PERGAMON

Vision Research 38 (1998) 2913–2932

Vision Research

Stereoscopic transparency: a test for binocular vision's disambiguating power[1]

Sergei Gepshtein [2], Alexander Cooperman *

Brain Research Center, Department of Neurobiology, The Weizmann Institute of Science, 76100 Rehovot, Israel

Received 20 May 1996; received in revised form 11 February 1997; accepted 19 February 1998

Abstract

It has been suggested that to resolve ambiguities implicit in binocular perception of complex visual scenes, the brain adopts a continuity constraint assuming that disparities change smoothly with eccentricity. Stereoscopic transparency is characterized by abrupt changes of binocular disparity across retinal locations. The focus of the present study is how the brain uses the continuity constraint in the perception of stereoscopic transparency despite the presence of abrupt disparity changes. Observers viewed random-dot stereograms of overlapping transparent plane and cylindrical surfaces and had to distinguish between two orientations of the cylindrical surface under conditions of strictly controlled depth fixation. Surprisingly, maximal dot density of the transparent plane at which perception is still veridical dramatically decreases as depth separation between the surfaces grows. Persistence of this relationship, when binocular matching processes at each surface are separated to on and off brightness channels, suggests at least two stages in the underlying computation binocular matching and inter-surface interactions. We show that these phenomena cannot be accounted for by either higher severity of matching with high dot densities or the ability of the denser surface to pull vergence to its depth. We also measure contrast sensitivity and near–far symmetry of the underlying mechanism and propose a model of competitive interactions between dissimilar disparities.

1. Introduction

Each object in the binocular visual field is seen from two different vantage points and creates an image in the retina of either eye. These images make up the two retinal arrays, when there are several objects in the binocular scene. The retinal arrays are disparate, when the visual scene contains objects at different depths. To extract the depth from the binocular disparities between the two retinal arrays, the brain matches individual images before it elaborates their shapes [1]. Each image of one retinal array could be matched with several images of the second retinal array. All those pairings but one are false matches, or 'ghosts'. The brain must solve the problem of matching ambiguity called correspondence problem or binocular matching problem to eliminate the false matches.

A first step towards understanding how the brain solves the correspondence problem is to hypothesize constraints that the brain may use to reduce the difficulty of the problem. One such constraint is the continuity constraint (CC) [2], which reflects the facts that (a) matter is separated into objects, and that (b) depth-variations of object surfaces are generally small compared to their distance from the observer. The CC hypothesis claims that the brain tries out many matches between the two retinal arrays and preserves only those that are consistent with a smooth shape in depth.

What happens when the visual scene contains a transparent surface, or when we view the scene through a fence or the branches of a tree? In these situations, called stereoscopic transparency[3], the visual objects cast

* Corresponding author. Fax: +972 8 9344140; e-mail: alex@nisan.weizmann.ac.il.

[1] A part of this study was reported at the 19th European Conference on visual perception, Strasbourg 1996.

[2] Department of Psychology, Gilmer Hall, University of Virginia, Charlottesville, VA 22903, USA.

[3] Also, these situations are often called partial stereoscopic transparency (semitransparency) or a screen-door problem e.g. [37] to remind us that nearer objects or elements of the nearer surfaces occlude some elements of the rear surfaces. Stereoscopic transparency is different from 2D transparency [41,42].

0042-6989/98/$19.00
PII: S0042-6989(98)00091-1

retinal images with binocular disparities that change abruptly at nearby retinal locations. Nevertheless, our visual system readily resolves stereoscopic transparency even in random-dot stereograms (RDSs), where all the dots are identical, and binocular disparity is the only depth cue.[4] In RDSs images of elements of one surface can match with images of elements of other surfaces, which further aggravates the correspondence problem.

To figure out whether the CC is compatible with the perception of stereoscopic transparency, let us consider the computational algorithms of stereopsis which use the CC. In these algorithms, the CC is implemented in one of two ways:

- using inhibitory (competitive) interactions between the disparity detectors to suppress false matches [2,3]
- allowing binocular matching only in the borders of disparity gradient limit [4] (PMF), [5], which is a limiting rate of disparity change across retinal locations [6].

The first class of algorithms fails to resolve stereoscopic transparency, because the matches for one of the surface are treated as false matches (matching noise) and suppressed. On the other hand, neither the PMF nor the Prazdny algorithms has inhibitory interactions between disparities that violate the CC, and both resolve stereoscopic transparency. The contrast between the two classes of algorithms might lead us to infer that inhibitory interactions are incompatible with the perception of stereoscopic transparency. However, both psychophysical [7–9] and neurophysiological [10] evidence indicates the existence of inhibitory interactions between disparity detectors tuned to dissimilar disparities (henceforth dissimilar disparities).

How does the brain reconcile the apparent contradiction between the CC and the perception of stereoscopic transparency?

In this study we tackle this question and present the observers with a stimulus which contains both smooth surfaces and stereoscopic transparency. The stimulus is a sparse RDS which depicts two overlapping surfaces: a plane and a convex cylinder. The cylinder, which is the target, is positioned behind the transparent plane. The observers' task is to report the orientation of the target, which may be either vertical or horizontal. We vary the dot density of the plane (ρ_p) and measure the ρ_p at which discrimination of target orientation is no better than at chance, for different depth separations (δ) between the surfaces.

We find that as δ grows, it takes less dots in the transparent surface to mask the target. In other words, the perception of stereoscopic transparency diminishes as the separation between the surfaces increases. We call this phenomenon farther-worse effect.

The observer may facilitate the perception of stereoscopic transparency by sequentially fixating at the depth of each surface. Indeed, previous studies suggested a critical role of eye movements between the depths (vergence eye movements) in the perception of stereoscopic transparency. In those studies, observers were asked to separate and order in depth several planes depicted in RDS of varying dot density [7,11,12]. Most observers needed long presentation times to achieve the impression of transparency and typically reported that it required a considerable 'effort' for the overlapping surfaces to pop out in depth. The role of time in resolving overlapping surfaces implied a critical role of vergence: while fixation is maintained at an appropriate depth, binocular matching may be resolved at each surface separately. Also, the strength with which binocular stimuli attract vergence depends on the strength of the stimulus, e.g. its intensity or dot density [13,14]. A denser surface may 'lock' vergence at its depth in an RDS depicting overlapping surfaces of varying dot density.

To minimize the role of vergence in the perception of stereoscopic transparency, we (a) use an enforced fixation at the depth of one of the overlapping surfaces, and (b) present the stimulus for durations comparable to the time required to initiate a vergence eye movement. This does not allow the observers to fixate at the depths of each of the surfaces during presentation of the stimulus. The results show that the farther-worse effect is independent of the fixation depth, which suggests that the deterioration of performance with greater δ is independent of vergence.

We then address the question, How can the visual system use the inhibition between dissimilar disparities to resolve stereoscopic transparency?

To answer this question, we need to distinguish between two effects which might hamper resolution of stereoscopic transparency: (a) lower stereoscopic efficiency [15], where higher dot densities make the correspondence problem more severe, and (b) possible interactions between binocular disparities, such as inhibition between dissimilar disparities. Stereoscopic efficiency reduces as the dot density grows in an RDS depicting non-transparent surfaces, due to the difficulty to extract a depth signal from the noise of false targets. Similarly, higher dot densities of RDS hinder separation of surfaces in stereoscopic transparency [7]. We

[4] Matching primitives, more complex than dots, such as oriented tokens of different size, can paradoxically render RDS an almost unambiguous stimulus for stereopsis under specific conditions [34,38]: A limited number of identical features remain in the output if a dense RDS is passed through monocular linear filters tuned to a range of orientations and spatial scales. However, dotbased matching is necessary, e.g., to resolve transparency in an RDS where each even row (column) portrays one depth surface, while each odd row (column) portrays another surface [40]. Evidence regarding which matching primitives are used in biological vision is inconsistent [36]. Perhaps, vision utilizes different matching primitives as well as matching strategies depending on the stimuli and the tasks it faces.

analyze the geometry of binocular matching under the CC and under another common matching constraint the ordering constraint (OC) and show that matching noise can, at least qualitatively, account for the farther-worse effect. This account predicts that increasing the dot density of the target surface will also hamper performance. We find, however, that the performance is facilitated with higher density of the target surface.

Furthermore, reversing the depth order of the two surfaces changes characteristics of the farther-worse effect, which cannot be explained by the destructive influence of matching noise. We also use another test prompted by the recent finding that on and off brightness channels [16] are independent at the level of binocular matching [17]. We apply opposite contrast polarity to the dots of different surfaces and thus separate binocular matching at each surface to on and off. This allows us to separately examine the effects of stereoscopic efficiency and interactions between the surfaces. We find that the farther-worse effect persists with opposite contrast polarities, although the performance is facilitated when the dots of the two surfaces have different luminances. The observed facilitation is quantitatively different from the predictions of the OC. The persistence of the farther-worse effect indicates that it is independent of the false matches between the images of different surface elements. We conclude that the farther-worse effect characterizes an interaction between the surfaces which occurs separately from binocular matching at each surface.

We offer a model of inhibitory connections at the hyperglobal[5] level of inter-surface interactions, consistent with the concept of disparity gradient limit. In the model, active binocular unit inhibit the units tuned to dissimilar disparities so that the zone of inhibition has a conelike shape in the disparity map: The lateral extent of the inhibition grows at higher depth separations.

Finally, the model led us to perform other experiments (section Experiment 5–6), where we ruled out matching noise explanation of the farther-worse effect, and investigate:

- The role of target strength in the perception of stereoscopic transparency: In agreement with our model, an increment of the target dot density improves performance, which further rules out the false matches between different surface images as an explanation of the farther-worse effect.
- The symmetry of the inter-surface interactions in the domains of crossed (near) and uncrossed (far) disparities: We show that the inter-surface interactions also operate when the target is positioned in front of the masking surface, but the characteristics of the inhibition differ between the near and far domains.

The rest of the paper has the following structure: After a detailed explanation of the experimental settings in General methods, we present the experiments which reveal the father-worse effect with different fixation depths and different luminances of the overlapping surfaces (Experiments 1–4). Then, we turn to inhibition in the disparity domain and a model of inhibitory interactions. Finally, we present the experiments about the strength of the target surface and the near–far asymmetry of the father-worse effect (Experiments 5–6).

2. General methods

Apparatus stereograms were displayed in a darkened room on a Mitsubishi color monitor with a screen of 1280×1024 pixels subtending about $25.6 \times 20.4°$ with the viewing distance of 90 cm. Stimulus generation and display were controlled by a Silicon Graphics Crimson workstation using a Silicon Graphics Reality Engine system. The disparate images, displayed using a flicker-free method (StereoGraphics system, 120 fields/s rate), were viewed with CrystalEyes shutter glasses synchronized with the presentation of the left and the right images (half-images) by infra-red emitter.

2.1. Observers

The observers were one of the authors (SG) and three paid subjects, naive as to the purpose of the experiment. All the observers had normal stereoscopic vision and normal or corrected-to-normal visual acuity.

