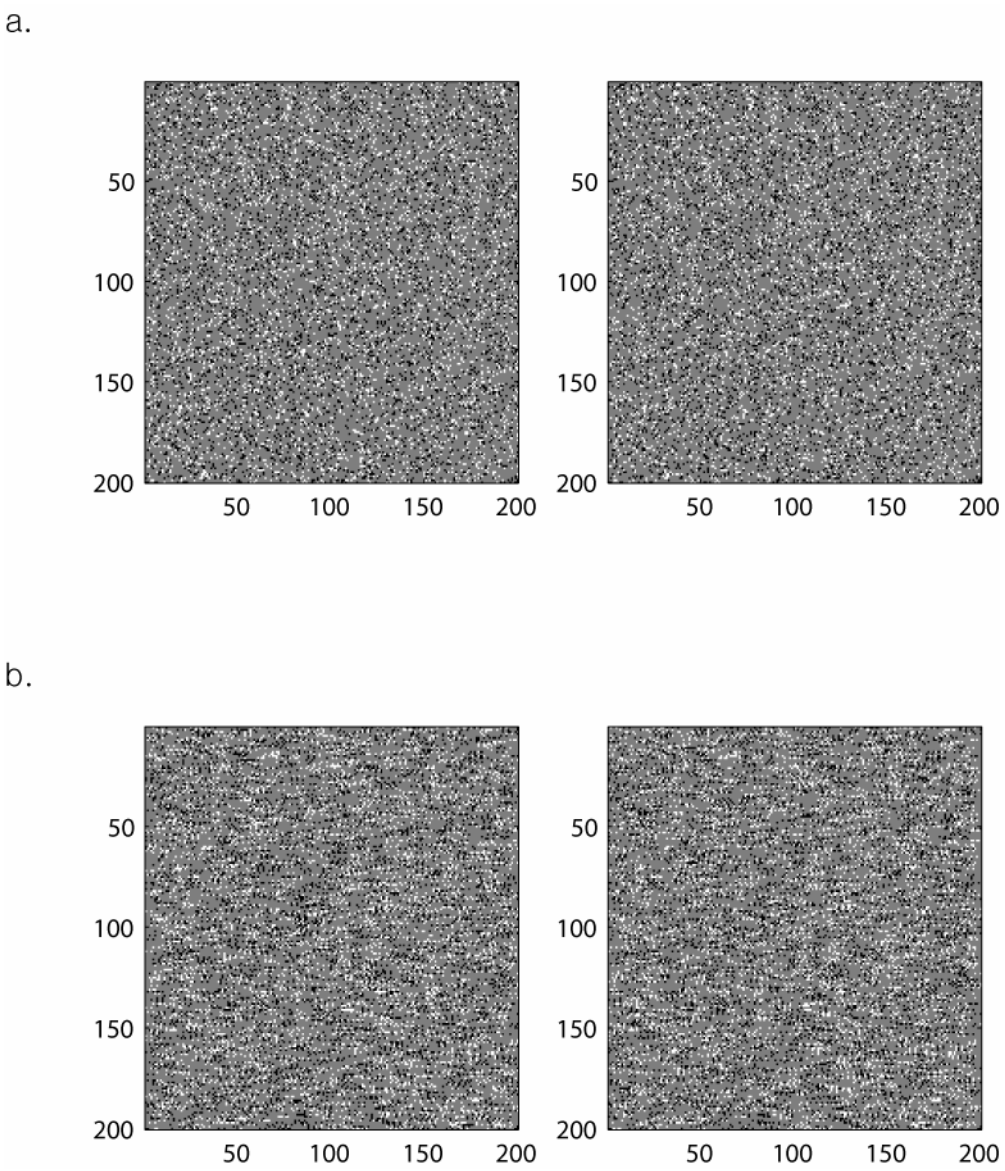


Figure 4.22. RDSs used in the model simulation. a. A 200 x 200 RDS composed of two layers of random-dots at two different depths. b. A 200 x 200 RDS that contains PLC-defined transparent surfaces.

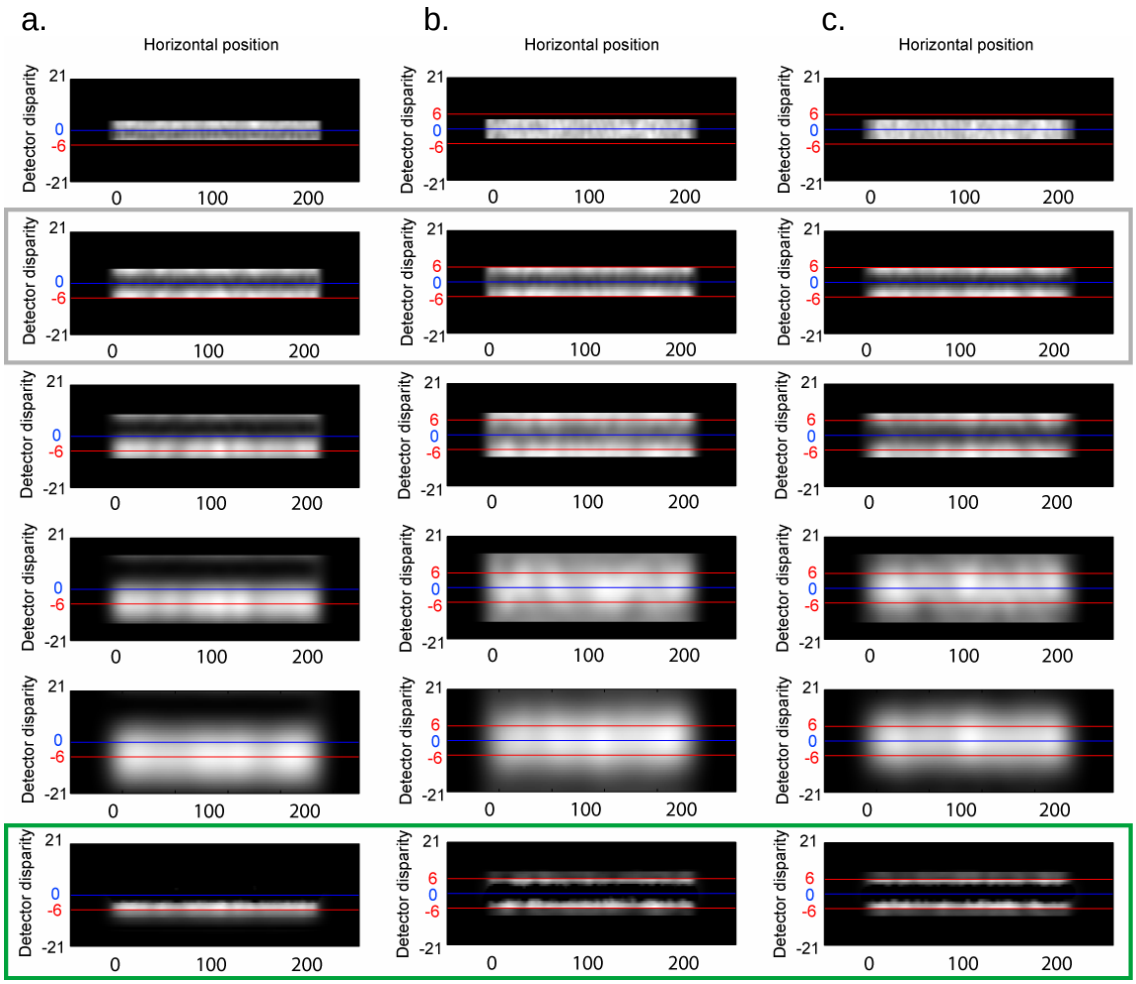


Figure 4.23. Disparity maps of the transparent RDSs. RF period: 8, 12, 18, 28, and 42 pixels from top to bottom row (red line: stimulus disparity). Response of our improved model at the bottom row in the green rectangle. a. Disparity map of non-transparent surface for comparison. b. Disparity maps of double-layered RDS. c. Disparity maps of PLC RDS.

The disparity maps of double-layered RDS and PLC RDS of identical stimulus disparities are very similar.

