Atten Percept Psychophys (2015) 77:190–206
DOI 10.3758/s13414-014-0746-8

Magnitude of perceived depth of multiple stereo transparent surfaces

Saori Aida · Koichi Shimono · Wa James Tam

Published online: 14 August 2014

Abstract According to the geometric relational expression of binocular stereopsis, for a given viewing distance the magnitude of the perceived depth of objects would be the same, as long as the disparity magnitudes were the same. However, we found that this is not necessarily the case for random-dot stereograms that depict parallel, overlapping, transparent stereoscopic surfaces (POTS). The data from five experiments indicated that (1) the magnitude of perceived depth between the two outer surfaces of a three- or a four-POTS configuration can be smaller than that for an identical pair of stereo surfaces of a two-POTS configuration for the range of disparities that we used (5.2–19.4 arcmin); (2) this phenomenon can be observed irrespective of the total dot density of a POTS configuration, at least for the range that we used (1.1–3.3 dots/deg^2); and (3) the magnitude of perceived depth between the two outer surfaces of a POTS configuration can be reduced as the total number of stereo surfaces is increased, up to four surfaces. We explained these results in terms of a higher-order process or processes, with an output representing perceived depth magnitude, which is weakened when the number of its surfaces is increased.

Keywords Stereo transparent surface · Perceived depth magnitude · Depth discrimination · Dot density · Higher-order process · Cross-correlation analysis

When a human observer binocularly views a three-dimensional object, images of the object are formed on the retinae, which differ in their positions along the horizontal direction because of the horizontal alignment of the eyes. This difference of the retinal images is referred to as *horizontal binocular disparity*, which is a known cue for depth perception—that is, binocular stereopsis. Binocular stereopsis was discovered in the early 1830s by Wheatstone, who invented a mirror stereoscope to present a disparate picture to each eye of a single viewer (e.g., Wade & Ono, 2012). Since the discovery of binocular stereopsis, much effort has been devoted to understand its basic mechanism in different fields of research, such as psychology, physiology, and computer science (e.g., Howard & Rogers, 2002, 2012). This may be because the function of binocular stereopsis—to localize objects in depth in a three-dimensional visual space—is one of the major functions of the visual system.

Although it is well established that disparity is a depth cue, it has been an open question whether or not the visual system can utilize it to localize a three-dimensional object veridically (e.g., Glennerster, Rogers, & Bradshaw, 1996; Howard & Rogers, 2002, 2012; Nakamizo & Shimono, 2001). In binocular stereopsis, the geometric relation of binocular disparity (δ), viewing distance (D), interocular distance (I), and depth magnitude (d) between objects can be approximately expressed as follows (for a derivation, see Ono & Comerford, 1977):

$$d \fallingdotseq \frac{\delta D^2}{I}. \qquad (1)$$

However, Eq. 1 indicates the geometrical relation among these variables and does not indicate whether the human visual system actually "calculates" the depth as predicted by the equation. In the literature, studies have indicated that the magnitude of perceived depth may covary well with the disparity value when the viewing distance is constant (e.g., Richards, 1971; Ritter, 1977, 1979; Shimono, Tam, Stelmach,

S. Aida (✉) · K. Shimono
Graduate School of Marine Science and Technology, Tokyo University of Marine Science and Technology, Tokyo, Japan
e-mail: saori.aida.t@gmail.com

W. J. Tam
School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Ontario, Canada

& Hildreth, 2002), and with the viewing distance when the disparity value is constant and relatively small (e.g., Glennerster et al., 1996; Johnston, 1991; Ritter, 1977; Shimono & Nakamizo, 1990). These studies suggest that the visual system may take into account the disparity value and the viewing distance to "calculate" the magnitude of perceived depth as predicted by Eq. 1, although the magnitude of perceived depth can be affected by the stimulus properties or viewing conditions (see Howard & Rogers, 2002, 2012).

Recently, Aida and Shimono (2010) reported a phenomenon indicating that even when the values of δ and D are constant, the magnitudes of perceived depth can be different. This is not expected from Eq. 1. The phenomenon can be observed for random-dot stereograms (RDS) that depict two parallel, overlapping, transparent stereoscopic surfaces (POTS); the perception observed with a RDS of POTS is that human observers can perceive multiple surfaces simultaneously at different depths in the same visual direction (i.e., stereo transparency), and the magnitudes of perceived depth for the closest and farthest surfaces of the POTS depend on the number of surfaces in the set of POTS.