2.2. Stimuli

The three stimuli fixation pattern, the test stimulus and the mask were presented on a background with luminance of 2 cd/m^2 except for Experiment 3. The viewing distance in all experiments was 90 cm. Each trial consisted of a fixation pattern, a test stimulus and a mask, shown in a rapid succession:

2.2.1. Forced fixation

To control the depth of fixation, a stereogram of two vertical lines (luminance 94 cd/m^2; size $1.2 \times 84'$ arc), on the background of a sparse random-dot plane, preceded each test stimulus. When stereoscopically fused, one of the lines was perceived as slightly closer to the observer. The depth separation between the lines

[5] Following [27,39], we distinguish between the three levels in stereopsis: (1) local, where information from the left and the right eyes is first combined at each retinal location disregarding the disparities at other locations, (2) global, where the disparities at different locations interact to produce a smooth solution for the correspondence problem, and (3) hyperglobal, where several global solutions exist in the same visual direction. Stereoscopic transparency is a case of hyperglobality where the overlapping surfaces make (simultaneous) global solutions in the same visual direction.

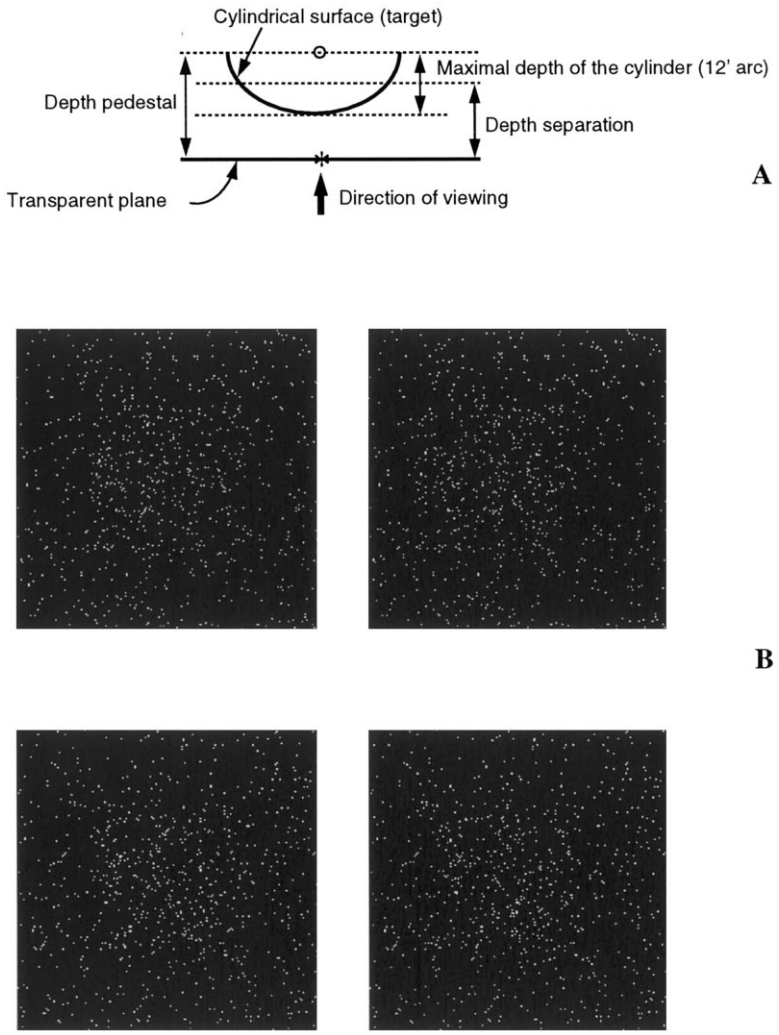

Fig. 1. A. Horizontal cross-section of the test stimulus with the vertical cylinder (target). To compute the depth separation between the surfaces (δ we subtracted half of the maximal cylinder disparity from the disparity of its pedestal. The asterisk marks the depth of fixation in Experiment 1, which was changed in Experiment 2 (open circle). B. Two random-dot stereograms illustrate stereoscopic configurations similar to our test stimuli. When fused with crossed disparities, vertical (top) or horizontal (bottom) convex cylindrical surface is seen through the lace of the transparent plane.

was 2.4′ arc. The sparse RDS plane at the depth of one of the lines made fixation at the desirable depth easier and more stable.

At the end of the trial, the observers were asked to respond which of the lines left or right was closer. If the answer was wrong, the response of the observer to the test stimulus in that trial was discarded and a beep was heard. In Experiment 1, the lines and the plane of the fixation stereogram were at the depth of the transparent plane of the test stimulus (zero depth; Fig. 1A).

2.2.2. Test stimulus

The test stimulus was a sparse RDS occupying a central rectangular region ($18 \times 14°$) of the display. The dots subtended $1.2 \times 1.2'$ arc and had luminance of 94 cd/m^2 (the latter was changed in Experiments 3–4). When fused, the test stereogram was perceived as a convex vertical, or horizontal, cylindrical surface (target) seen through a semitransparent plane (Fig. 1B).

We computed the binocular disparities between the two half-images of the test stereograms in two independent cyclic processes, each producing one of the two surfaces the plane or the cylinder (Fig. 1A):

- The plane. In one of the processes, over the whole stereogram, each dot was repeated in the left and the right half-images with the same co-ordinates to produce a stereogram of a plane at zero depth ($18 \times 14°$). The dot density of the cylinder was 1% in all the experiments, except for Experiment 5.
- The cylinder. In the second process, over a smaller central patch of the stereogram ($7.6 \times 5.6°$), each dot was repeated in the left and the right half- images with a horizontal displacement computed according to the current co-ordinates of the dot to produce a

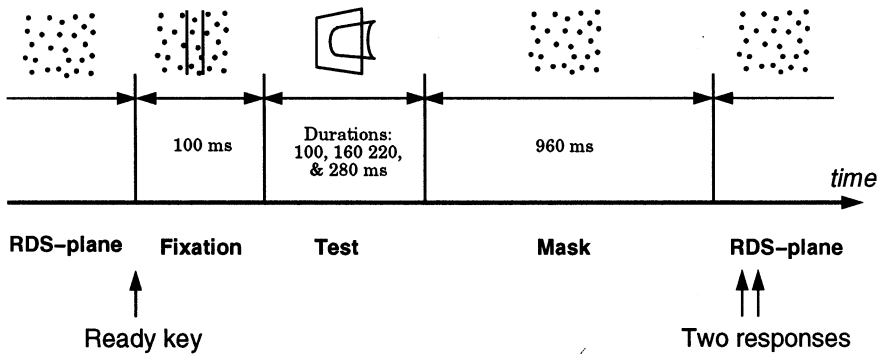

Fig. 2. The trial temporal sequence. After the observers pressed the ready key in response to presentation of the random-dot stereogram of a plane, fixation pattern, test stimulus, and mask were shown in a rapid succession without intervals. In the end of the trial, the observers were asked to give two responses: first, whether the left or the right vertical line of the fixation pattern was closer, and second, was the target surface in the test stimulus vertical or horizontal.

stereogram of a central vertical or horizontal convex cylinder, mounted on a depth pedestal[6]. Maximal depth of the cylinder (Fig. 1A) was 12′ arc. The vertical and horizontal cylinders had the same width and height (the same 'perimeter') in the frontal plane. We followed the convention that when the plane was seen in front of the cylinder, the depth pedestal had positive binocular disparity. The dot density of the transparent plane varied as explained below.

Within each iteration of each of the two processes, a random number generator was employed twice to produce the horizontal and vertical co-ordinates of a dot. This resulted in a roughly uniform distribution of the dots over a chosen part of the stereogram. To compute the dot density, we divided the number of dots in a surface by the area circumscribed in the 'perimeter' of that surface, i.e. by the total number of all possible dot positions inside the area. For example, 300 random-dots in 100×200 array make dot density of $300/(100 \times 200) = 1.5\%$.

In each experimental block we wanted to compare observer's performance with two different dot densities of the transparent plane. To do that, we generated a new set of four test stimuli, two with vertical and two with horizontal cylinders, at the beginning of each experimental or pilot block. The test stimuli with equally oriented cylinders differed by the dot density of the transparent plane.

2.2.3. Mask

The mask was an RDS of 20×20 planar rectangular patches, each at a random depth. The overall size and maximal depth of the mask, as well as the size and luminance of the dots, were close to those of the test pattern. The dot density of the mask and the test were similar to ensure that their overall luminances were close.

[6] A strong transparency effect may be produced by multiple repetition of micropatterns, each similar to the Panum's limiting case ([35]) or double-nail (dot) micropatterns ([12]). We did not use these configurations here because dot densities for individual surfaces cannot be manipulated independently in both cases.

2.3. Procedure

The trial temporal sequence is shown in Fig. 2. The observers viewed an RDS of a plane at zero depth and pressed the ready key to call the fixation pattern, which was shown for 96 ms. The test stimulus followed the fixation for variable duration, and was immediately succeeded by the mask for 960 ms. (There was no delay between the test and the mask.) An RDS of a plane at zero depth identical to the plane shown at the beginning of the trail followed the mask, which prompted the observers to give two responses and initiate the next trial. The two responses at the end of the trail were: (1) which of the lines of the fixation pattern-the left or the right was closer, and (2) what was the orientation vertical or horizontal of the target. There were no gaps in the sequence of stereoscopic stimuli, which provided the observers with a continuous clue for vergence and helped them to correctly solve the fixation task.

In pilot experiments, we established the durations of the test stimulus (henceforth test durations) to plot a psychometric curve of correct answer rate as a function of test duration for each of the four test stimuli in a block, i.e. so that at the shortest test duration the observers performed at the chance level, while at the longest test duration the performance was near perfect. We used four different test durations in each experimental block: 96, 160, 224, and 288 ms. Only the two largest test durations permitted vergence eye movements during the presentation of the test stimulus.

The distribution of binocular disparities in the cylinder resulted in an uneven density of the dots over one of the half-images. This inequality could serve as a monocular cue to the cylinder orientation. This prob-

lem could be avoided in a task of convex versus concave horizontal cylinders, instead of the task vertical versus horizontal convex cylinders, which we used. We did not use the convex versus concave task, because it would produce a high variability of the intersurface depth separations, the parameter we wanted to control. Control experiments showed that with the dot densities and test durations used in the experiments, the performance was at chance level when the stimuli were presented either monocularly or with inactive shutter stereo-glasses, and the density inequalities were the only clue about orientation of the target.

Out of four different test patterns presented to the observer in each experimental block (60 presentations of each kind), two contained vertical and two horizontal cylinders. Typically, the two test stimuli with equally oriented cylinders differed by the dot density of the transparent plane, while the depth separation between the plane and the cylinder was the same for four test stimuli.

Each session typically comprised several short pilot blocks and three to five experimental blocks, each 240 trials long (four test durations × four test conditions × 15 trials).

3. Experiments 1–4

3.1. Experiment 1: shape discrimination drops at higher depth separations between the surfaces

In introduction we discussed the computational advantages of the continuity constraint in solving the correspondence problem. We said also that we readily perceive stereoscopic transparency, even though it contains abrupt depth changes at nearby retinal locations. To figure out how the brain could use the continuity constraint in this case, we wanted to quantify the limitations of the human observers in perception of stereoscopic transparency, using a stimulus which contains smooth overlapping surfaces. Specifically, in this experiment our purpose was to explore how perception of a rear surface depends on the dot density of the transparent surface at various depth separations between the surfaces.