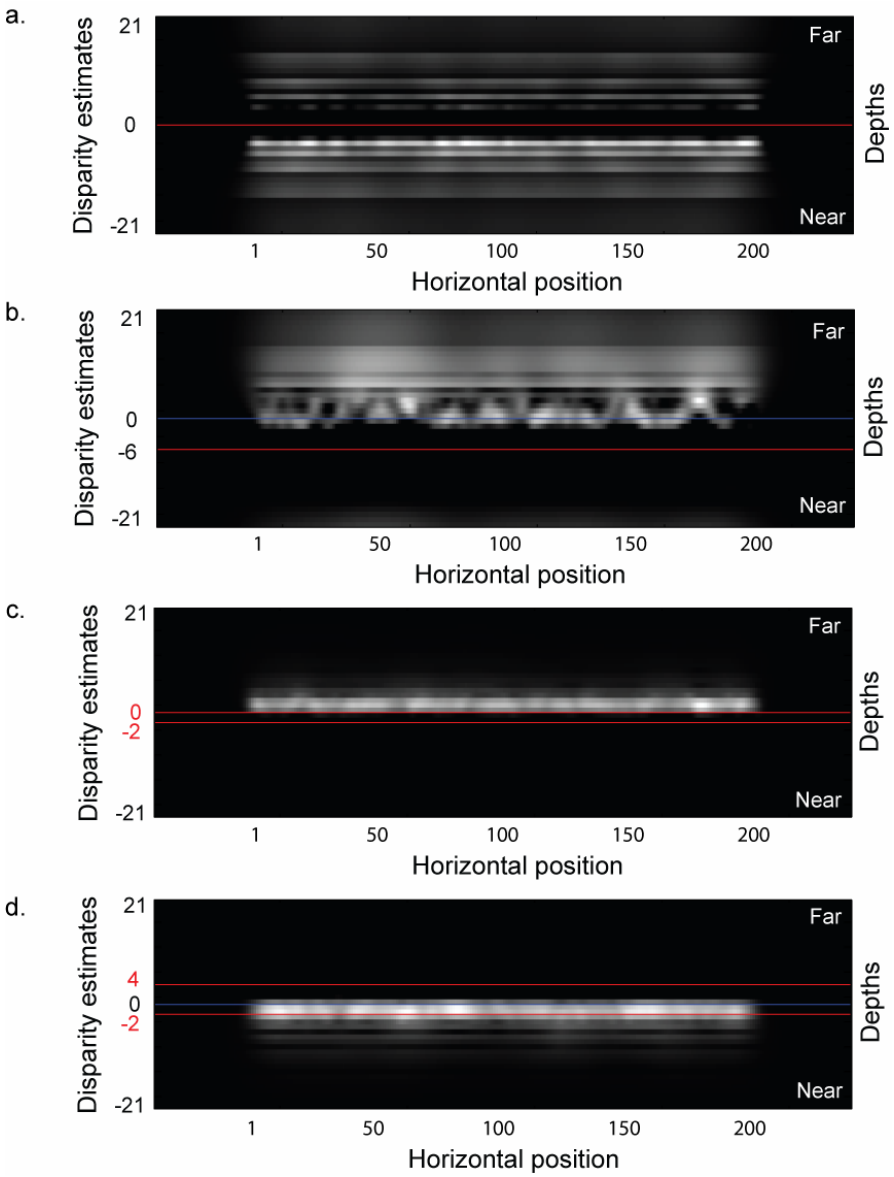


Figure 4.24. Disparity maps of anti-correlated RDSs (red line: stimulus separation). a and b. non-transparent RDS. c and d. transparent PLC RDS. a. The stimulus disparity was 0 pixel. The response spreads toward both directions. b. The stimulus disparity was -6 pixels. The response is on the other side of the stimulus disparity and spreads over wide range. c. The PLC separation was 2 pixels. At a small separation, the disparity estimates appeared to be a single layer whose depth is opposite to the stimulus geometry. b. The PLC separation was 6 pixels. At a greater separation, the disparity estimates of the anti-correlated layer show some degree of spreading and its depth was inverted.

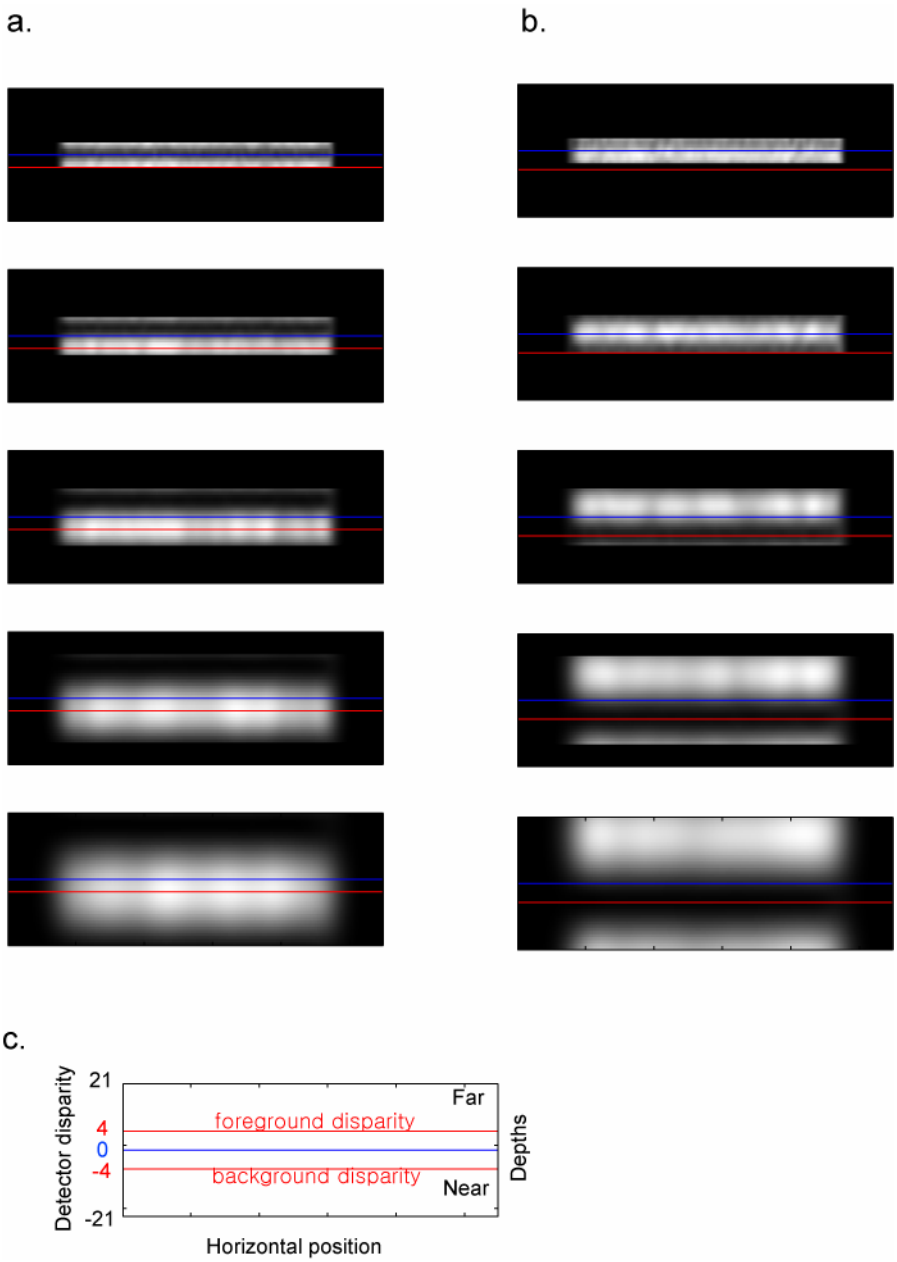


Figure 4.25. Wrap-around of disparity estimates in the phase-disparity detection mechanism prior to applying π-inhibition. Stimulus disparity is 4 pixels in both cases. The conventions and the axis are described in c. a. The raw disparity map of single layer correlated RDS. b. The raw disparity map of single layer anti-correlated RDS. The reduced activations that peak on the other side of the disparity estimates are due to the roll-over. The roll-over occurs because the phase-disparity detector does not distinguish $\pi+\alpha$ and $-\pi+\alpha$.

CHAPTER 5

GENERAL DISCUSSION

5.1 Biological Version of the Correspondence Problem

The uniqueness constraint is a popular strategy of sorting out the true matches among false matches in stereo algorithms in computer vision. In this thesis, we describe evidence that our visual system does not employ the uniqueness constraint in computing the disparity of surfaces. We further showed that even opposite contrast matches give rise to depth perception in some cases, which is an absolute violation of the most fundamental rule applied to binocular matching algorithms. Here we reconsider the biological version of the so-called correspondence problem in stereo vision, the problem involved in finding local image structure in the two retinal images that originate from the same object in the world.

To detect the binocular disparity or to identify the correlated/uncorrelated features in the two eyes, establishing the proper matches between the image features in the two eyes is the most fundamental problem in stereo vision. In the case of random-dot stereograms, the correspondence problem is often stated as identifying which dot in the left image matches which dot in the right image. Because every dot is identical, any two dots – one from each eye – can be potentially matched, and the visual system faces the enormously difficult problem of sorting out the true matches from the huge number of false matches. This is not the case; however, in a scheme that incorporates matching

image patches, for example, in the binocular energy model. The distribution of the random-dots in an image patch is more likely unique. The task of matching becomes conceptually simple, computing the cross-correlation between the two patches at different locations (or phases within the same patch) and identifying the patches of maximum correlation.

The double-matching in Panum's limiting case (PLC) is not an issue anymore in patch-wise matching, whereas the anti-correlated matches are because it produces inversed cross-correlation. Although occurrence of the anti-correlated retinal images (binocular matches of opposite contrast) is most improbable in natural scene, the existence of the neuronal signal related to the anti-correlated matches and the depth reversal in our novel RDS is consistent with patch-wise matching but not with point matching. Therefore, the biological version of the correspondence problem requires identification of new characteristics of the relevant visual signals from the irrelevant ones. We think the different patterns recognized in response to the correlated and anti-correlated stimulus provide one of the clues, thus the coherence of the disparity estimates across scale is crucial in stereo matching.

5.2 Biologically Plausible Process of Combining Elementary Disparity Signals

It is well known that the elementary signals found in the early cortical areas (V1) cannot account for our conscious depth perception and suggested that stereopsis requires additional processing beyond the initial disparity detection.

Incorporating the binocular energy model which well-describes the physiological characteristics of some disparity detectors in the V1, we have proposed a post disparity

detection mechanism that combines the signals from the individual disparity detectors. We employ a biologically plausible process of simple mutual inhibition between disparity detectors. The inhibitory process eliminates the monocular response from the disparity detectors and enhances the disparity-tuning. The enhanced signal is then summed across scale to increase signal to noise ratio and to cover arbitrary disparity range.

This second stage process provides a disparity map whose depth description is closer to our perceptual experience in RDS. The most significant progress includes the following three observations. First, the disparity estimates are much narrower than those of the original ODF model; therefore, the model distinguishes the sharp transition of the depth boundaries. Second, the narrow disparity estimates also lead to discrimination of the transparent surfaces in both double-layered RDS and PLC RDS. Third, the process is free of wrap-around effect intrinsic to the phase-disparity mechanism and applicable to an arbitrary range of disparities.