Stereo transparency or perception obtained with the POTS configuration is interesting from an ecological point of view, because a mechanism mediating stereo transparency may be closely related to that operating for a natural scene, in which human observers view objects, such as tree trunks, branches, and leaves that are in front and those (at different depths) through "transparent" gaps among them (Tsirlin, Allison, & Wilcox, 2008). Because it is a matter of life or death to be able to find food or escape from a predator in the three-dimensional scene, it is reasonable to assume that a mechanism to detect or get information about an object(s) hiding in the scene may have developed in the course of evolution. Thus, by examining stereo transparency, more light can be thrown on this mechanism.

Most studies concerning stereo transparency have examined which properties of a POTS configuration can produce perceived transparency (e.g., Akerstrom & Todd, 1988; Anderson, 1992; Gepshtein & Cooperman, 1998; McKee & Verghese, 2002; Stevenson, Cormack, & Schor, 1989, 1991; Tsirlin, Allison, & Wilcox, 2008, 2012; Wallace & Mamassian, 2004). Stereo transparency can be obtained with POTS configurations having binocular disparity within a range of disparities from 3 to 30 arcmin (Stevenson et al., 1989) and with POTS having up to six different depth surfaces for a given binocular disparity (Tsirlin et al., 2008). The number of dots presented in the depth surface (Akerstrom & Todd, 1988; Tsirlin et al., 2008, 2012; Wallace & Mamassian, 2004), the color of dots in the depth surface (Akerstrom & Todd, 1988), and the stimulus luminance (Gepshtein & Cooperman, 1998) are also known factors that can affect perceived stereo transparency. Moreover, although fusion is thought to break down when the binocular disparity gradient is over unity (Burt & Julesz, 1980), McKee and Verghese (2002) have reported that stereo transparency can be obtained even when the binocular disparity gradient between two dot elements to be fused (i.e., their binocular disparity divided by their angular separation) on different depth surfaces is beyond unity. Moreover, Aida and Shimono (2010) found that the magnitude of perceived depth (between the closest and farthest surfaces) for a POTS configuration consisting of three surfaces can be smaller than that for a POTS configuration consisting of only two surfaces, even when the disparities of the two configurations and the viewing distances are the same. This suggests that the perceived depth magnitude is affected by one or more stimulus properties that are specific to POTS.

The aim of the present study was to confirm the (depth reduction) phenomenon that Aida and Shimono (2010) reported, and to further examine the factor(s) that might play a role in the phenomenon. We believe that the phenomenon reflects a property of a mechanism mediating stereo transparency and that it is worth studying in more detail. In Experiments 1 and 2, observers were asked to perform a depth reproduction task and a depth discrimination task, respectively, for a two-POTS configuration and a three-POTS configuration with the same disparities for the closest and farthest surfaces, in order to show that the phenomenon is reproducible using different methods. In Experiments 3 and 4, observers performed a depth reproduction task using stimuli with different dot densities and with different numbers of surfaces, respectively, to examine the impact of each stimulus property on the phenomenon. In Experiment 5, observers performed a depth discrimination task in which the magnitude of perceived depth for a three-POTS configuration and that for a four-POTS configuration were compared, to examine the impact of the number of surfaces on the phenomenon further by using a method different from that of Experiment 4.

General method

Apparatus

The stimuli were generated with a ViSaGe (Cambridge Research Systems) controlled by a computer (Endeavour, MT7500, EPSON), and were displayed on a screen (SP-100, IZUMI-COSMO) using a 3-D projector (DepthQ HDs3D-1 or DepthQ-WXGA, IT Co). The stimuli were viewed with LCD shutter glasses (60GX, NuVision) at a frame rate of 120 Hz. Due to limitations of access, the DepthQ HDs3D-1 3-D projector was used for Experiment 1 and for the smaller disparity condition of one of two disparity conditions in Experiment 5. The DepthQ-WXGA 3-D projector was used for Experiments 2, 3, and 4 as well as for the larger-disparity condition in Experiment 5. The 3-D projector was located behind the observer, with the center of the lens set at a height of

120 cm above the ground. The resolution of the 3-D projector was 1,264 × 632 pixels for the DepthQ-WXGA projector and 1,024 × 512 pixels for the DepthQ HDs3D-1 projector. Each stimulus consisted of a rectangular area and its size differed among the five experiments. The midpoint of the stimulus on the screen was positioned 157 cm above the ground. The observer sat in a chair and viewed a RDS depicting POTS. The height of the seat of the chair was 40 cm from the ground. The size of the screen was 199 × 149 cm and the screen was viewed from 380 cm, so that it subtended 29 × 22 arcdeg. During the data collection, all lights in the experimental room were switched off and the room was darkened expect for the dim illumination of the light from the monitor.