3.1.1. Method

Observers initiated each trial and viewed the sequence of fixation, test and mask patterns. The fixation pattern contained two vertical lines at slightly different depths; the fixation task was to tell which line is closer. The test stimulus depicted a vertical or horizontal cylindrical surface (target) seen through a transparent plane; the test task was to tell the orientation of the target. The fixation pattern was positioned at a depth of the transparent plane of the test stimulus, and the observers had to fixate at the required depth to be able to solve the fixation task. At the end of each trial observers gave two answers: to the fixation and to the test tasks.

We measured the proportion of correct answers as a function of test duration, as we varied the two parameters depth separation between the surfaces (δ), and the dot density of the transparent plane (ρ_p). The dot density of the target (ρ_c, where index c stands for cylinder) was constant and equal to 1%. At any given δ, we increased ρ_p through the experimental blocks until the limiting density (ρ_p^*, see definition below) was reached. We repeated this procedure across δ in order to plot the relationship between ρ_p and δ.

We measured ρ_p^* as follows. First, we obtained close to perfect performance at low ρ_p (Fig. 3A). Second, we increased ρ_p over experimental blocks, with other parameters held constant, until the discrimination between the two orientations of the target fell below a criterion of 75% correct. We defined ρ_p^* as the highest ρ_p at which discrimination between the two orientations of the target was veridical. High ρ_p^* means that the task is easy, and many random-dots need to be added to the transparent plane to reduce performance below the 75% criterion. In other words, ρ_p^* measures the efficiency of target discrimination.

In each block, we applied two different ρ_p's to the test stimuli with the same orientation of the target. Each session began with a number of pilot blocks, 48 or 80 trials long, used to estimate the operative range of ρ_p. In subsequent experimental blocks, 240 trials long, ρ_p differed between the two test patterns, or between successive blocks, by several tenths percent, which defined the precision of our measurements. This procedure appeared robust for two reasons: (1) sensitivity-when ρ_p^* was approached, an increment of ρ_p by 0.1–0.2% typically caused a dramatic deterioration of

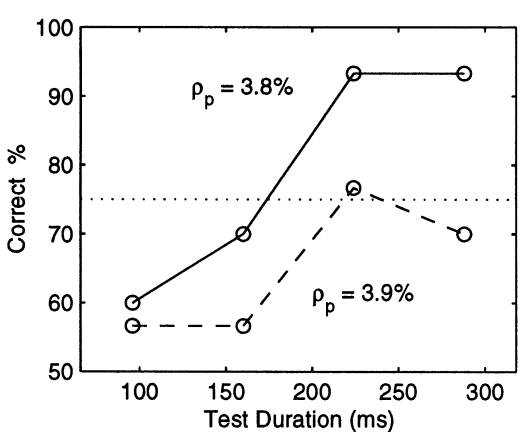

Fig. 3. Typical results from one experimental block (observer OK). We plot correct answer rates averaged for the vertical and the horizontal targets as a function of test duration. The dotted line marks the 75% criterion level. With $\rho_p = 3.9\%$ performance is much worse (dashed line) than with $\rho_p = 3.8\%$ (solid line).

performance; (2) reproducibility-when we repeated the experiment in the same observer on the next day or even a week after the first measurement, ρ_p^* typically was very close (within 0.1–11.2%) to ρ_p^* obtained in the previous measurement, except for the domain of small δ, where the measurements were noisier (see Results section). An example of the data from one experimental block is shown in Fig. 3.

3.1.2. Results

In Fig. 4, we plot the $\rho_p^*(\delta)$ relationships (circles) for three observers, with the test duration of 220 ms. Each circle in the figure corresponds to at least 80 trials. The inverse relationship between ρ_p^* and δ for all the observers indicates that shape discrimination of the cyclopean target dramatically deteriorates as δ grows. The data show that the procedure is highly sensitive to changes of ρ_p, measured by tenths of a per cent at the discrimination threshold. Henceforth, we refer to the decreasing relationship of ρ_p^* versus δ as the farther-worse effect.

3.1.2.1. Time considerations. We obtained the results shown in Fig. 4 with the test duration of 220 ms. With shorter test durations 100 and 160 ms the task was difficult for all the observers: they tired quickly even with low ρ_p, although veridical perception was still possible in short blocks of 80 trials or less (which is equivalent to five trials or less for each data point of the psychometric curve). The test duration of 220 ms was the lowest duration at which correct performance was consistent across the experimental blocks.

3.1.2.2. The range of depths. The large range of binocular disparities over which the observers could correctly identify the target, with test duration as short as 220 ms, is consistent with the findings that stereoscopic fusion tolerates disparities up to 2° [18,19].

3.1.2.3. Variance. At low δ, the measurements of ρ_p^* were noisier: we attribute this effect to pyknostereoscopic (from Greek for dense[20]) depth averaging, which typically occurs when depth separation between the surfaces is small [20,21]. We believe that the depth averaging effect impedes extraction of the disparities associated with the target, and thus reduces the reliability of our measurements.

3.1.3. Discussion

Binocular matching produces more spurious pairings when the dot density of the RDS is high [15]. This might explain why performance deteriorates as ρ_p grows at any particular δ, which is captured by limiting dot density ρ_p^*. We also found that ρ_p^* drops as we increased δ; we refer to this phenomenon as farther-worse effect. Could reduction of performance with higher dot densities account for the farther-worse effect?

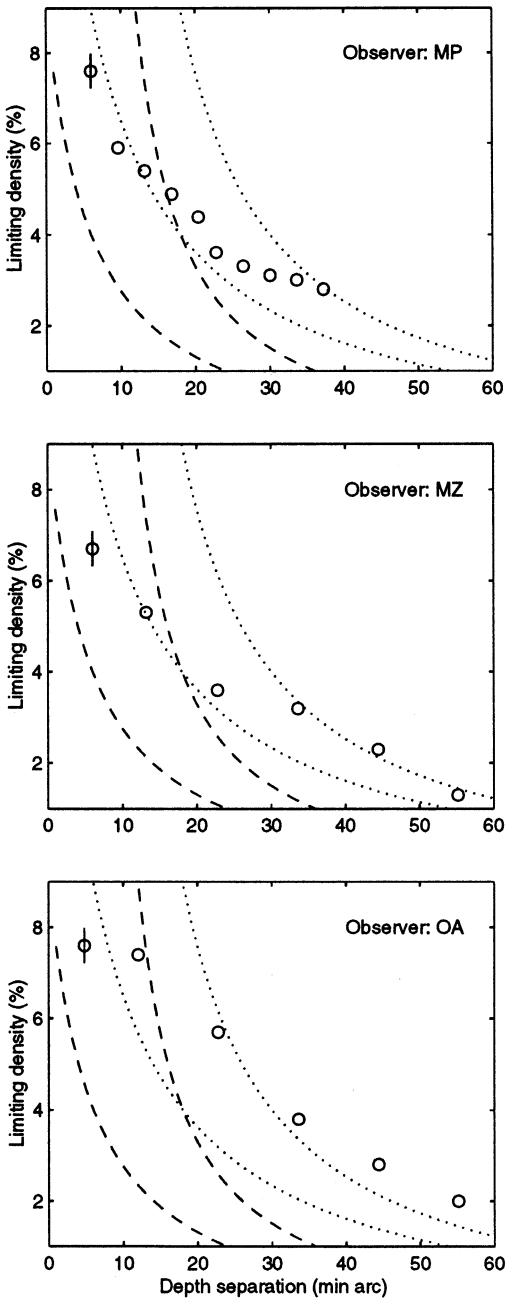

Fig. 4. Results of Experiment 1 for three observers. Each graph represents the limiting densities (ρ_p^*) plotted with circles as a function of depth separation (δ) between the surfaces of the test stimulus with the test duration of 220 ms. When δ was greater than 6′ arc, standard errors did not exceed 0.15% (not shown in the graphs because of their small size). The standard errors increased to 0.3–0.5% with low δ, which may be accounted for by the pyknostereoscopic averaging effect (see text). The dashed lines delimit the range of limiting densities predicted by the effect of continuity constraint (CC) on resolution of stereoscopic transparency (see text). The dotted lines delimit the range of limiting densities predicted by the ordering constraint (OC) on resolution of stereoscopic transparency (see text).

To answer the question of whether matching noise alone might cause the farther-worse effect, we assume that the continuity constraint (CC) is used in binocular matching; that is, the brain prefers small binocular disparities and searches to minimize the overall disparity in the reconstructed surfaces. First, let us consider how images of the transparent plane and the cylinder dots are distributed on the retinas. The geometry of binocular vision facilitates our task, because the corresponding images in the two retinas fall onto corresponding lines called epipolar lines and matching is reduced to a one-dimensional problem. Let P and C stand for the retinal images of the dots of the transparent plane and of the cylinder, respectively, the superscripts of these symbols stand for the eye (L for the left and R for the right) of the images, and the subscripts stand for the sequential order of the surface elements along the epipolar lines. For example, C_i^{L}represents the left image of the ith dot of the cylinder, and C_{i+1}^{R} represents the right image of its $(i+1)$th dot. In our test stimuli, P images have the same epipolar co-ordinates in both retinas. The epipolar co-ordinates of C_i^{L} and C_i^{R} images are disparate: Their binocular disparity varies in the range of $\Delta = \delta \pm 6'$ min. As δ grows, the disparity between C_i^{L} and C_i^{R} increases and at some point the disparity of the correct match $[C_i^{\mathrm{L}}\ C_i^{\mathrm{R}}]$ will exceed the disparity of one of the two false matches: $[C_i^{\mathrm{R}}\ C_{i+1}^{\mathrm{L}}]$ or $[C_i^{\mathrm{R}}\ P_j^{\mathrm{L}}]$. The CC assumption predicts that in this situation the false matches with the smaller disparities will have a stronger weight in the computation of the global shape than the correct match.[7]

Based on the CC assumption, we can compute the $\rho_{\mathrm{p}}^*(\delta)$ relationship predicted by matching noise alone. Remember that to produce a monocular array of the test stereogram, we first created a sparse array of random-dots (plane-dots) with the dot density of ρ_{p}. Then, we added more random-dots (target-dots) to the target area, so that the resulting dot density in this area became $\rho_t = \rho_{\mathrm{p}} + \rho_{\mathrm{c}}$ and the average horizontal inter-dot distance became $\bar{\mathrm{h}} = 100s/\rho_t$, where s stands for the dot width in angular minutes, and ρ_t is measured in per cent. Next, we copied the resulting array into the second half of the stereogram and shifted its target-dots laterally by to produce a distribution of binocular disparities yielding perception of a cylindrical surface in the fused stereogram. When Δ exceeds $\bar{\mathrm{h}}/2$, the binocular disparities of most correct matches $[C_i^{\mathrm{R}}\ C_i^{\mathrm{L}}]$ will exceed the disparity of one of the two types of false matches: $[C_i^{\mathrm{R}}\ C_{i+1}^{\mathrm{L}}]$ $[C_i^{\mathrm{R}}\ P_j^{\mathrm{L}}]$. The CC predicts that at that point the false matches with the smaller disparities will be preferred, and the resulting matching noise will destroy the correct reconstruction of the target surface. Thus, the observers' ability to resolve the stereogram will drop when Δ approaches $\bar{\mathrm{h}}/2$. From $\bar{\mathrm{h}} = 100s/\rho_t$ and $\Delta = \bar{\mathrm{h}}/2$ follows that the matching noise will encumber resolution of the target when ρ_{p} reaches $\rho_{\mathrm{p}}^* = (100s/2\Delta) - \rho_{\mathrm{c}}$. We compute the ρ_{p}^* and plot them, as a function of Δ, with two dashed lines in Fig. 4: The dashed lines demarcate the range of predicted ρ_{p}^* with two extreme values of Δ; that is with $\Delta = \delta - 6'$ when only the disparities on the top of the convex cylinder systematically produce false matches (the right dashed curve), and with $\Delta = \delta + 6'$ arc, when all the target disparities are false.