There are, yet, several assumptions that require additional improvement or explanations for broader implication. The first is intrinsic to the phase-disparity detection mechanism. This mechanism shows rapid attenuation in the response off the central RF region. For example, a small stimulus at large disparity produces very little activation in the ODF model. The signal is almost negligible compared to another stimulus at the center of the RF (see figure 4.24d for attenuation). The pi-inhibition improves this issue but encompassing the position-disparity mechanism, most probably a hybrid of the two models, fundamentally corrects this problem. The second is the application or extension to natural images. We used several simplifications to focus on the problem of simple contrast matching, that is correlated and anti-correlated matches. For this reason, we used

the luminance values of only -1, 0 and 1. When extended, to work with wider range of luminance, preprocessing of the image is required; however, this has an issue for degrading the unique signatures in patch-matching. Another simplification we made was employing 1-D network, which limited the scope of the simulation by excluding vertical disparity and orientation.

Besides those limitations, the core merit of the network is the proposition of a biologically plausible way of combining the elementary disparity signals, whose response is closer to our depth experience. This process is simple but robust enough so that it works independent of the front end mechanisms as long as it provides physiologically probable disparity signals.

LIST OF REFERENCES

Anzai A, Ohzawa I, Freeman RD (1999) Neural mechanisms for processing binocular information II. Complex cells. Journal of Neurophysiology 82:909-924

Anzai A, Ohzawa I, Freeman RD (1999) Neural mechanisms for processing binocular information I. Simple cells. Journal of Neurophysiology 82:891-908

Barlow HB, Blakemore C & Pettigrew JD (1967) The neural mechanism of binocular depth discrimination. Journal of Physiology 193: 327-342

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

Burt P, Julesz B (1980) A disparity gradient limit for binocular fusion. Science 208:615-617

Campbell FW & Robson J (1968) Application of Fourier analysis to the visibility of gratings. Journal of Physiology 197: 551-566

Cumming BG & Parker AJ (1997) Responses of primary visual cortical neurons to binocular disparity without depth perception. Nature 389: 280-283

Cumming BG, Shapiro SE & Parker AJ (1998) Disparity detection in anticorrelated stereograms. Perception 27: 1367-1377

DeValois R L, Albrecht DG & Thorell LG (1982) Spatial frequency selectivity of cells in macaque visual cortex. Vision Research 22:5 45–559

Fender D & Julesz B (1967) Extension of Panum's fusional area in binocularly stabilized vision. Journal of the Optical Society of America 57: 819-830

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

Frisby JP (2001) Limited Understanding of Panum's Limiting Case. Perception 30: 1151-1152

Gettys C F & Harker GS (1967) Some observations and easurements of the panum phenomenon. Perception and Psycho-physics 2: 87–395

Gillam B, Blackburn S, Cook M (1995) Panum's Limiting Case: Double Fusion, Convergence Error, or 'da vinci stereopsis‘. Perception 4: 333-346

Gillam B & Borsting E (1988) The role of monocular regions in stereoscopic displays. Perception 17: 603-608

Gillam B, Cook M, Blackburn S (2003) Monocular discs in the occlusion zones of binocular surfaces do not have quantitative depth - a comparison with Panum's limiting case. Perception 32: 1009-1019

Graham N & Nachmias J (1971) Detection of gratings patterns containing two spatial frequencies: a comparison of single-channel and multiple channel models. Vision Research 11: 251-259

Hampton DR & Kertesz AE (1983) The extent of Panum's area and the human cortical magnification factor. Perception 12: 161-165

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

Howard IP & Rogers BJ (2002) Seeing in Depth (Toronto: I Porteous)

Julesz B (1960) Binocular depth perception of computer generated patterns. Bell System Technical Journal 39: 1125-1162

Kaufman L & Pitblado C (1965) Further observations on the nature of effective binocular disparities. Am J Psychol. 78: 379-391

Krol JD, Grind WA van de (1980) The double-nail illusion: experiments on binocular vision with nails, needles, and pins. Perception 9: 651-669

Krug K, Cumming BG & Parker AJ (2004) Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. J Neurophysiol. 92: 1586-1596

Liu L, Stevenson SB & Schor CM (1994) Quantitative stereoscopic depth without binocular correspondence. Nature 367: 66-69

Mallot HA, Gillner S & Arndt PA (1996) Is correspondence search in human stereo vision a coarse-to-fine process? Biological Cybernetics 74: 95–106

Marr D (1982) Vision: A Computational Investigation into the Human Representation and Processing of Information. (San Francisco, Freeman)

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

Masson GS, Busettini C & Miles FA (1997) Vergence eye movements in response to binocular disparity without depth perception. Nature 389: 283-286

McKee SP, Bravo MJ, & Smallman HS (1995) The ‘Uniqueness Constraint' and Binocular Masking. Perception 24: 49-65

Mikaelian S & Qian N (2000) A physiologically-based explanation of disparity attraction and repulsion. Vision Research 40: 2999–3016

Mitchell DE (1966) Retinal disparity and diplopia. Vision Research 6: 441-451

Nakamizo S, Shimono K, Kondo M & Ono H (1994) Visual directions of two stimuli in Panum's limiting case. Perception 23:1037-1048

Nakayama K & Shimojo S (1990) da Vinci stereopsis: depth and subjective occluding contours from unpaired image points. Vision Research 30: 1811-1825

Ogle K N (1952) On the limits of stereoscopic vision. Journal of Experimental Psychology 44: 253-259

Ogle K N (1962) The optical space scene. in *The Eye* volume 4, Ed. H Davson (New York: Academic Press) 374-375

Ohzawa I, DeAngelis GC & Freeman RD (1990) Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. Science 249: 1037-1041

Ono H, Shimono K, Shibuta K (1992) Occlusion as a depth cue in the Wheatstone-Panum limiting case. Percept Psychophys 51:3-13

Palmer DA (1961) Measurement of the horizontal extent of Panum's area by a method of constant stimuli. Optica Acta 8: 151-169

Parker AJ & Yang Y (1989) Spatial properties of disparity pooling in human stereo vision. Vision Research 29:1525-1538

Poggio GF (1995) Mechanisms of stereopsis in monkey visual cortex. Cerebral Cortex 5: 193-204

Poggio GF & Fisher B (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. Journal of Neurophysiology 40: 1392-1405

Poggio GF, Motter BC, Squatrito S & Trotter Y (1985) Responses of neurons in visual cortex (V1 and V2) of the alert macaque to dynamic random-dot stereograms. Vision Research 25: 397-406

Poggio GF & Poggio T (1984) The analysis of stereopsis. Annual Review of Neuroscience 7: 379-412

Prince SJ, Cumming BG & Parker AJ (2002) Range and mechanism of encoding of horizontal disparity in macaque V1. J Neurophysiol. 87: 209-221

Read JCA and Eagle RA (2000) Reversed Stereo Depth and Motion Direction with Anti-Correlated Stimuli. Vision Research 40: 3345-3358

Richards W & Foley JM (1971) Interhemispheric processing of binocular disparity. Journal of the Optical Society of America 61: 419–421.

Qian N (1994) Computing Stereo Disparity and Motion with Known Binocular Cell Properties. Neural Computation 6: 390-404

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

Sanger TD (1988) Stereo disparity computation using Gabor filters. Biological Cybernetics 59: 405-418

Shapley R & Lennie P (1985) Spatial frequency analysis in the visual system. Annual Review of Neuroscience 8: 547-583

Shimojo S & Nakayama K (1990) Real world occlusion constraints and binocular rivalry. Vision Research 30: 69-80

Shimono K, Tam WJ & Nakamizo S (1999) Wheatstone-Panum limiting case: occlusion, camouflage, and vergence-induced disparity cues. Percept Psychophys 61: 445-455

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

Tsai JJ & Victor JD (2003) Reading a population code: a multi-scale neural model for representing binocular disparity. Vision Research 43: 445-466

Wang Z, Wu X, Ni R & Wang Y (2001) Double fusion Does Not Occur in Panum's Limiting Case: Evidence from Orientation Disparity. Perception 30: 1143-1149

Westheimer G (1986) Panum's phenomenon and the confluence of signals from the two eyes in stereoscopy. Proc R Soc Lond B Biol Sci. 228: 289-305

Westheimer G & Tanzman IJ (1956) Qualitative depth localization with diplopic images. Journal of the Optical Society of America 46: 116–117

Wilson HR, Blake R & Halpern DL (1991) Coarse spatial scales constrain the range of binocular fusion on fine scales. Journal of the Optical Society of America A 8: 229-236

Zhu YD & Qian N (1996) Binocular receptive field models, disparity tuning, and characteristic disparity. Neural Comput. 8: 1611-1641

APPENDIX: SEEING STEREOCOPIC DEPTH

A.1 Binocular Depth Cues

There are two types of depth cue that play important roles in stereoscopic depth perception. Here we explain some terminology used in stereoscopic depth perception or stereopsis.

A.1.1 Retinal Disparity (Binocular Disparity)

Due to the lateral separation of the two eyes, objects at different distances from the eyes project images on to the two retinae that differ in their horizontal positions. The difference between the two retinal positions is called retinal disparity or binocular disparity. Binocular disparity encodes the third dimension into a pair of 2-D retinal images. Binocular disparity provides two crucial piece of information about depths; the depth direction and the depth magnitude. figure A.1 and figure A.2 illustrate the relationship between the points in space and the corresponding projections onto the two retinae.

Our eyes converge when we look at nearby objects and diverge when we look at distant objects. The convergence and the divergence of the eyes bring the object of interest as close as possible to the center of each retina where the resolution of vision is maximized. In figure A.1, the eyes rotate toward each other to point to a dot (F) in space. The image of dot F falls on to the center of the two retinae. The red disk (U) behind the

fixation dot (F) projects its image on to the nasal side, and the blue triangle (C) in front of the fixation dot (F) projects its image onto the temporal side. From the retinal point of view, the direction of disparity provides information about which points are nearer and which are farther than the fixation point in space. Conventionally, positive values represent uncrossed disparity and negative values represent crossed disparity.