Stimuli

The stimuli used in the experiments were random-dot stereograms (RDS) that consisted of rectangular elements (see Fig. 1). The sizes of each rectangular element of the RDS were 2.6 × 2.6 arcmin for Experiment 1, 3.2 × 3.2 arcmin for Experiments 2, 3, and 4, and 2.6 × 2.6 arcmin for a small-disparity condition and 3.2 × 3.2 arcmin for a large-disparity condition in Experiment 5. The luminance of the rectangular element was 1.67 cd/m^2, and that of the background was 0.22 cd/m^2. Each RDS, when fused, depicted several overlaid planes of dots or POTS. Each of the stereo surfaces was a rectangular area (20×13 arcdeg in Exps. 1, 3, and 4, 20 × 10 arcdeg in Exp. 2 and for the larger-disparity condition in Exp. 5, or 20 × 11 arcdeg for the smaller-disparity condition in Exp. 5). The dot positions in each of the surfaces were randomly assigned. In this study, three configurations were used: a two-POTS configuration with two surfaces, a three-POTS configuration with three surfaces, and a four-POTS configuration with four surfaces. (Fig. 1b schematically illustrates the expected observer's percepts from the top view for the two-, three-, and four-POTS configurations.) The three-POTS configuration was made by adding a third surface with zero disparity between the two surfaces of a two-POTS configuration. The first surface (closest to an observer) was presented with crossed disparities with respect to the screen, and the end surface (farthest to an observer) was presented with uncrossed disparities with respect to the screen. The middle surface of the three-POTS configuration was presented at the screen distance (i.e., with zero disparity). The four-POTS

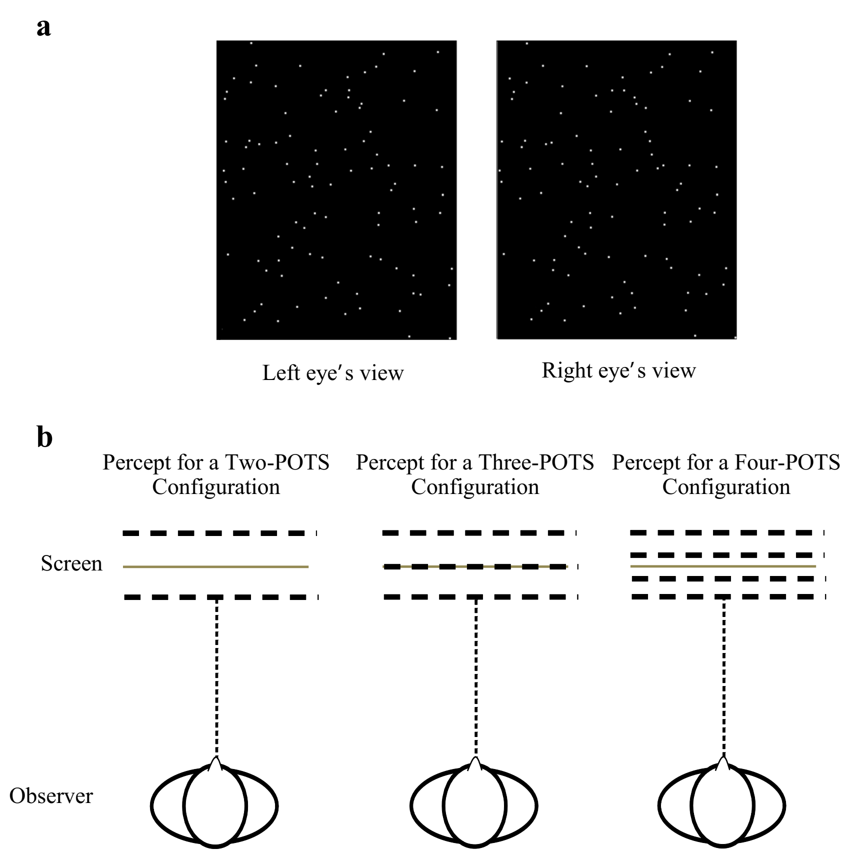

Fig. 1 (**a**) Sample random-dot stereogram (of a two-POTS configuration). When the stereogram is fused, two surfaces would be seen in depth and in the same visual direction. To provide a simple illustration, the aspect ratio and dot density are different from those of the actual stimuli used in this study. (**b**) Schematic illustration of observers' percepts from the top view for the two-, three-, and four-POTS configurations. For the two-, three-, and four-POTS configurations, observers would perceive two, three, and four stereo surfaces, respectively (dashed lines), in the same visual direction