As clear from the simulation, the CC predicts that performance will deteriorate as a function of δ, but the rate of the deterioration will be steeper than we found in Experiment 1. Ordering constraint (OC) is another constraint which is often considered in the context of the effect of matching noise on the reconstruction of stereoscopic depth.[8] The OC forbids binocular matching which violates the order of images in the two retinas. To use our notation, most of the matches will violation the order when Δ exceeds $\bar{h}$, rather then $\bar{\mathrm{h}}/2$, and $\rho_{\mathrm{p}}^* = (100s/\Delta) - \rho_{\mathrm{c}}$. The dotted lines in Fig. 4 represent the OC predicted ρ_{p}^* for $\Delta = \delta - 6'$ (the right dotted curve), and $\Delta = \delta + 6'$. As clear from the figure, the OC permits resolution of transparency over a broader range of δ than the CC does. Results of Experiment 1 fall close to the range of ρ_{p}^* predicted by the OC. Note, however, that the results of Experiment 1 are based on a less-than-perfect performance (75% correct). In our simulation of the effect of matching noise, we assume that most of the binocular matches must turn false to hamper reconstruction of a surface. Since we do not know what proportion of false matches will reduce performance to the level of 75% correct, we can only draw a qualitative conclusion that the matching noise predicted by the OC might account for our findings in Experiment 1. Also, our computations of the theoretical curves was based on a number of assumptions, which weakens the predictive power of the simulations. To decide the issue, we search for other evidence in the following experiments and discover that the hypothesis of the OC makes an unlikely explanation of the farther-worse effect:

- In Experiment 5 we find that discrimination of the target is facilitated with higher ρ_{c}, the dot density of the target surface. By contrast, the ordering constraint hypothesis predicts that increasing dot density of either surface will hamper performance.

[7] Notice that the *CC* agrees with a more general principle of interaction between the visual stimuli. The relative proximity between the stimuli typically controls their interaction in the static e.g. [44] and dynamic e.g. [43] visual scenes. The *CC* implies that the binocular matches between the images with the closest inter-ocular co-ordinates are preferred. Thus, the *CC* may be viewed as a binocular equivalent of the Gestalt principle of grouping by proximity.

[8] We thank one of the anonymous referees for drawing our attention to the fact that operation of the *OC* may account for the farther-worse effect.

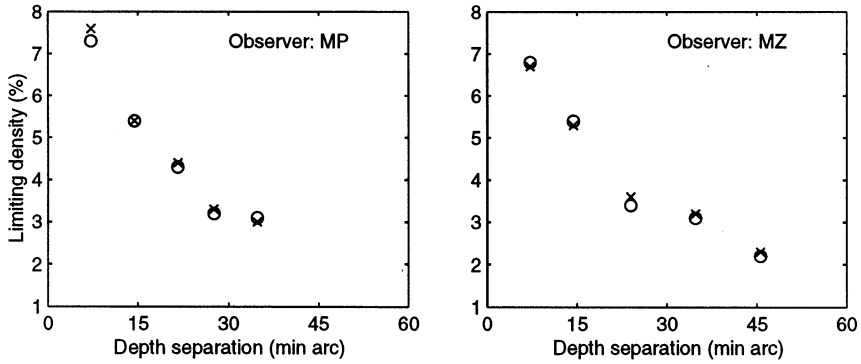

Fig. 5. Results of Experiment 2 (open circles) for two observers. Cross symbols (×) show data from Experiment 1. Standard errors were not larger than in Experiment 1. The results show that displacement of fixation pattern to the depth of the cylinder pedestal does not affect the basic farther-worse effect reported in Experiment 1 (Fig. 4).

- The ordering constraint hypothesis predicts that the depth order of the target and the masking surfaces will not affect the farther-worse effect, as long as the dot densities of the surfaces do not change. We find, however, in Experiment 6 that the farther-worse effect is influenced by reversing the depth order of the two surfaces.

Searching for other explanations of the farther-worse effect, we consider two further hypotheses: (i) the vergence hypothesis and (ii) the competitive interactions hypothesis. The vergence hypothesis rests on the evidence that a disparate stimulus in the visual field elicits vergence eye movements directed to its depth [13,14,22]. Along these lines, the stronger the disparate stimulus of the transparent plane (the strength may be a function of the dot density in the RDS [14]) the more time is required for vergence eye movement to reach the target surface.

The hypothesis of competitive interactions explains the lower performance at higher dot densities by inhibitory interactions between binocular disparity detectors (or pools of the disparity detectors) tuned to different disparities [7]. First, we rule out the vergence hypothesis in Experiment 2; then we address the hypothesis of competitive interactions and compare it with the predictions of the OC.

3.2. Experiment 2: farther-worse effect cannot be accounted for by the vergence hypothesis

In Experiment 1 we found that the resolution of stereoscopic transparency is hurt at larger depth separations (δ) between the surfaces; we call this phenomenon farther-worse effect. As we discussed in the introduction vergence eye movements might account for the farther-worse effect: Binocular disparities elicit vergence eye movements, and random-dot images provide more stimulation for the vergence system than line targets or single dots [13,14]. In a similar way, a depth plane with higher dot density provides more stimulation for fusion and vergence. The farther-worse effect may reflect this dependence between the pulling strength of the transparent plane and its dot density.

The purpose of this experiment is to examine whether an increase of transparent plane pulling strength, produced by its higher dot density (ρ_p), causes the decrease in the limiting density (ρ_p^*) at greater depth separations (δ) between the surfaces (farther-worse effect). If a denser transparent plane slows down vergence movements directed to another surface, then higher ρ_p will hamper target discrimination at the same δ and test duration. This possibility is consistent with the results of Experiment 1.

The vergence hypothesis also predicts that if the observers fixate at the depth of the target, then the farther-worse effect will either disappear or at least be significantly reduced. To test this prediction, the same procedure as in Experiment 1 was employed, but now the fixation pattern was placed at the depth of the cylinder pedestal. Thus, fixation depth varied across the blocks as we changed the depth separation between the plane and the cylinder.

3.2.1. Results

In Fig. 5, we plot ρ_p^* as a function of δ for two observers. The results of Experiment 1 with fixation pattern at zero depth are replotted for comparison. As clear from the figure, there is no significant difference between the two curves for both observers. Thus, the vergence hypothesis cannot explain the farther-worse effect. Vergence eye movements may still be involved in discrimination of the cyclopean target over the examined range of δ (remember that the test duration of 220 ms was required for consistently veridical perception), but the farther-worse effect is independent of vergence.

3.2.2. Discussion

After we have shown that the farther-worse effect is independent of vergence, we will consider the hypothesis of competitive interactions between dissimilar dis-

parities, i.e. between disparity detectors tuned to sufficiently different disparities.

Using a method of measuring statistical efficiency of stereoscopic perception, it was recently found that increments and decrements of brightness (transmitted through on and off neural channels, respectively; [16]) are processed independently in binocular vision, at least as far as the level at which information from the left and right eyes is first combined [17]. Furthermore, the increments and decrements of brightness are statistically independent in stereo matching, i.e. the elimination of false matches proceeds independently for the bright and dark elements [17]. These results presuppose that (a) local stereopsis is specific to contrast sign (contrast polarity), and (b) at least two distinct global processes exist (Fig. 6):

- where elimination of false matches occurs separately for increments and decrements of brightness.
- where the individual depth estimates derived separately at on and off channels are combined.

The evidence of separate binocular matching over the increments and decrements of brightness suggests a tool for discovering whether a stereoscopic phenomenon occurs at the level of global disparity processing. If a phenomenon persists when the image elements are of opposite contrasts polarities, then the underlying mechanism cannot be operating before global disparity processing. The converse—if the phenomenon is abolished, it depends on a local disparity processes—is not possible, because both local and global mechanisms work at the levels where the on and off channels are separated (Fig. 6). In other words, the application of opposite contrast polarities makes a test for stereoscopic globality. Since spurious matches between the surfaces are impossible when matching processes at each surface are separated, the test allows one to rule out the possibility that inter-surface matching noise, including the matches which violate the ordering constraint, contributes to the farther-worse effect. We use this test in Experiment 3.

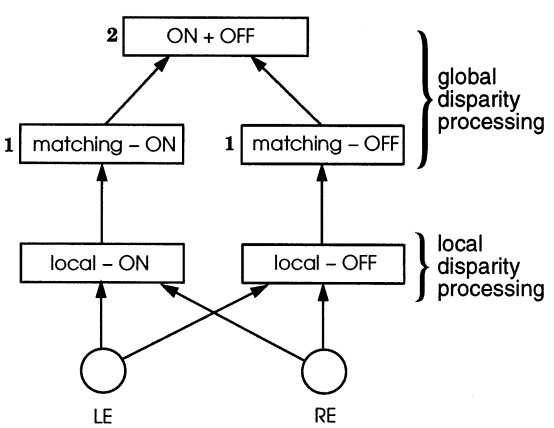

Fig. 6. A scheme of local and global stereoscopic mechanisms in relation to the on and off neural channels in early vision. Psychophysical evidence [17] suggests that local stereopsis is sensitive to the contrast polarity of the stimulus. Global stereopsis has at least two distinct levels, where binocular matching occurs independently in the on and off channels (1), and where the polarity-specific depth estimates integrate (2).

3.3. Experiment 3: farther-worse effect persists with dots of the two surfaces having opposite contrast polarities

In this experiment we apply opposite contrast polarities to the two surfaces of the test stimuli to separate the processes of binocular matching at each surface from possible interactions between the surfaces. This helps us (1) test whether monocular or local stereoscopic mechanisms can account for the findings of Experiment 1, and (2) estimate individual contributions of matching noise and inter-surface interactions into the farther-worse effect.

3.3.1. Method

We changed the contrast polarities of dots of the transparent plane and the cylinder as follows: We increased background luminance to 56 cd/m^2 (as compared to 2 cd/m^2 in Experiments 1–2) and decreased cylinder dots luminance to 2 cd/m^2 (as compared to 94 cd/m^2 in Experiments 1–2). The luminance of the transparent plane dots was 94 cd/m^2. This made the contrasts of the dots of the plane and the cylinder against the background 68 and 96%, respectively.

3.3.2. Results

The data for two observers are shown in Fig. 7 (open circles), compared to the results of Experiment 1 (cross symbols). The results show that application of opposite contrast polarities (a) facilitate the task, and (b) preserve the farther-worse effect.

Does the matching noise under the ordering constraint (OC) predict the results of this experiment? When dots of the two surfaces have opposite contrast polarity, matching noise is produced by false matches between the images of the same surface. As we showed in the Discussion of Experiment 1, the OC predicts that most of the matches will become false and performance will deteriorate, when δ exceeds the average monocular distance between the target images. Using the notation developed in Experiment 1, the limiting dot density becomes $\rho_p^* = 100s/\Delta$. We plot the resulting ρ_p^* with dotted lines in Fig. 7: the OC predicts that performance will be facilitated much less than we found in the present experiment.