Figure .2 illustrates that the amplitude of the binocular disparity provide information about how much nearer or farther the points are relative to the fixation point in space. In figure A.2, the projections of F on the two retinae align completely and therefore have zero disparity. The blue square (U_1) behind the fixation point F projects its image on to inward direction (uncrossed disparity). The red square (U_2) which is farther than the blue square (U_1) from the viewer projects its image on to the retinal location that is further away than projection U_1' from projection F'. Thus, the greater the difference in depth between the fixation point and the object is the greater the disparity of the object on the retinae.

A.1.2 Monocular Occlusion

When near objects partially occlude the far objects, the occlusion is different in the two eyes because of the different views of the two eyes. For example, when an opaque sheet is placed in front of a wall, part of the wall can be seen only by one eye but not by the other. Figure A.3 illustrates that the monocular region M_L in the scene is visible only to the left eye but not to the right eye. In the retinal projections, these monocular zones are always situated on the nasal side of the occluding binocular zone (B_a') and on the temporal side of the occluded binocular zone (B_b'). In the retinal

projections, the differential occlusion provides important binocular information about the depth relationship between the occluding and occluded surfaces. Furthermore, because the width of the occluded zone increases as the relative depth between the two surfaces increases, differential occlusion also provides information about the relative depth between the surfaces.

There is evidence that the monocular occlusion facilitates depth perception by reducing the time required to recognize depth discontinuity (Gillam and Borsting, 1988) and enhances impression of the depth edges (Shimojo and Nakayama, 1990). Beside those complementary roles, there is yet no conclusive evidence that such monocular region itself give rise to depth sensation in the absence of any type of binocular disparity (Liu, Stevenson and Schor, 1994; Gillam, Blackburn and Cook, 1995).

The features in the binocular zone have matching images in the two eyes that provides unambiguous depth once the proper matches are established. On the other hand, features in the monocular region provide depth information only about whether the monocular feature is farther or closer than the occluding surface. This is because the projections form the monocular zone on to the retina can originate from any point in space along the visual direction of the viewing eye (figure A.4).

In figure A.3, the monocular region M_L belongs to the farther surface so the depth of M_L can be referenced from the ownership, however, this is not always the case. In figure A.5, the monocular region M_2 is another surface on the left side of occluding object. Determination of the ownership in this case requires additional information for comparison, for example, texture, luminance contrast, shading, or blur.

A.2 Exercise of Free-fusion With Stereograms

Three dimensional depth perception can be stimulated with a pair of images called a stereogram. A stereogram is a pair of flat, 2-d images that give rise to a depth sensation when presented one for each eye. The stereogram consists of a pair of photographic images taken at different view points as if from the two eyes, or a pair of artificially designed images containing features whose presence and positions are manipulated to give rise to depth sensation. Figure A.6 is a simple example of a later case, in which one bar has a crossed disparity and the other has an uncrossed disparity relative to the rectangular frame. When free-fused, the bars appear at different depths. Free-fusion is a way of seeing stereoscopic depth in stereogram with bare eyes.

For free-fusion, one intentionally looks nearer or farther away than the depth of the plane where two images lie. Figure A.7 illustrates how free-fusion works. The red and blue dots correspond to the two dots above the image pair in figure A.6. Figure A.7a show a normal viewing situation. The eyes are fixating on the depth of the plane where the two images lie on. When the two retinae are superimposed, red and blue dots at corresponding positions completely overlap. Figure A.7b illustrates the situation of free-fusion. By looking farther than the depth of the images, the eyes fixate on an imaginary point behind the plane of the two images. Now, the point 2 in the left retina corresponds to the point 1 in the right retina while the point 1 in the left retina and point 2 in the right eye become unmatched. Because of these two unmatched dots, the observer sees three dots instead of seeing two dots as in normal case - one binocularly matched on corresponding position and two unmatched dots in the flanking positions. The unmatched

images are the side-effects of the free-fusion and one should focus on the central matched images to see the depths.

A.3 Depth Cues That Are Not Binocular

There are many different sources of depth information other than the stereoscopic information mentioned above. Even with one eye alone, one has a convincing depth experience from the scene. Photographs and motion pictures give rise to compelling sense of depths. This is possible because the visual system implicitly make certain heuristic assumptions about the nature of the visual scene and together with the retinal images derives the most probable and consistent conclusions. For example, because we know something about the size of a person from experience, we can judge a person's distance. Or if similar objects appear different in size, the smaller is assumed to be more distant. Pattern of light and dark that resembles the distribution of illumination and shadows give rise to a strong sensation of depth. Even simple edges provide depth information in some circumstances. In figure A.8, people invariably perceive a rectangle behind a disc. Under normal viewing conditions, there are various pictorial depth cues that do not require stereoscopic vision.

A.4 Random-dot Stereograms (RDS)

Random-dot stereogram (RDS) provides a stimulus that enables one to manipulate stereoscopic factors of depth perception devoid of other pictorial cues. An RDS is a pair of images containing thousands of randomly positioned dots whose positions are systemically shifted. The shifts follow the geometry of the binocular

disparity. When binocularly combined, one sees depth structures that are not visible with one eye closed. Figure 1.2 is an example a typical RDS first introduced into vision science by Julesz (1960). By free-fusion, one sees a square floating in the center. This type of depth perception is purely stereoscopic because one cannot detect the square if either eye is closed.

Retrieving depth from RDS accentuates a problem in stereoscopic vision, the so-called correspondence problem. Extraction of disparity information from the two images requires the proper matches to be identified between two images. Under normal viewing conditions, the features in each eye have unique signatures (color, luminance, shape, orientation, etc), so the candidates for the match in the other eye are rather limited. In RDS, however, with no unique features, every dot in one eye becomes a potential candidate of binocular match with every dot in the other eye. Nevertheless, our visual system finds the proper solution in the RDS that leads to a certain and consistent depth perception.

A.5 Da Vinci Stereopsis Stimulus and its Variants

Figure A.9 shows a da Vinci stereopsis stimulus and a variant. The vivid depth sensation observed in typical da Vinci stereopsis stimulus in figure A.9a does not occur in its variant in figure A.9b. The perceived depth would not be disturbed if our visual system uses the configural information to extract depth in this case. In figure A.10a, the white lines are assigned small disparities, an uncrossed disparity in the upper half and a crossed disparity in the lower half of the stereogram. The upper half appears farther away than the lower half when the stereogram is binocularly combined. Another stereogram in

figure A.10b contains identical geometry except that the white line is split into two third of white and one third of the black. The white lines are now vertically aligned and flanked with black thin monocular lines. The positions of the flanking black lines are on the opposite side in relation to the white lines for the upper groups and the lower groups. This is a da Vinci stereopsis configuration. The white lines that do not have obvious binocular disparity appear in two different depths when the two images are binocularly combined.

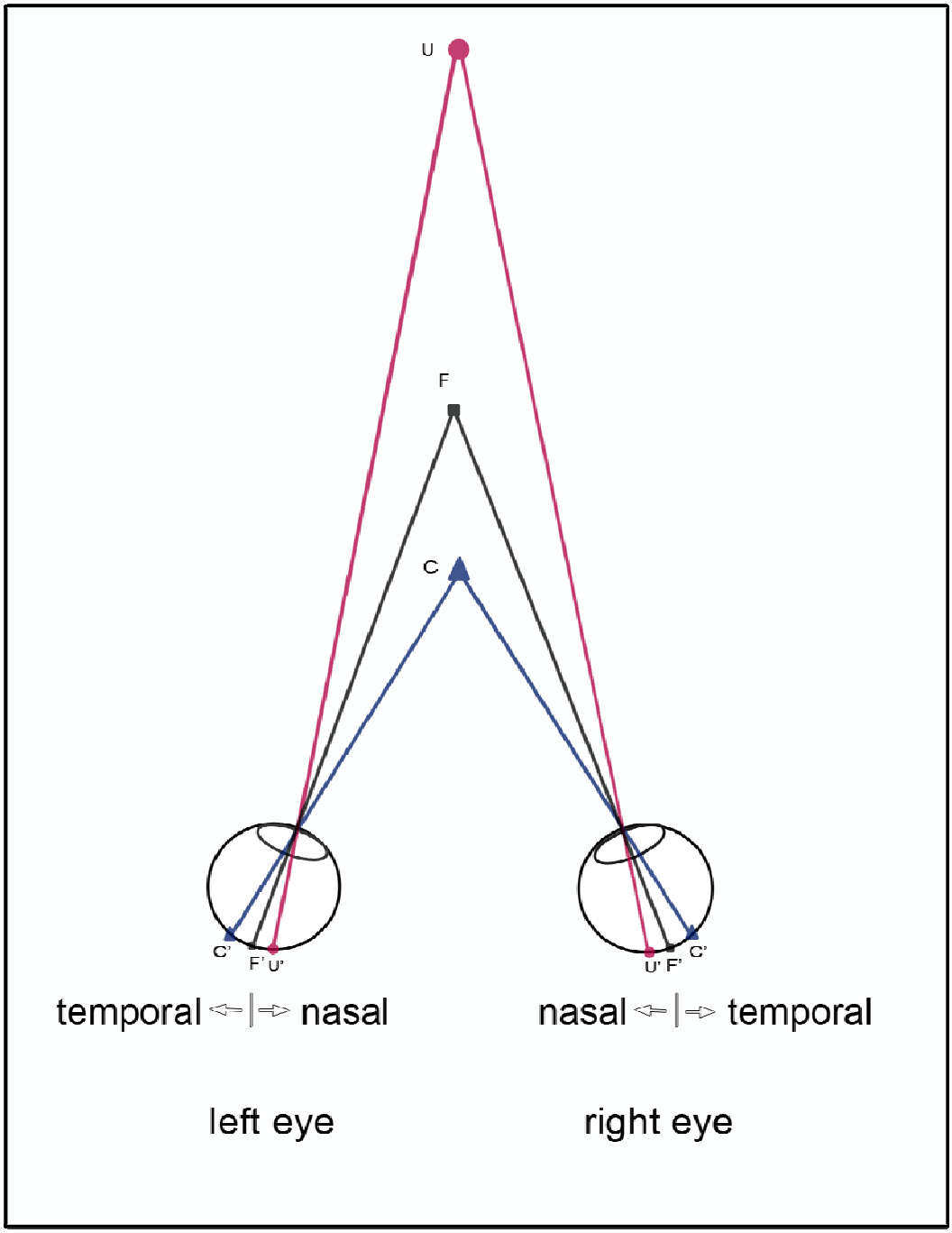


Figure A.1. Relationship between the depth sign (farness/nearness) and the direction of retinal disparity. The image of the object U behind the fixation point F falls on to the nasal side of the retina (U'), whereas the image of the object C in front of point F falls on to the temporal side of the retina (C').