configuration had four surfaces in which a two-POTS configuration was divided into three equal parts, so that the disparities between adjacent surfaces were equal. Note that throughout this article, the *disparity* of the POTS configuration refers to the disparity between its closest and its farthest surfaces. The total range of disparities for each POTS configuration used in the present study was between 3.2 and 22.6 arcmin, which is well within the fusional range of observers (Howard & Rogers, 2002, 2012) and within the disparity range in which stereo transparency can be observed (Stevenson et al., 1989). In the present study, dots could overlap between POTS configurations; however, the probability was very low for all of the conditions used (from .002 to .079), because of the relatively sparse dot densities that were used.

Experiment 1: Reproduction of perceived depth for a POTS configuration

In this experiment, we obtained perceived differences in depth by asking observers to reproduce the magnitude of perceived depth by estimating the depth between the closest and the farthest surfaces of a two-POTS configuration and a three-POTS configuration.

Method

Stimuli Each surface of the POTS contained horizontal disparities corresponding to a frontal surface, centered in the visual field. The disparity of the surfaces of the two-POTS configuration and that of the two outer surfaces of the three-POTS configuration were 5.2, 7.7, and 10.3 arcmin. The disparity pairs used to generate the stimuli were 2.6 and –2.6 arcmin, 3.9 and –3.9 arcmin, and 5.2 and –5.2 arcmin. A positive and a negative value represented a crossed and an uncrossed disparity, respectively, with respect to the screen surface. Each stereogram consisted of 600 dots (2.2 dots/deg^2). The *dot density* refers to the total dot density of the stimulus. In the two-POTS configuration, each surface consisted of 300 dots (1.1 dots/deg^2). In the three-POTS configuration, each surface consisted of 200 dots (0.7 dots/deg^2).

Procedure Participants performed in five experimental sessions, and the disparities and POTS configurations for each session were randomly selected from three different disparities and two different POTS configurations, respectively, with one repetition. Thus, each observer had 30 trials in total. Before an experimental session, each observer performed several training trials, which was randomly selected from trials in the experimental sessions, until the experimenter judged that observers understood the task. The observers' tasks were (1) to report the total number of surfaces in the RDS and (2) to reproduce with a tape measure the magnitude of the depth between the surfaces of the two-POTS configuration or between the two outer surfaces of the three-POTS configuration. We asked observers to report the number of stereo surfaces because it was a prerequisite in the present experiment that observers be able to perceive the stereo surface and its number "correctly," as depicted by the stimuli. The RDS was visible until the observers finished responding with their estimates.

In the training and experimental trials, the stimulus presentation time was unlimited, and observers could move their eyes freely. This stimulus presentation was arranged on the basis of our informal observations that (a) observers needed relatively long presentation times—more than a few seconds, at least—to perceive stereo transparency, and (b) when they fixated the center of a POTS configuration, they reported difficulty in perceiving clear stereo transparency in the peripheral visual field.

Observers Eight students from the university community, ranging in age from 19 to 23 years, participated. They had normal or corrected-to-normal visual acuity and good stereopsis. All of them were naive as to the purpose of the experiment.

Results and discussion

First, we examined whether the number of surfaces reported corresponded to that of the actual POTS configurations, and found that all observers reported the correct number of surfaces for the two- and three-POTS configurations in every trial. This suggests that they did observe stereo surfaces, as expected. Second, we transformed the reported value to a logarithmic depth value for every trial, because some observers assigned a relatively large number to the depth value. By this transformation, we attempted to make the distribution of the depth values normal. The transformed values for each disparity and each POTS condition were averaged for each observer, and the averaged values were used as "basic" scores for analysis. The validity of this procedure may be seen by referring to Fig. 2a, which shows the basic scores for the three-POTS configuration against those for the two-POTS configuration on linear–linear axes. As can be seen in Fig. 2a, there was no unusual or skewed distribution of the transformed mean values at each disparity condition.