The persistence of the farther-worse effect with binocular matching at each surface separated to different neural sites points to the global nature of the farther-worse effect (stereoscopic globality test). This observation further supports the argument that a lower efficiency of binocular matching at higher ρ_p cannot

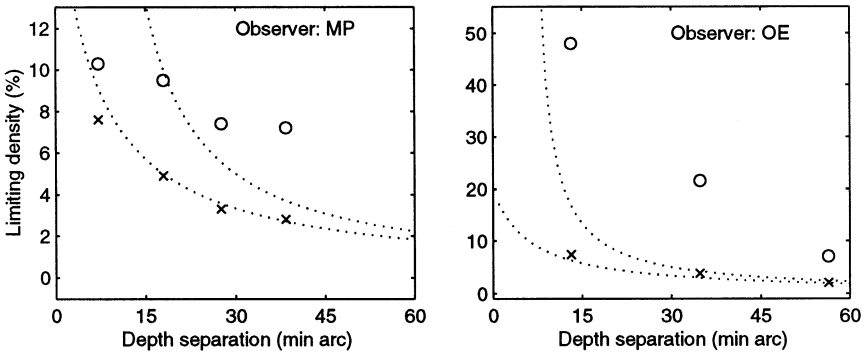

Fig. 7. Results of Experiment 3 (open circles) for two observers. Standard errors were not larger than in Experiment 1. Cross symbols (×) show data from Experiment 1. The farther-worse effect persisted with the opposite contrast polarities of the elements of two surfaces. In the same time, the limiting densities were significantly larger than in Experiment 1. The dotted lines stand for the predictions of the ordering constraint, where the performance is limited by violations of matching order between the images of the target surfaces alone.

account for the farther-worse effect. We believe that the inverse relationship between ρ_p^* and δ reveals inter-surface interactions which occur at a stage later than the binocular matching, i.e., at or above the site designated as on and off in Fig. 6.

The magnitudes of ρ_p^* in this experiment are significantly larger than when elements of the two surfaces had the same contrast. It is unclear, however, whether the facilitation with opposite contrast polarities is due to the separation between binocular matching at individual surfaces, or to a mere contrast difference between the two surfaces. In other words, the question is whether the underlying mechanism is sensitive to contrast polarities or contrast difference. We explore contrast sensitivity of the farther-worse effect in Experiment 4.

3.4. Experiment 4: limiting density is sensitive to the luminance of transparent surface elements

In this experiment we examine whether the facilitation observed in Experiment 3 was caused by the separation of the binocular matching processes at each surface, or that a contrast difference between elements of the two surfaces is sufficient for easier discrimination of the target surface.

3.4.1. Method

The dot contrast of the plane and the cylinder were of the same polarity and we measured ρ_p^* at different dot luminance of the transparent plane. Luminance of the background was 2 cd/m^2 as in Experiments 1–2. In this experiment we were only able to use inter surface contrast differences smaller than the ones we used in Experiment 3; otherwise the dots of the transparent plane would have been invisible against the dark background. While the inter-surface luminance difference used in Experiment 3 was 92 cd/m^2, the maximal luminance difference used in this experiment was 80 cd/m^2 for observer OE and 60 cd/m^2 for observer SG. In this luminance range, the dots were well visible and transparency was readily perceived. For the observers OE and SG, δ was 35 and 24′ arc min, respectively.

3.4.2. Results

In Fig. 8, we plot ρ_p^* with open circles as a function of dot luminance of the transparent plane. The results show that ρ_p^* is sensitive to luminance difference between the surfaces. To show that the results agree with the notion of signal strength in the disparity domain (see below), we fit the data with the quadratic function (the solid lines). Again, we cannot measure ρ_p^* at lower luminances of the transparent plane, where the experience of stereoscopic transparency becomes questionable.

3.4.3. Discussion

As we have shown in Experiment 3, discrimination of cyclopean shape is much easier with opposite contrast polarities of the transparent and the target surfaces. The magnitude of this facilitation is close in Experiments 3–4 for observer OK, the only observer who participated in both. Thus, the facilitation observed in Experiment 3 is attributable to the large difference in luminances (92 cd/m^2) at opposite contrasts polarities, rather than to opposite contrasts polarities per se. We conclude that contrast differences between the surfaces facilitate matching within the surfaces, while the farther-worse effect is due to a mechanism independent of within-surface matching (Experiment 3).

What architecture of inhibitory interactions between disparities could both resolve the severe matching ambiguities in perception of stereoscopic transparency and agree with our observations? In the next section we suggest a scheme of inhibitory connections which uses the concept of disparity gradient limit. First, we will briefly review psychophysical and physiological evidence of inhibitory interactions in the disparity domain

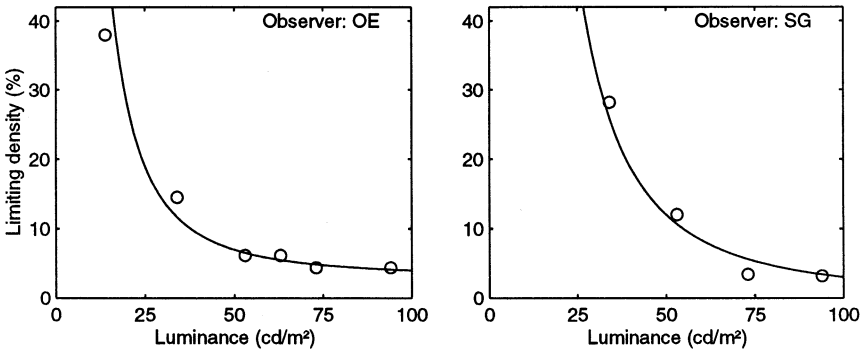

Fig. 8. Results of Experiment 4 for two observers. Limiting dot density of the transparent plane is plotted as a function of the luminance of its dots. The solid lines represent quadratic approximations to the data. The results show that with greater differences between the dots luminances of the two surfaces, the limiting density may be increased by tens per cent.

and, second, review the evidence which suggests that disparity gradient limit is brought about by a global mechanism.

4. Model of inhibitory connections in the cyclopean domain

Based on our findings in Experiments 1 , we offer a model of inhibitory interactions between binocular units, which may account for the farther-worse effect. In this section, we will first review the evidence of the inhibitory interactions between dissimilar disparities, and then turn to the design of the model. The model suggests two tests to distinguish between the hypothesis of competitive interactions and the OC hypothesis. In Experiments 5–6 we discover that these tests support the hypothesis of competitive interactions as an account of the farther-worse effect.

Inhibition in the disparity domain; evidence from psychophysics. The spaces between the dots of the transparent surface are perceived at the depth of the rear surface. By contrast, in random-dot stereograms of opaque surfaces, the background regions are typically perceived 'filled in' at the depth of the surface. Perhaps, this perceptual 'filling in' is inhibited in stereoscopic transparency in sparse RDS [7]. What happens at the boundaries between opaque surfaces, where sharply defined subjective contours are perceived? Akerstrom and Todd [7] hypothesized that:

- neurons tuned to similar disparities co-operatively interact over relatively large neighborhoods of visual space (global excitation associated with slowly changing depth),
- whereas competitive interactions between neurons tuned to different disparities are restricted to a much smaller vicinity (local inhibition over abrupt changes in disparity).

This scheme explains why locally detected disparity information can propagate into neighboring regions, and why this propagation is inhibited when two regions of sufficiently different disparity come in contact with one another: The lateral propagation of disparity information is inhibited both at the boundaries between opaque surfaces and in stereoscopic transparency.

In another study of interactions between disparities, the stimulus was a random-dot plane emerging from a dots cloud [8]. Inter-ocular correlation was used as a measure of strength of the stereoscopic signal. The strength of the signal was proportional to the square of dot contrast, which supports our observation in Experiment 4. The authors interpreted the influence of dot contrast on signal strength as evidence of inhibitory interactions between dissimilar disparities. Similarly, inter-ocular correlation may be used as measure of signal strength, and weight, of each surface in inter-surface inhibitory interactions in transparency.

In a study attempting to reveal binocular disparity channels, the disparity-tuning functions of the channels had opponent-type profiles [9], which also implies inhibition between dissimilar disparities.

Inhibition in the disparity domain; evidence from physiology. A convenient way to visualize the operation of global stereopsis is to use the metaphor of disparity map (or projection field; [23]). In the disparity map, each element of the binocular visual field is associated with a location in the disparity map through a pair of retinal receptive fields (Fig. 9). Each intersection of projection lines in the disparity map represents binocular units activated by the currently viewed stimulus. Note that the depth layers ranging from 'far' to 'near' (as shown in Fig. 9) are not anatomically segregated to layers in the visual cortex, rather, the segregation is provided by the connectivity between binocular neurons tuned to different disparities.

The notion of disparity map provides a convenient framework to illustrate the known properties of binocular neurons (binocular units). The typical profiles of neural responses (Fig. 9, insets) were obtained in experiments with extracellular recording from neurons of