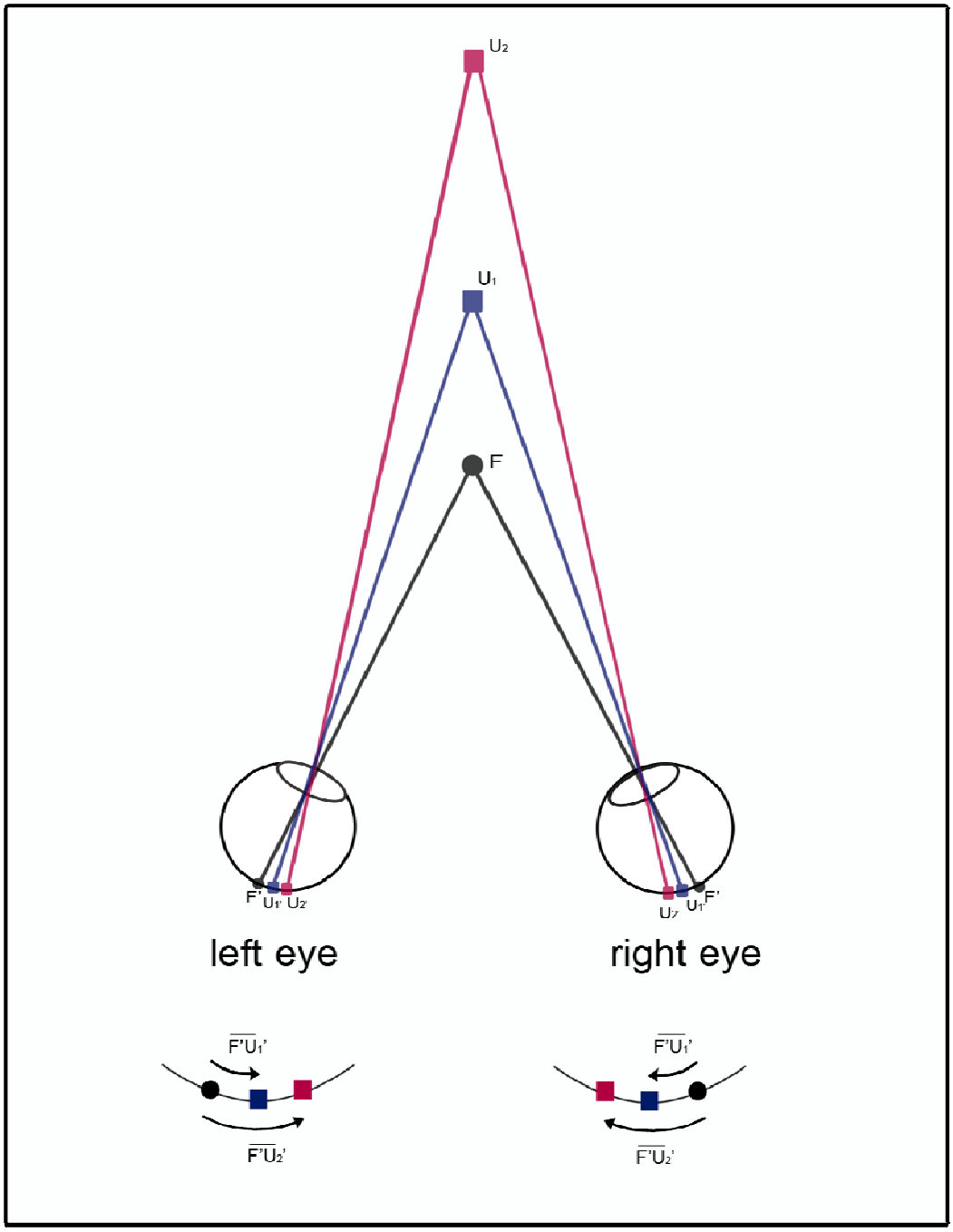


Figure A.2. Relationship between the depth in space and the amplitude of the retinal disparity. The farther or the nearer from the fixation point an object is, the further away from the center of the retina the image of the object falls on to the retinae. The retinal distance of FU_2' is greater than FU_1' because U_2 is farther away form point F in space.

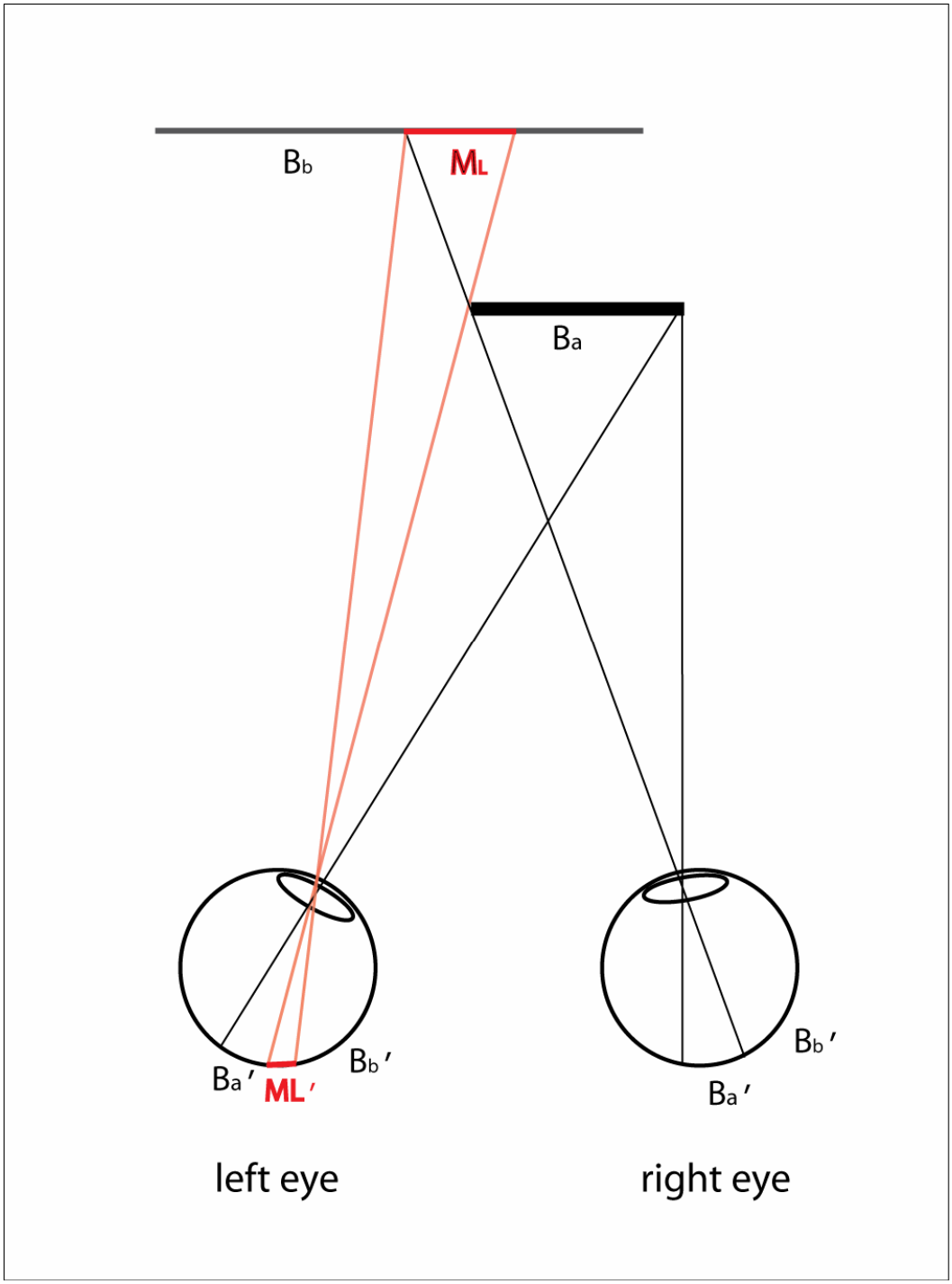


Figure A.3. Monocular occlusion and the retinal projections. The orange lines are the lines of visual direction that encapsulate the monocular region M_L that is visible only to the left eye. The binocular region B_a and B_b are projected on to both eyes.