We conducted a two-way repeated measures analysis of variance (ANOVA; 2 POTS × 3 disparities) on the basic score. We found that the main effects of POTS [$F(1, 7) = 14.96$, $p < .01$] and disparity [$F(2, 14) = 92.46$, $p < .01$] were statistically significant, whereas their interaction was not. Post hoc analyses (Tukey test) showed that the mean transformed values were significantly different between the two POTS configurations in each disparity condition.

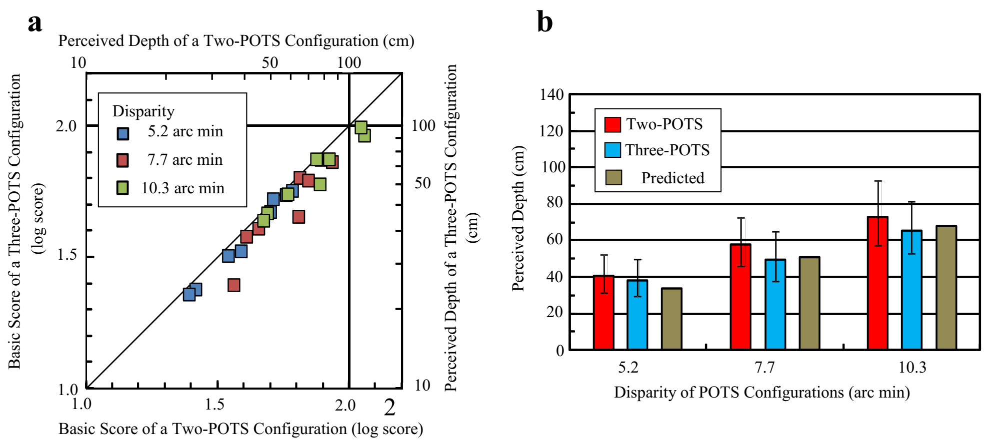

Fig. 2 Experiment 1. (**a**) Basic score of a three-POTS configuration as a function of that of a two-POTS configuration for each observer, as a function of disparity. The diagonal line is a reference line indicating the case when the score of the three-POTS configuration is the same as that of the two-POTS configuration. The right *y*-axis indicates the magnitude of perceived depth in centimeters, which was transformed back from each basic score. (**b**) Mean perceived depths of two- and three-POTS configurations and the depth magnitude predicted from Eq. 1, in centimeters, as a function of disparity. The mean for each disparity condition and each POTS configuration was calculated by averaging the basic scores over eight observers and then transforming the average back by using the equation mean = $10 \wedge (\sum_{n=1}^{8} \log_{10}(\text{Reported Depth})_n/8)$. The vertical lines attached to the bars indicate 95 % confidence intervals that were transformed back from those for the basic scores

The significant main effect of POTS can be seen in Fig. 2a, in which we drew a diagonal line where the scores for the two-POTS and the three-POTS configurations would be the same. Figure 2a shows that most scores fall below the diagonal, suggesting that the perceived depth for the three-POTS configuration is smaller than that for the two-POTS configuration. To estimate how much smaller the magnitude of perceived depth for the three-POTS configuration was than that for the two-POTS configuration, we averaged the scores among the eight observers for each condition, transformed back the averaged score, and plotted the transformed-back values as a function of disparity in Fig. 2b. Figure 2b shows that the mean transformed-back value or the mean magnitude of perceived depth (38.2 cm) for the three-POTS configuration was smaller than that (40.6 cm) for the two-POTS configuration in the small-disparity condition (5.2 arcmin); the mean magnitude of perceived depth (49.7 cm) for the three-POTS configuration was smaller than that (57.9 cm) for the two-POTS configuration in the middle-disparity condition (7.7 arcmin); and the mean magnitude of perceived depth (65.6 cm) for the three-POTS configuration was significantly different from that (72.9 cm) for the two-POTS configuration in the large-disparity condition (10.3 arcmin). The ratios of the mean perceived depth of the two-POTS configuration to that of the three-POTS configuration were 0.94, 0.86, and 0.90, for the small-, middle-, and large-disparity conditions, respectively. These results are consistent with the phenomenon reported by Aida and Shimono (2010) that the magnitude of perceived depth in the three-POTS configuration can be reduced as compared with that in the two-POTS configuration, even if they have the same disparity and are viewed at the same distance.