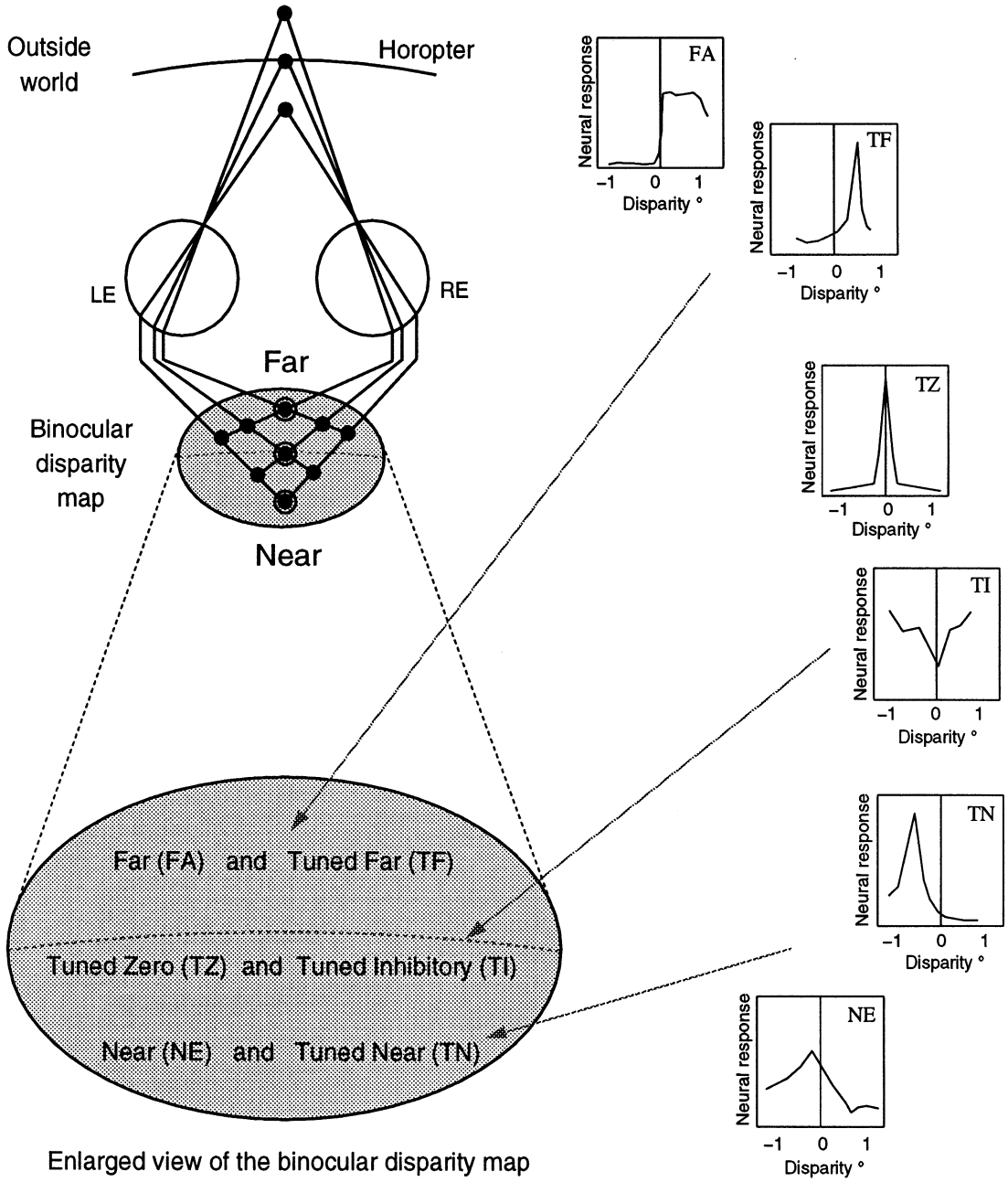

Fig. 9. Binocular disparity map and response profiles of binocular neurons. Each position in the binocular visual field is associated with a particular location in the disparity map, through a pair of receptive fields. The grey ellipses represent a horizontal sections of the disparity map. Large open circles in the disparity map (top) represent correct matches. Relative arrangement of six types of binocular neurons are roughly shown in the enlarged view of the disparity map (bottom). In the insets, the typical response profiles of these cells [25] are schematically plotted as a function of horizontal disparity.

striate (area V1) and prestriate (areas V2, V3, and V3a) cortex of awake, visually attentive, behaving monkeys [24,25]. Three general types of binocular neurons were evident with both solid figure and random-dot stereograms as stimuli [10]:

- excitatory and inhibitory neurons tuned to about zero (TZ, excitatory and TI, inhibitory), crossed (TN, tuned near), and uncrossed (TF, tuned far) disparities,
- reciprocal neurons responding (a) by excitation to a range of crossed disparities and by inhibition to a range of uncrossed disparities (NE, near neurons), and (b) in reversed manner (FA, far neurons),
- flat neurons that are not sensitive to a specific horizontal disparity (not shown in Fig. 9).

All types of stereoscopic neurons have been found in the explored regions of the visual cortex. Tuned excitatory neurons (TZ, TN, TF) were, on average, twice as

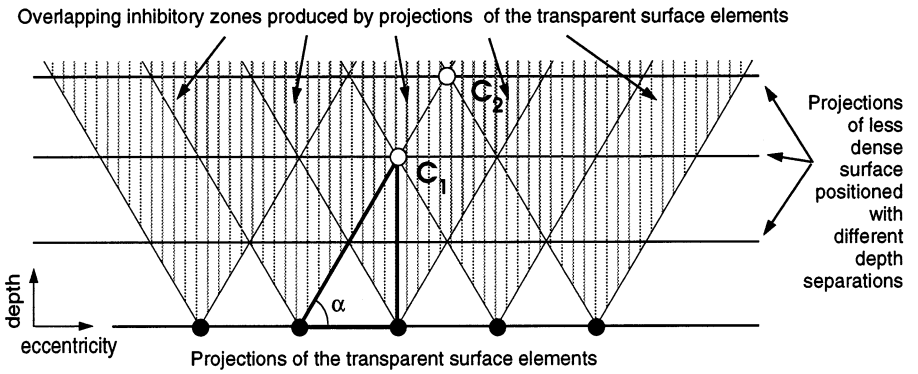

Fig. 10. Scheme of binocular projection field in stereoscopic transparency. The open and filled circles represent binocular units activated by elements of the target and the transparent surfaces, respectively. The activated units inhibit other units tuned to dissimilar disparities so that the resulting inhibitory (suppression) zones have a cone-like profile in the binocular disparity map: The larger the difference in binocular disparities the wider the inhibitory zone. Projection of the target surface will be more immersed into the overlapping inhibitory zones at larger depth separations (δ). Let $\nabla = \tan \alpha$ stand for the slope of the suppression zone boundary, and C_1 for the units activated by the rear surface positioned at depth i. The depth of C_1 is the minimal distance, at which all the rear surface projections will fall into at least two inhibitory zones and will be fully inhibited if the inhibitory weights $W = 1/2$. In this situation, $\nabla = \delta/\overline{S}_p = \delta/(2(\overline{S}_p/2))$. If $W = 1/3$, all the elements at the minimal depth of C_2 will be fully inhibited, and $\nabla = \delta/3(\overline{S}_p/2)$. By induction, $\nabla = \delta/(1/W)(\overline{S}_p/2) = 2W(\delta/\overline{S}_p)$. Thus, one can compute the limiting density ρ_p^*, at each δ, where all the target elements are inhibited.

numerous as reciprocal neurons (NE, FA). Response profiles obtained with isolated bars (solid figure stereograms) and random-dot stimuli were qualitatively the same, but the magnitude of the response and the range of effective disparities might be different [10].

To summarize, there is psychophysical and physiological evidence for the existence of inhibitory interactions in the cyclopean domain. The connections which mediate these influences might constitute the neural substrate for the farther-worse effect. We did not find any physiological evidence regarding the lateral extent of the inhibitory influences; the main aspect of the connectivity between cortical binocular neurons we discuss below.

The fact that the inhibitory influences are found in all the studied visual areas (V1–V3a) suggests that global disparity processing occurs at several levels, probably including those where results of disparity processing from successive depth fixations integrate. This is consistent with the view that several levels of global disparity processing exist in the cyclopean domain, prior to elaboration of a 'cleaned' cyclopean image [20], where matching noise is entirely eliminated.

Disparity gradient limit. Gradients of disparity steeper than a particular value cannot be processed stereoscopically [6]. This limitation, called disparity gradient limit allows large variations in depth if they occur to a sufficiently large spatial extent. There has been a debate in the literature whether disparity gradient limit is a global or a local phenomenon (see the review by Tyler [20]). Both psychophysical and physiological evidence suggests that disparity gradient is constrained by global stereopsis; that is it depends on interactions between local disparity detectors.

4.1. Design of the inhibitory connections

We offer a model of inhibitory connections between dissimilar disparities that agrees with the concept of global disparity gradient limit, and may account for our findings. In the model, each activated disparity detector, or a pool of detectors (binocular unit), inhibits binocular units tuned to certain disparities so that the resultant inhibitory (suppression) zone has a coneshaped profile in the disparity map: The inhibitory zones grow wider as the depth increases (Fig. 10). The scheme predicts that an increasingly large part of the rear surface will enter the overlapping zones of suppression, as the depth separation (δ) between the rear and the transparent surfaces grows. Alternatively, at any given δ, an increasing dot density of the transparent plane (ρ_p) will cast more suppression on the rear surface.

Now we will introduce several simplifying assumptions and simulate operation of the model to show that it predicts the farther-worse effect in resolution of stereoscopic transparency. The assumptions are:

- an activated binocular unit will be inhibited to its background activity level and will not contribute to further disparity processing, if its activity decreases by 1,
- veridical perception occurs until most elements of the less dense surface are inhibited to their background activity level,
- the weights of inhibitory connections are equal and have a magnitude $W < 1$,
- the boundary of the suppression zone has a constant slope (the oblique line in Fig. 10).

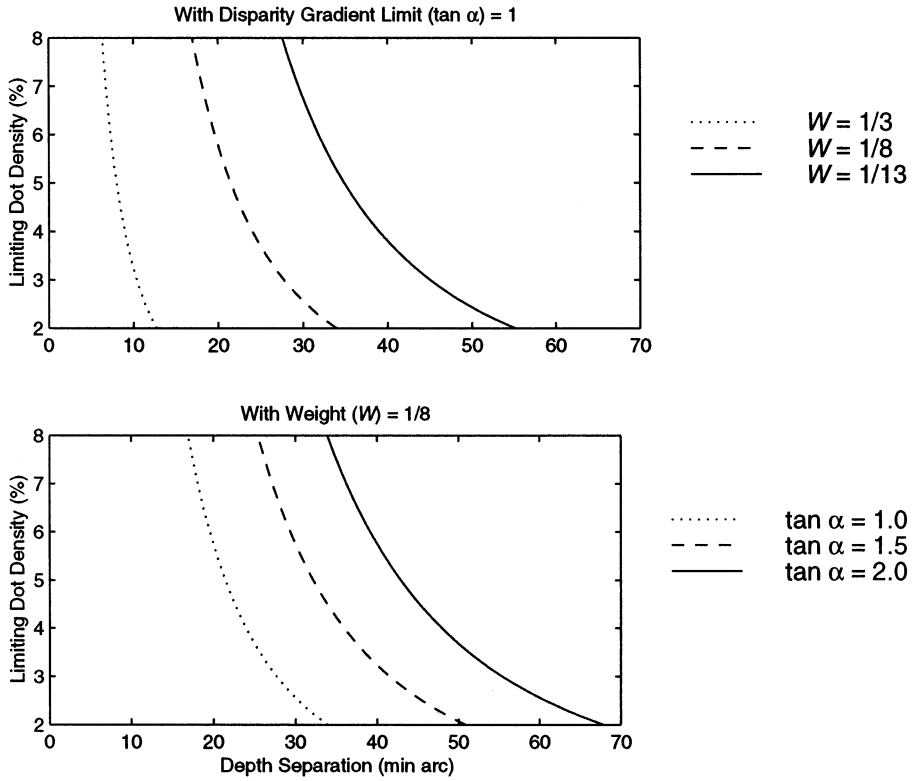

Fig. 11. Simulations of the relationship between the limiting density of the transparent plane and the inter-surfaces depth separation, with various weights of inhibitory connections (top) and slopes of suppression zone boundaries (bottom).

Consider a visual scene with two overlapping surfaces P (for plane) and C (for cylinder), where the dot density of P is higher than of C. Let P' and C' stand for the binocular units activated by P and C dots. The activity of a C', which occurred in the overlapping inhibitory zones of N units P', will decrease by $N \times W$. Let $W = 1/K$, where $K > 1$. N units P' will suppress one C' unit to its background activity level if $N = K$. Note that the C' unit will also inhibit several P's, each by $W = 1/K$, but this inhibition becomes negligible when P has a sufficiently higher dot density than C ($\rho_p \gg \rho_c$).