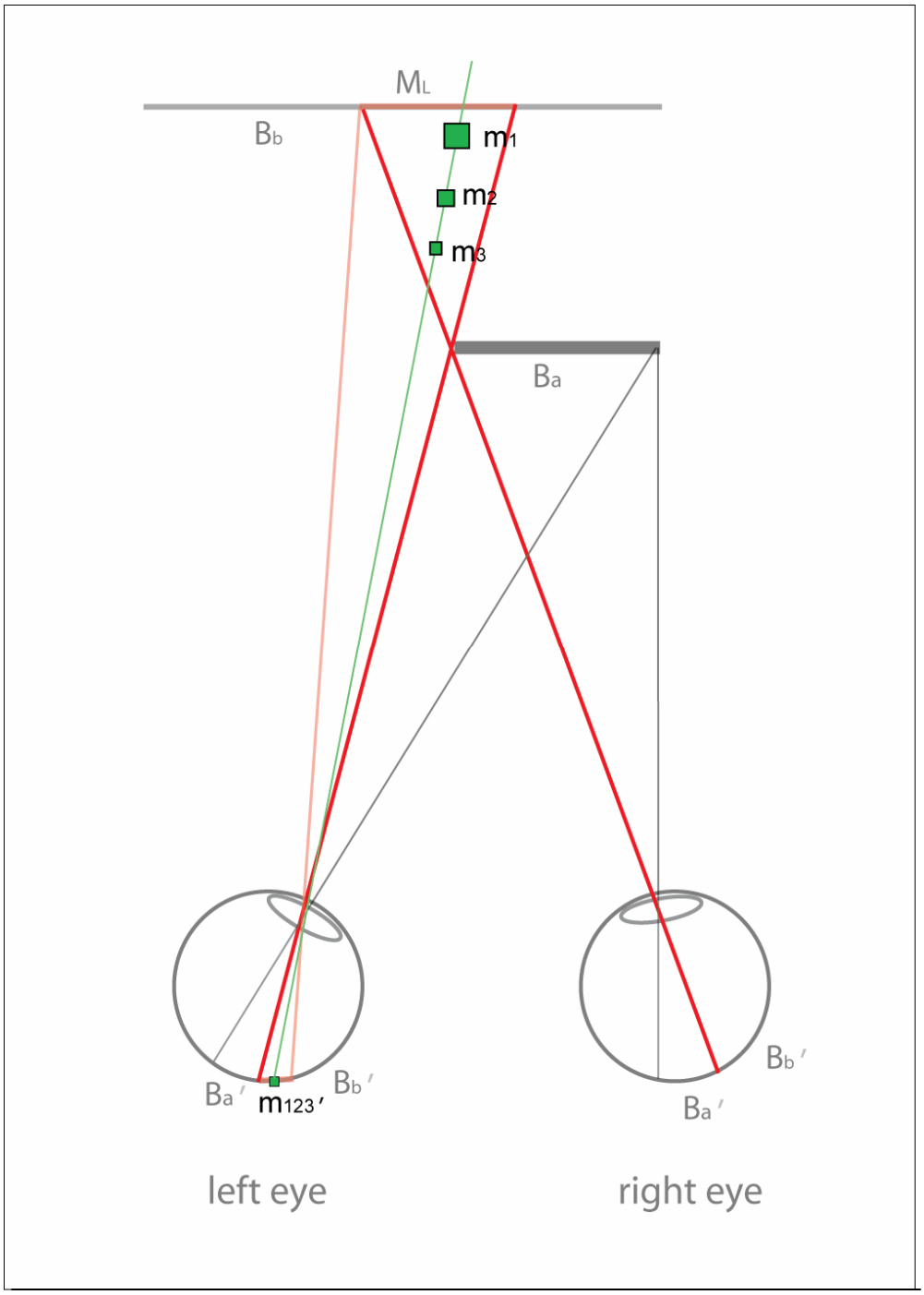


Figure A.4. Ambiguity of the depth magnitude for the monocular features. Three objects (m1, m2, and m3) that are at different depths but within a monocular zone project their images on to the same retinal location (m_{123}'). From the retinal point of view, the depth magnitude of the object is not determinable.

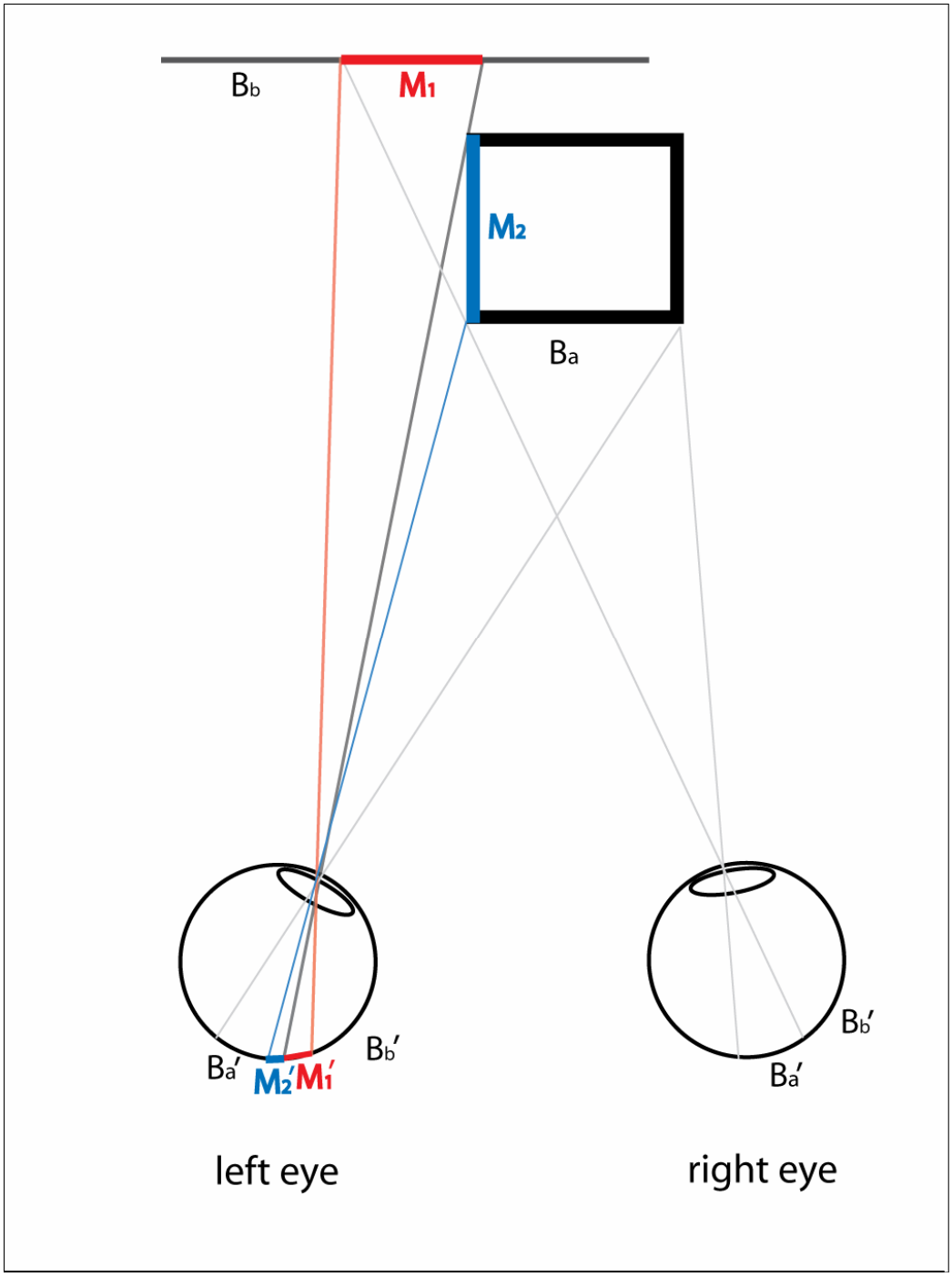


Figure A.5. Ambiguity of the ownership in monocular occlusion. The monocular zones (M_1 and M_2) on the retina do not necessarily belong to the occluded surface (B_b) although they are at a depth farther than the occluding surface (B_a).

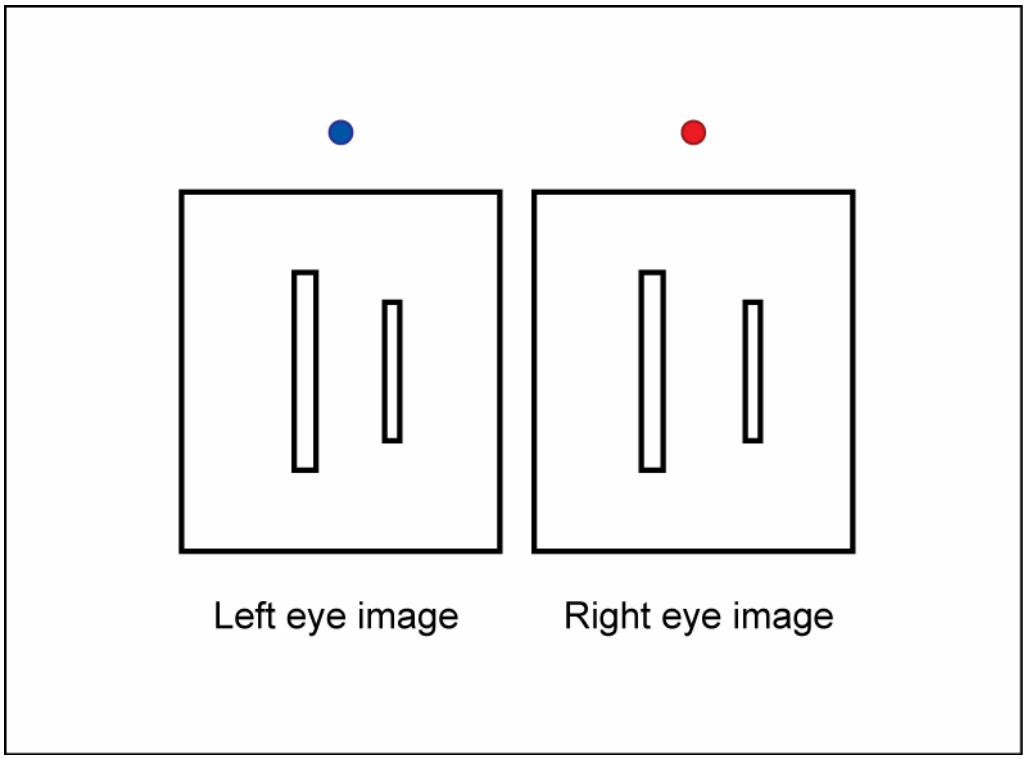


Figure A.6. A simple stereogram. When these two images are binocularly presented and combined, the two bars appear at different depths.