The significant main effect of disparity can be seen in Fig. 2b, in which the magnitude of perceived depth increases as a function of disparity for both the two- and three-POTS configurations. Furthermore, Fig. 2b shows that the magnitudes of perceived depth for the two POTS configurations correspond well with those predicted from Eq. 1. The predicted magnitudes from Eq. 1 were 33.8, 50.7, and 67.6 cm for the 5.2-, 7.7-, and l0.3-arcmin disparities, respectively. Follow-up *t*-tests performed on the differences between observers' basic scores and those logarithmically transformed from the values predicted by Eq. 1 indicated that the difference in every disparity condition was not statistically significant for both the two- and three-POTS configurations.

Why was there a difference in perceived depth between the two- and three-POTS configurations? One may think that a difference in stimulus properties between the POTS configurations could play a role in the perceived depth. The dot density for each surface in the two-POTS configuration (1.1 dots/deg^2) was somewhat larger than for the three-POTS configuration (0.7 dots/deg^2), whereas the total dot densities of the two configurations were the same (2.2 dots/deg^2). Wallace and Mamassian (2004) showed that as the dot density of a POTS configuration increases from 0.9 to 57.0 dots/deg^2, the efficiency (i.e., the ratio of human to ideal performance) in a depth discrimination task decreases. Furthermore, Akerstrom and Todd (1988) found that observers had difficulty reporting stereo transparency when the dot density was relatively high (10.4–31.3 dots/deg^2). However, it is not known whether or not dot density affect the magnitude of perceived depth. Thus, we examined the effect of dot density on perceived depth in Experiment 3.

Mean disparity gradients were also slightly different between the two- and three-POTS configurations. For the 5.2-arcmin disparity condition, the mean disparity gradients for the two- and three-POTS configurations were 0.34 and 0.28, respectively; for the 7.7-arcmin disparity condition, the mean disparity gradients for the two- and three-POTS configurations were 0.51 and 0.41, respectively; for the 10.3-arcmin disparity condition, the mean disparity gradients for the two- and three-POTS configurations were 0.68 and 0.55, respectively. Each mean disparity gradient in this experiment was calculated as a ratio using the largest disparity difference between the two outer surfaces of the POTS and the average distance between the dot elements on the surfaces, as in Tsirlin et al. (2008). Although it is not known yet whether such differences in mean disparity gradients can affect the magnitude of perceived depth, Bülthoff, Fahle, and Wegmann (1991) reported that the perceived relative depth between two fused stimuli with a constant disparity, presented side by side of a fixation stimulus, decreases as a function of their horizontal distance or disparity gradient, when the disparity gradient is more than 0.3. On the contrary, McKee and Verghese (2002) reported that disparity gradients less than 2.0 did not affect the magnitude of perceived depth in stereo transparency. We will discuss this issue further in Experiment 3, because the mean disparity gradient depends partly on dot density.

Experiment 2: Discrimination of perceived depth for POTS configuration

In this experiment, we estimated a disparity value of the outer surfaces of a two-POTS configuration that would give rise to the same amount of perceived depth as a three-POTS configuration with a given disparity.

Method

Stimuli The two-POTS and three-POTS configurations were displayed laterally on both sides of the midsagittal plane, and they were horizontally separated 3.2 arcdeg from each other. The disparity of the two outer surfaces of the three-POTS configuration was 6.5, 9.7, and 19.4 arcmin. The disparity pairs used for the outer surfaces of the "reference" three-POTS configuration were 3.2 and –3.2 arcmin, 4.8 and –4.8 arcmin, and 9.7 and –9.7 arcmin; the middle surface was at the screen distance with zero disparity. A positive or a negative value represents a crossed or an uncrossed disparity, respectively, with respect to the screen. The disparity of the two-POTS configuration was varied (3.2 to 9.7, 6.5 to 12.9, or 16.2 to 22.6 arcmin) every five steps, whereas that of the reference three-POTS configuration (6.5, 9.7, or 19.4 arcmin) was kept constant. Each stereogram consisted of 300 dots (1.5 dots/deg^2), and thus, each surface consisted of 150 dots (0.7 dots/deg^2) and 100 dots (0.5 dots/deg^2) in the two- and three-POTS configurations, respectively.

Procedure On each trial, observers were asked (1) to report the total number of surfaces of the POTS configuration presented on the left side of the screen, (2) to judge the perceived depths between the outer surfaces of the two configurations, and (3) to indicate which had a larger depth magnitude. The stimuli were visible until the observers finished responding with estimates. Before the experiment, observers were given several training trials, which were randomly selected from those used in the experimental session, until the experimenter judged that observers understood the task. The experiment consisted of three blocks, and for each block the disparity of the three-POTS configuration was different. Each block included ten sessions, in each of which the disparities of the two-POTS configuration and its location were randomly selected from the five disparities and two locations (right and left), respectively, with one repetition. Thus, each observer performed 300 trials in total for the whole experiment. In the training and experimental sessions, the stimulus presentation time was unlimited and observers could move their eyes freely, as in Experiment 1.