Now we can derive an expression to relate ρ_p^* to δ. We begin from the observation that each C' will be inhibited by 1 if it falls within N suppression zones, where the weights of the connections are equal to $1/N$. Thus, the degree of inhibition at certain δ depends on ρ_p; this point provides the desired relation between δ and ρ_p. Let $\overline{S}_p = A_p/n_p$ stand for the average inter-dot distance of P, where A_p and n_p represent the area and the number of dots of the plane, respectively. As should be clear from Fig. 10, the minimal depth separation (δ^*) at which inhibition is sufficient to interfere with veridical perception of the cylinder becomes: $\delta^* = \nabla(\overline{S}_p/2W)$, where $\nabla = \tan\alpha$ stands for disparity gradient limit.

4.2. Results of the simulation

The profile of the suppression zone agrees with the idea of disparity gradient limit. We start our simulation from the slope of the suppression zone boundary equal to 1, which is the magnitude of a disparity gradient limit measured psychophysically [26,6].[9]

In Fig. 11 (top), we show the results of the simulated relationship between ρ_p^* and δ with three different weights of the inhibitory connections. We chose the weights to keep the range of simulated δ in agreement with the results.[10] As clear from the simulation, the

[9] AIternatively, in the PMF algorithm, the magnitude of the disparity gradient limit was a trade-off between the disambiguation power and the ability to deal with as wide a class of scene structures as possible. Pollard and Frisby [28] examined how the PMF deals with stereograms of transparent surfaces used by Weinshall [11] and found that with various RDS densities (in Weinshall's stereograms, the dot density of all participating surfaces were coupled) the algorithm could be tuned to a better performance with different values of the disparity gradient limit (see below).

[10] To compute δ we subtracted half of the maximal disparity of the cylindrical surface from the disparity of the cylinder pedestal. High reproducibility of our results suggests that the observers used large and presumably the same regions of the target cylinder to identify its orientation. For simplicity, we assumed that the test stimuli had a single δ.

lower weights of inhibitory connections permit veridical perception over larger ‘, i.e. lower weights facilitate resolution of transparency.

The effect of variable inhibitory weights may explain our findings in Experiment 4, where we showed that ρ_p^* is sensitive to the contrast of the transparent surface: It is generally assumed that binocular units perform local cross-correlation at different locations in the binocular visual field e.g. [27]. Accordingly, different contrasts of binocular objects activate binocular units to various extents, which in turn may influence the contribution of these units to global disparity processing, i.e. by changing either the weights of inhibitory connections or the lateral extent of the suppression zones. Above, we have shown how the weights of inhibitory connections affect performance. What about the slope of the suppression zone boundary? Fig. 11 (bottom) illustrates the simulated relationships between ρ_p^* and δ with different slopes of boundaries of the inhibitory zones: the steeper disparity gradient limit the broader the range of δ at which transparency can be resolved. This conclusion is consistent with our expectations, because low disparity gradient limits strengthen smooth surfaces and suppress perception of surfaces varying rapidly in depth. Note that we used the magnitude of limiting disparity gradient reported by Tyler [26] and Burt and Julesz [6] with respect to fusional limits. Limiting disparity gradient may be different in global stereopsis, or the limiting value of the disparity gradient may depend on the number of involved disparate stimuli [28].

4.3. Discussion

So far, we have focused on the algorithms which solve the correspondence problem using the local criterion of disparity gradient. Another class of algorithms, based on the Bayesian approach, assumes that several smooth surfaces may exist in the same visual direction [29,30]. These algorithms compute and minimize the ‘cost functions’, which weigh alternative solutions of the stimulus. Resolution of transparent surfaces is limited by employing in the cost function a ‘line process field’, which represents depth discontinuities, where disparity changes too abruptly. For example, in the global algorithm of Madarasmi et al. [30] the activity at each depth layer is weighed by the number of units activated at that depth, a feature that may provide a link between the dot densities of the participating surfaces and the ability to resolve transparency. However, it is unclear without simulation whether the model will exhibit worse performance at larger δ.

Another interesting approach to global stereopsis is the stereo-matching neural network model of Marshall et al. [31]. The model incorporates ‘surface patch’ binocular units which

- are sensitive to small pieces of surfaces (coherent groups of close binocular disparities),
- project reciprocal excitatory connections to the first-stage binocular units.

Due to these back-projections, the model occupies an intermediate position on the local-to-global matching scale. It would be very interesting to see how our results could be incorporated into the theories described above.

Global limitations may operate both at the level of feature-specific binocular matching and when these individual reconstructions integrate. Our scheme may be applied at both global stages, i.e. through parsimonious recurrent inputs. Alternatively, the further-worse effect may be attributed only to a separate level of hyperglobal integration, suggested in Experiment 3. In this case, binocular matching processes would proceed separately at each surface and be followed by a hyperglobal reconstruction process characterized by the farther-worse effect. Now, the participating surfaces have already been ‘cleaned’ in earlier global computations and the inhibitory interactions may serve to eliminate the spurious surfaces interpolated between those false matches which survived at earlier levels (cf. the conjugacy principle in [20]).

The inhibitory interactions at a separate level of hyperglobal integration will bring to a ‘rarefication’ of all participating surfaces, which is consistent with the known parsimony and interpolating ability of vision, when a limited number of disparate elements is sufficient to produce a strong percept of a surface.

To summarize, we proposed a model of inhibitory connections between dissimilar disparities which may be verified using physiological methods [10] and tested by extending the existing stereo-algorithms. Global stereopsis also involves co-operative (excitatory) interactions between disparities satisfying the disparity gradient requirement (as in the PMF and Prazdny’s algorithms). Here, we have not addressed the excitatory aspect of connectivity between binocular units to emphasize that the existence of competitive interactions is compatible with the perception of stereoscopic transparency.

5. Experiments 5–6

5.1. Experiment 5: linear relationship between limiting density and dot density of the rear surface

In discussion section of Experiment 1 we showed that the ordering constraint (OC) produces matching noise which hampers performance in a manner similar to our farther-worse effect. The OC hypothesis predicts that dot density increment of either the transparent plane or the target surface will hurt performance. On the other

hand, our model of inhibitory connections between dissimilar disparities predicts that performance will improve with higher dot densities of the target surface (ρ_c): an increment of ρ_c will cause a fraction of the rear surface dots to fall into those locations in the disparity map where the overlap between the suppression zones is not sufficient for inhibition. As a result, the observers will be able to identify cylinder orientation with higher limiting densities of the transparent plane (ρ_p^*).

In the present experiment we pit the hypothesis of competitive interactions against the OC hypothesis.

5.1.1. Method

We kept δ constant (28′ arc for observer MP and 35′ arc for observer MZ) and measured ρ_p^* with various ρ_c. The fixation pattern was presented at the depth of the cylinder pedestal (34′ arc) for the observer MP, and at zero depth (the depth of the transparent plane of the following test stimulus) for the observer MZ.

5.1.2. Results and discussion

The results for two observers show that ρ_p^* directly depends on ρ_c (Fig. 12). This finding rules out the OC hypothesis as an account of the farther-worse effect, and agrees with the qualitative predictions of the hypothesis of competitive interactions and of our scheme of inhibitory connections.

The results of Experiments 4–5 are supported by a study which investigated the depth averaging mechanism involved in vergence control [14]. In that work, the pulling strength of each depth plane was considered as a weight to the averaged signal power of two depth planes depicted in RDS. The signal power was directly proportional to:

- the number of dots in the depth plane, in agreement with results of the present experiment,
- the square of the dot contrast, in agreement with results of Experiment 4.

The agreement between the results further supports the view that disparity representations for depth perception and vergence control are shared (see discussion in [14]).

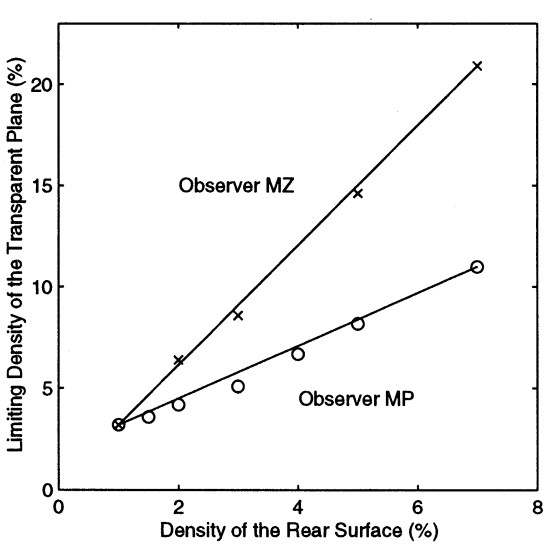

Fig. 12. Results of Experiment 5 for two observers. Limiting dot density (ρ_p^*) of the transparent plane is plotted as a function of the dot density of the target (ρ_c). We drew the straight lines through the extreme data points to emphasize the linear character of the relationship.

5.2. Experiment 6: near–far asymmetry in stereoscopic hyperglobality

The OC hypothesis predicts that the depth order of the target and the masking surfaces will not affect the farther-worse effect, as long as the dot densities of the surfaces do not change. The hypothesis of competitive interactions allows different inhibitory weights, for example, in the domains of crossed and uncrossed disparities (see discussion below). Also, the view of stereoscopic transparency perception as a hyperglobal process suggests that a surface with a higher dot density (a masking surface) will hamper discrimination of another surface (a target), even when the target is closer to the observer. In other words, we expect that the farther-worse effect will persist in a visual configuration with overlapping surfaces, when the target is closer to the observer than the masking surface. These conditions transcend stereoscopic transparency: Now the task does not require the perception of the rear surface.

5.2.1. Method

To explore this possibility we reversed the depth order of the two surfaces the target cylinder and the masking plane in the stereograms otherwise identical to the test stimuli in Experiment 1. We measured the limiting dot density of the plane (ρ_p^*) with different negative depth separations (δ) between the surfaces. The fixation pattern was presented at zero depth, which was the depth of the masking plane of the test stimulus. This procedure is identical to Experiment 1 with negative δ.

5.2.2. Results

The observers reported that discrimination of the cylinder orientation was more difficult than in Experiment 1. Now, the range of used negative depth separations was limited by (pykno-stereoscopic) disparity averaging at small depth separations and increasing difficulty of the task at high depth separations.

The results for two observers are plotted with open circles in Fig. 13. Cross symbols represent results from Experiment 1 with positive δ. The dashed line illustrates the prediction of perfect symmetry of the farther-worse effect relative to the imaginary plane of zero δ (dotted line). The results show that:

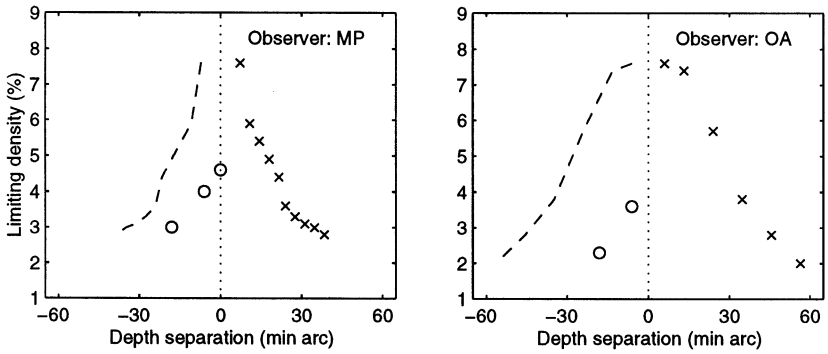

Fig. 13. Results of Experiment 6 for two observers. Limiting dot density of the transparent plane (ρ_p^*) is plotted with open circles as a function of depth separation between the surfaces (δ). Results of Experiment 1 with positive depth separations are replotted with crossed symbols. If the farther-worse effect were completely symmetrical relative to zero depth separation (dotted line), then the limiting densities obtained with negative depth separations would fit the dashed lines.