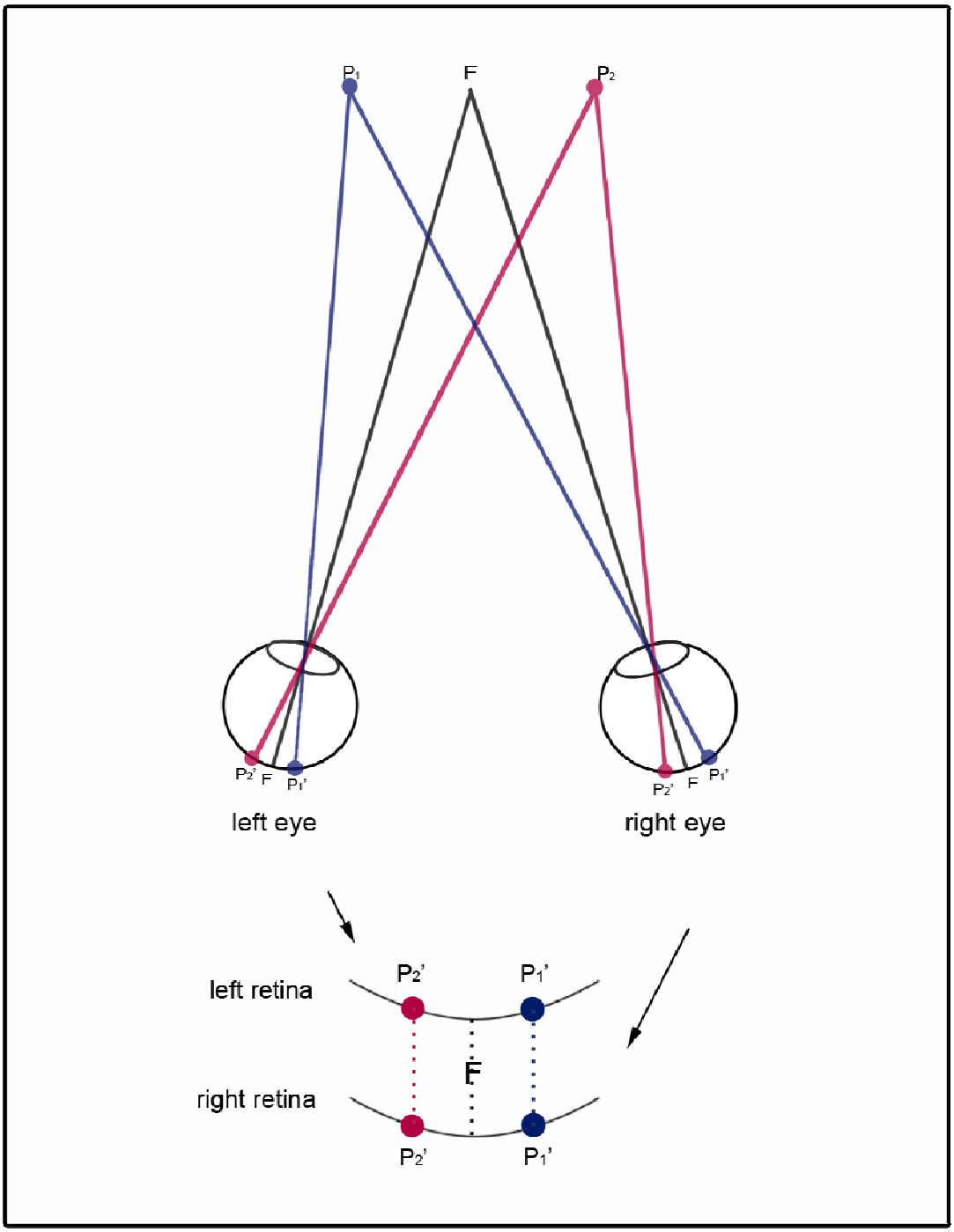


Figure A.7. Exercise of free-fusion. a. Retinal images under normal viewing condition. The retinal images of P_1' and P_2' completely overlap when one fixates at the depth of the two dots.

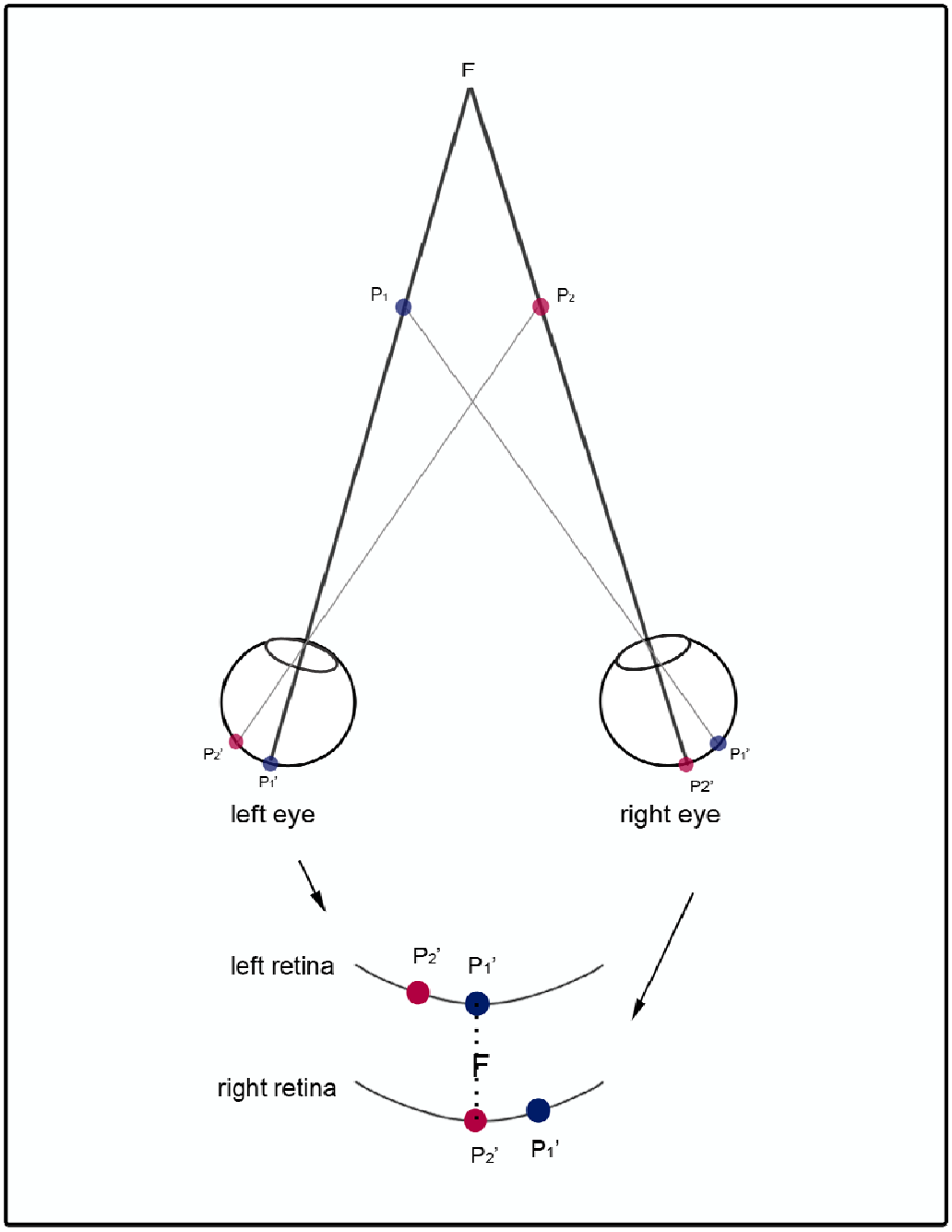


Figure A.7. Exercise of free-fusion. b. Uncrossed free-fusion. When one looks farther than the depth of the plane on which the two dots lie, the image of the blue dot in the left eye overlaps with the image of the red dot in the right eye. The stereogram below the dots in figure 1.7 then becomes fused and depth sensation occurs.