Observers Seven students from the university community, ranging in age from 21 to 23 years, participated. They had normal or corrected-to-normal visual acuity and good stereopsis. One observer was an author of this study, and the others were naive as to the purpose of the experiment.

Results and discussion

First, we examined the number of surfaces as in Experiment 1 and found that the observers reported the correct number of surfaces for both the two- and three-POTS configurations. From the obtained data we calculated a "matched disparity" for each observer and each disparity condition for the three-POTS configuration. The matched disparity was defined as the disparity of the surface of the two-POTS configuration that produced the same magnitude of perceived depth as that of the two outer surfaces of the three-POTS configuration, for each of the three disparity conditions (6.5, 9.7, or 19.4 arcmin). It was calculated from the psychometric function of the percentage of the responses, in which the depth of the surfaces obtained with the two-POTS configuration was perceived to be larger than that of the three-POTS configuration, as a function of disparity of the two-POTS configuration. The psychometric function was fitted to the sigmoid function and we regarded the disparity value of the two-POTS configuration that produced 50 % response in the fitted function as the matched disparity.

To test whether the perceived depth for observers agree when the two-POTS and three-POTS configurations have the same disparity, we compared the matched disparity in the two-POTS configuration with the disparity of the reference three-POTS configuration. Additional *t*-tests showed that the matched disparity for the two-POTS configuration was significantly different from that for the three-POTS configuration, $t(6) = 5.81$, $p < .01$, $t(6) = 2.55$, $p < .05$, and $t(6) = 2.99$, $p < .05$, in the 6.5-, 9.7-, and 19.4-arcmin disparity conditions, respectively.

The significant statistical results can be seen in Fig. 3, in which each bar represents the matched disparity averaged over seven observers for each disparity condition of the three-POTS configuration. Figure 3 shows that the matched disparity of the two-POTS configuration is smaller than the disparity of the three-POTS configuration for every disparity condition, suggesting that the perceived depth between the outer surfaces for the three-POTS configuration appears smaller to observers than that for the two-POTS configuration when the surfaces of the two POTS configurations have the same disparity. The ratios of magnitude of perceived depth of the two-POTS configuration to that of the three-POTS configuration were 0.82, 0.97, and 0.96, for the 6.5-, 9.7-, and 19.4-arcmin disparity conditions, respectively (see Appendix A for how we calculated the ratios). Thus, the results from Experiment 2, in which the method of depth discrimination was used, and the results from Experiment 1, in which the method of depth reproduction was used, provide robust support for the phenomenon as previously reported by Aida and Shimono (2010).

Because the multi-POTS configurations were viewed without time limits and eye movements were not controlled in the present study, one might think that eye movements, particularly vergence eye movements, could have played a role in the depth reduction phenomenon. Indeed, Akerstrom and Todd (1988) argued that vergence eye movements are important for perceiving stereo transparency, even though they found that stereo transparency can be perceived for a two-POTS configuration with presentation time that was shorter than the vergence latency. However, we do not think that the vergence eye movement is an important factor on the depth reduction phenomenon. This is because the two- and three-POTS configurations were presented side by side simultaneously in this experiment, and, thus, any vergence eye movements would affect equally the magnitude of perceived depth of both the two- and three-POTS configurations. These arguments suggest that even if vergence did play a role in perceiving stereo transparency, it is difficult to explain the depth reduction phenomenon in terms of vergence eye movements only.

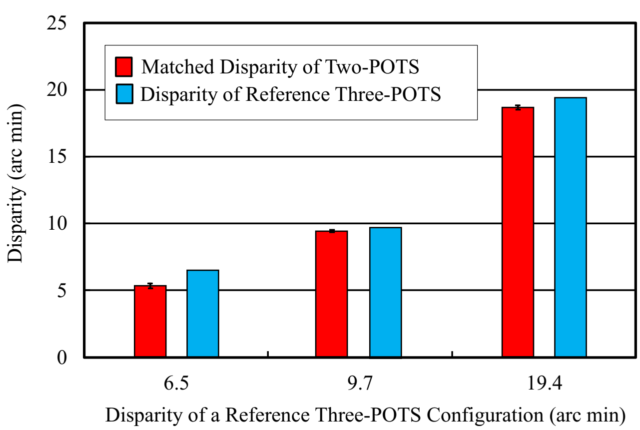

Fig. 3 Matched disparities of a two-POTS configuration as a function of the disparity of a reference three-POTS configuration in Experiment 2. The matched disparity is indicated by the left bar in each pair, and the disparity of the reference configuration is depicted by the right bar. The vertical lines attached to the left bars indicate the *SE*s