- The farther-worse effect persists when a surface with a higher dot density is positioned behind the target and is irrelevant to the task. The further-worse effect then transcends the conditions of stereoscopic transparency and may be thought of as a signature of hyperglobal disparity processing,
- The farther-worse effect with target closer to the observer occurs in a narrower range of δ than when the target is positioned behind the masking surface. We refer to this observation as a near–far asymmetry in the perception of hyperglobal stereoscopic stimuli. near–far asymmetry supports the idea that binocular disparity is processed in separate channels (see below),
- The near–far asymmetry is inconsistent with the prediction of the OC hypothesis.

5.2.3. Discussion

Based on studies of anomalous binocular vision, Richards proposed a pool hypothesis according to which the human visual system contains three pools of disparity sensitive neurons: one for near-zero disparities, one for crossed disparities, and one for uncrossed disparities [32]. Alternatively, psychophysical evidence suggests a continuum of overlapping depth channels, more sharply tuned to binocular disparity [9].

The near–far asymmetry may be explained in light of the channel-wise organization of stereopsis: the inhibitory connections between (or within) different pools of binocular cortical neurons may have a common architecture, while particular characteristics of the connectivity e.g. the weights of the connections are different across the pools. This difference could result from a (learned) strategy to fixate the visual objects so as to bring more details within the region of uncrossed disparities.

What is a possible neural substrate of the hyperglobal interactions? We have reported above that the resolution of stereoscopic transparency is more difficult with test durations below 220 ms, which suggests that the visual system 'prefers' to fixate at each surface separately before it relates the surfaces to each other and reconstructs the whole scene. On the other hand, the farther-worse effect is independent of fixation depth, as Experiment 2 showed. The evidence from Experiment 2 agrees with our conclusion in Experiment 3 that the separate reconstructions of the overlapping surfaces interact at a functional level above binocular matching at each surface. Perhaps, the individual surface fixations are required for the optimal matching at each surface, a process separate from a the higher-level interaction between the surface representations. The vergence-independent farther-worse effect requires binocular cells which are not sensitive to particular disparities. A class of binocular flat neurons [24] was found to respond equally strongly at all examined disparities. We suggest that an ensemble of the flat neurons makes a good candidate for an early neural substrate of the farther-worse effect.

The requirement of integration between separate depth 'snapshots' is not unique to stereoscopic transparency. The oculo-motor system constantly initiates vergence eye movements to compensate the imprecision of fixation (about 3′ arc) and the depth shifts produced by head movements. Under these conditions, the visual system needs a (presumably early) mechanism to combine the depths measures from the same objects, taken during different fixations in depth [33].

6. Conclusions

Our findings support and add to the growing evidence on inhibitory interactions in the disparity domain.

1. Using cyclopean stimuli depicted in sparse random-dot stereograms, we report a stable vergence-independent relationship between the depth separation and

the dot density (farther-worse effect) of two overlapping surfaces, one of which is perceived as transparent (Experiments 1–2).

2. The farther-worse effect persists:

1. when the elements of the participating surfaces are of the opposite contrast polarities (Experiment 3),
2. and in a purely hyperglobal task, where the target surface was positioned in front of the suppressing (masking) surface (Experiment 6).

3. The results with opposite contrast polarities (2a) show that the basic phenomenon persists when binocular matching at each surface is separated to independent neural sites in the visual pathway [17]. This suggests the independent processes of

1. binocular matching at each surface,
2. inter-surface suppressing interactions.

4. The sensitivity of the farther-worse effect to absolute luminance difference between elements of participating surfaces is well approximated by a quadratic relationship, which provides a link with the studies addressing the concept of signal strength in the disparity domain [8,14].

5. The farther-worse effect cannot be accounted for by matching noise generated due to operation of either the continuity or ordering constraint (Experiments 1–6).

6. We propose a scheme of inhibitory interactions between (pools of) cortical binocular neurons that is based on the concept of disparity gradient limit, which was first characterized psychophysically and found to be computationally advantageous in resolving stereoscopic transparency [4,5]. We simulate performance of the scheme and find that it agrees with psychophysical observations (Experiments 1–3–6).

7. Although based on competitive (inhibitory) interactions between dissimilar disparities, the scheme permits veridical perception of the scenes with overlapping surfaces in the same visual direction.

Acknowledgements

We are grateful to Dov Sagi for his continuous support throughout the study, to Yoram Bonneh for helpful advice on computer graphics, to Michael Kubovy for his substantial contribution to clarity of this article, and to two anonymous referees for providing helpful comments about the manuscript. The research was supported in part by The Ministry of Science and Technology and The Ministry of Absorption of Israel to A.C., and the Feinberg Graduate School of The Weizmann Institute of Science to S.G.

References

[1] Julesz B. Binocular depth perception of computer-generated patterns. Bell Syst Tech J 1960;39:1125–62.
[2] Marr D, Poggio T. Cooperative computation of stereo disparity. Science 1976;194:283–7.
[3] Grimson WEL. Computational experiments with a feature based stereo algorithm. IEEE Trans Pattern Anal Mach Intell 1985;1:17–34.
[4] Pollard SB, Mayhew JEW, Frisby JP. PMF: A stereo correspondence algorithm using a disparity gradient limit. Perception 1985;14:449–70.
[5] Prazdny K. Detection of binocular disparities. Biol Cyber 1985;52:93–9.
[6] Burt P, Julesz B. A disparity gradient limit for binocular fusion. Science 1980;208:615–7.
[7] Akerstrom RA, Todd JT. The perception of stereoscopic transparency. Percept Psychophys 1988;44:421–32.
[8] Cormack LK, Stevenson SB, Schor CM. Interocular correlation, luminance contrast and cyclopean processing. Vis Res 1991;31:2195–207.
[9] Cormack LK, Stevenson SB, Schor CM. Disparity-tuned channels of the human visual system. Vis Neurosci 1993;10:585–96.
[10] Poggio GF. Stereoscopic processing in monkey visual cortex: a review. In: Papathomas TV, Chubb C, Gorea A, Kowler E, editors. Early Vision and Beyond. Cambridge: MIT Press, 1995:43–55.
[11] Weinshall D. Perception of multiple transparent depth planes in stereo vision. Nature 1989;307:737–9.
[12] Weinshall D. Seeing 'ghost' planes in stereo vision. Vis Res 1991;31:1731–48.
[13] Rashbass C, Westheimer G. Disjunctive eye movements. J Physiol 1961;159:339–60.
[14] Mallot HA, Roll A, Arndt PA. Disparity-evoked vergence is directed towards average depth. Tech Rep 019, Max Planck Institut fur Biologische Kybernetik, Tubingen, Germany,(1995).
[15] Harris JM, Parker AJ. Efficiency of stereopsis in random-dot stereograms. J Optic Soc Am A 1992;9(1):14–24.
[16] Schiller PH. The on and off channels of the visual system. Trends Neurosci 1992;15:86–92.
[17] Harris JM, Parker AJ. Independent neural mechanisms for bright and dark information in binocular stereopsis. Nature 1995;374:808–11.
[18] Erkelens CJ. Adaptation of ocular vergence to stimulation with large disparities. Exp Brain Res 1987;66:507–16.
[19] Erkelens CJ. Fusional limits for a large random-dot stereogram. Vis Res 1988;28:345–53.
[20] Tyler CW. Cyclopean vision. In: Regan D, editor. Binocular Vision. London: Macmillian Press, 1991.
[21] Parker AJ, Yang Y. Spatial properties of disparity pooling in human stereo vision. Vis Res 1989;29:1525–38.
[22] Saye A, Frisby JP. The role of monocularly conspicuous feature in facilitating stereopsis from random-dot stereograms. Perception 1975;4:159–71.
[23] Boring EG. The Physical Dimensions of Consciousness. New York: The Century Co., 1933.
[24] Poggio GF, Gonzalez F, Krause F. Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. J Neurosci 1988;8:4531–50.
[25] Poggio GF. Physiological basis of stereoscopic vision. In: Regan D, editor. Binocular Vision. London: Macmillian Press, 1991.
[26] Tyler CW. Depth perception in disparity gratings. Nature 1974;251:140–2.
[27] Tyler CW. Cyclopean riches: co-operativity, neurotropy, hysteresis, stereoattention, hyperglobality, and hypercyclopean processes in random-dot stereopsis. In: Papathomas TV, Chubb C,

Gorea A, Kowler E, editors. Early Vision and Beyond. Cambridge: MIT Press, 1995.
[28] Pollard SB, Frisby JP. Transparency and the uniqueness constraint in human and computer stereo vision. Nature 1990;347:553–6.
[29] Clark JJ, Yuille AL. Data Fusion for Sensory Information Processing Systems. Boston: Kluwer, 1990.
[30] Madarasmi S, Kersten D, Pong TC. The computation of stereo disparity for transparent and for opaque surfaces. In: Giles CL, Hanson SJ, Cowan JD, editors. Advances in Neural Information Processing Systems 5. San Mateo, CA: Morgan Kaufmann, 1993.
[31] Marshall JA, Kalarickal GJ, Graves EB. Neural model of visual stereomatching: Slant, transparency, and clouds. Network: Computation in Neural Systems 1996;7:635–69.
[32] Richards W. Anomalous stereoscopic depth perception. J Optic Soc Am 1971;61:610–4.
[33] Collewijn H, Steinman RM, Erkelens CF, Regan D. Binocular fusion, stereopsis and stereoacuity with a moving head. In: Regan D, editor. Binocular Vision. London: Macmillian Press, 1991.
[34] Anderson BL, Nakayama K. Toward a general theory of stereopsis: binocular matching, occluding contours, and fusion. Psychol Rev 1994;101:414–45.
[35] Braddick O. Multiple matching in stereopsis.(1978) (Not published). Report of work supported by NSF grant No 77-07569-MCS to Richards W and Marr D.
[36] Cavanagh P. Pathways in early vision. In: Pylyshyn ZW, editor. Computational Processes in Human Vision. Norwood, NJ: Ablex, 1988.
[37] Churchland PS, Sejnowski TJ. The Computational Brain. Cambridge: MIT, 1993.
[38] Jones J, Malik J. Determining three-dimensional shape from orientation and spatial frequency disparities. Proc ECCV, Santa Margherita Ligure, Italy, (1992) 661–9.
[39] Julesz B. Global stereopsis: co-operative phenomena in stereoscopic depth perception. In: Held R, Leibowitz H, Teiber HL, editors. Handbook of Sensory Physiology, vol. VII. Berlin: Springer, 1978.
[40] Julesz B. Dialogues on Perception. Cambridge: MIT, 1995.
[41] Marr D. Vision. New York: WH Freeman, 1982.
[42] Metelli F. The perception of transparency. Sci Am 1974;230:91–8.
[43] Ramachandran VS, Anstis SM. Perception organization in multistable apparent motion. Perception 1985;14:135–43.
[44] Wertheimer M. Principles of perceptual organization. (Originally published in German, 1923.) In: Ellis WD, editor. A Source Book of Gestalt Pschology. London, Routledge and Kegan Paul, 1955.