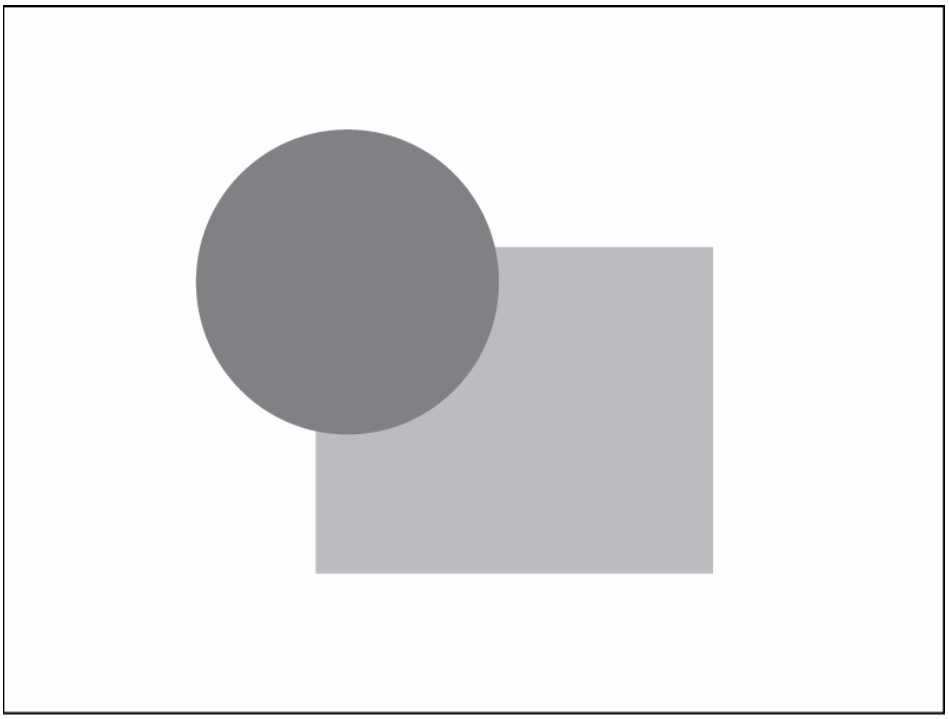


Figure A.8. Occlusion as a depth cue. Certain type of edges, junctions in particular, provides information about the relative depth defined with occlusion. In this case, the disc appears to stand out in front of the rectangle.

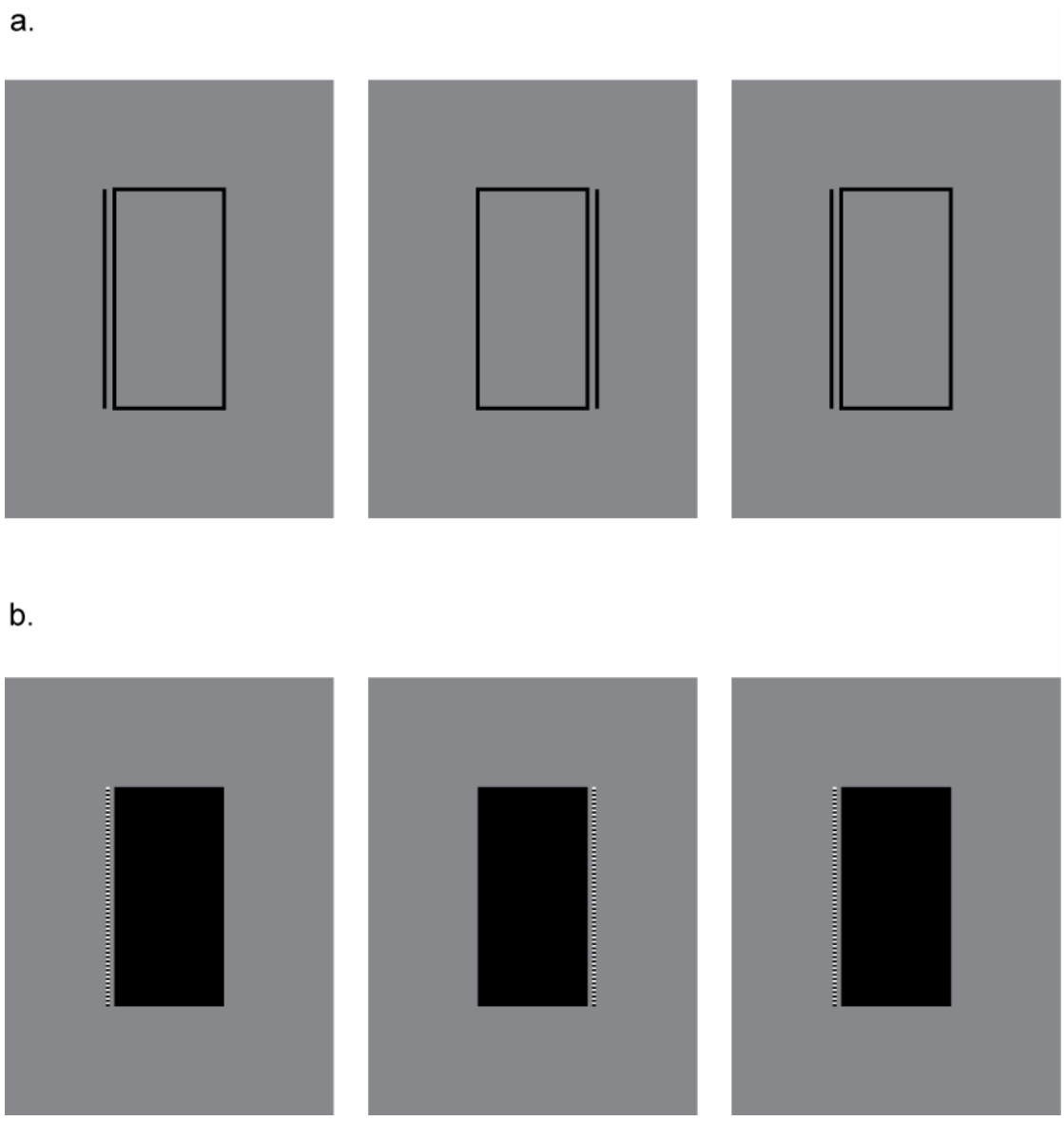


Figure A.9. Da Vinci stereopsis stimuli. a. A typical da Vinci stereopsis stimulus adopted from Nakayama and Shimojo (1990). The monocular line adjacent to the binocular rectangle appears in depths different from that of rectangle. b. A variant of the da Vinci stereopsis stimulus. The geometry is similar to the stimulus above a, however, the line composed of black and white elements does not stands out in depth as in a.

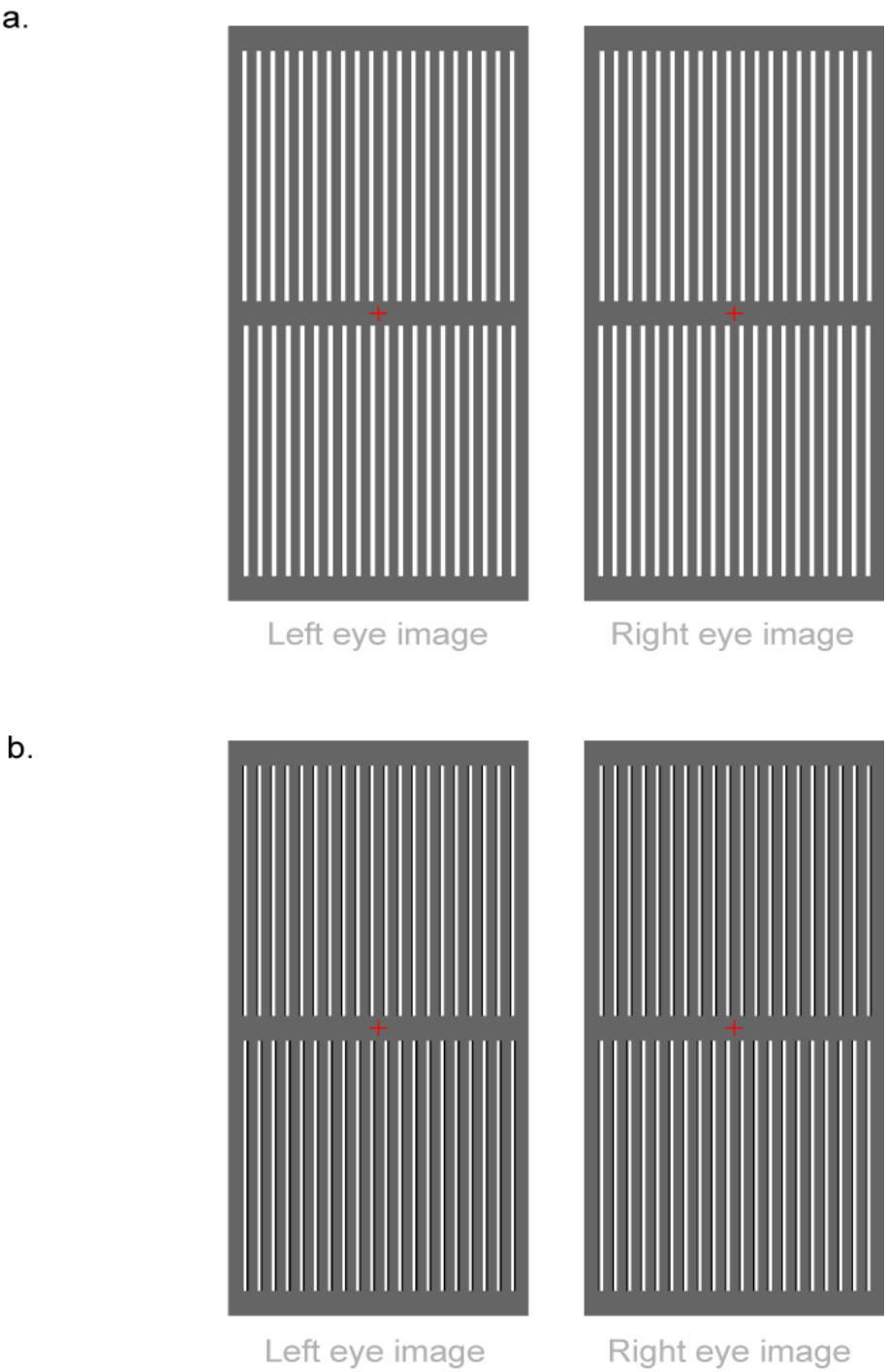


Figure A.10. Depth impression with two different configurations. a. Different disparities assigned to the white lines in the upper and bottom half of the stereograms produces two different depth. b. Stimulus with Da Vinci stereopsis configuration. The black monocular components insert apparent depths to the binocular white lines that have no obvious disparity.