Experiment 3: Effect of dot density

The aim of this experiment was to examine the effect of dot density on the depth reduction phenomenon. We asked observers to reproduce the magnitude of perceived depth, as in Experiment 1. We kept the total dot density of a two-POTS configuration constant and varied that of a three-POTS configuration.

Method

Stimuli Each surface of the POTS configurations contained horizontal disparities corresponding to a frontal surface, centered in the visual field. The disparity of the surfaces of the two-POTS configuration and that of the outer surfaces of the three-POTS configuration were the same, and was either 6.5 or 12.9 arcmin. The disparity pairs used to generate the stimuli were 3.2 and –3.2 arcmin and 6.5 and –6.5 arcmin. A positive or a negative value represents a crossed or an uncrossed disparity, respectively. The two-POTS configuration consisted of 600 dots with a dot density of 2.2 dots/deg^2 of visual angle. The three-POTS configuration consisted of 300, 600, or 900 dots, with dot densities of 1.1, 2.2, or 3.3 dots/deg^2 of visual angle, respectively. The dot density refers to the total dot density of the stimulus. Thus, each surface consisted of 300 dots (1.1 dots/deg^2) in the two-POTS configuration, and 100 (0.4 dots/deg^2), 200 (0.7 dots/deg^2), and 300 (1.1 dots/deg^2) dots in the three-POTS configuration.

Procedure Observers performed in six experimental sessions, and the disparities and POTS configurations for each session were randomly selected from the two different disparities and four different POTS configurations (a two-POTS configuration and a three-POTS configuration with three different dot densities), with one repetition. Thus, each observer performed 48 trials in total. In the sessions, observers' tasks were the same as those in Experiment 1; they reported the total number of surfaces in the RDS and reproduced the magnitude of the depth for the two- or three-POTS configuration by using a tape measure.

Observers Ten students from the university community, ranging in age from 20 to 24 years, participated. They had normal or corrected-to-normal visual acuity and good stereopsis. All of them were naive as to the purpose of the experiment.

Results and discussion

The procedure for data analysis was identical to that conducted for Experiment 1. First, we examined the number of surfaces reported as in Experiments 1 and 2 and found that the observers reported the correct number of surfaces for the two- and three-POTS configurations in every trial, suggesting that they observed stereo transparent surfaces as expected. Second, we transformed the reported depth value for every trial to a logarithmic value, because some observers assigned a relatively large number to the depth value. The transformed values for each disparity, each POTS, and each dot density condition were averaged for each observer, and the averaged values were used as basic scores for analysis. As in Fig. 2a, the validity of the procedure can be seen in Fig. 4a, which shows observers' mean transformed values for the three-POTS configurations with three different dot densities against those for the two-POTS configuration on linear–linear axes.

We conducted a two-way repeated measures ANOVA (4 POTS × 2 disparities) on the scores, and found that the main effects of POTS [$F(3, 27) = 13.59$, $p < .001$] and disparity [$F(1, 9) = 53.97$, $p < .001$], as well as their interaction [$F(3, 27) = 3.18$, $p < .05$], were all statistically significant. Post hoc analyses (Tukey test) showed that the scores for the three-POTS configurations with 300, 600, and 900 dots were significantly different from that of the two-POTS configuration in both the 6.5- and 12.9-arcmin disparity conditions ($ps < .05$), whereas any combinations of two scores for the three means were not significantly different.

The significant main effect of POTS can be seen in Fig. 4a, in which we drew a diagonal line as a reference to indicate where the scores from the two-POTS and the three-POTS configurations would be the same. Figure 4a shows that most scores lie below the diagonal, suggesting that the perceived depth for the three-POTS configuration is smaller than that for the two-POTS configuration. As in Experiment 1, we averaged the scores among ten observers for each condition and transformed back the mean scores in order to estimate the magnitudes of perceived depth of the two- and three-POTS configurations. The transformed-back values are plotted as a function of disparity in Fig. 4b, which shows that each of the transformed-back values or magnitudes of perceived depth for the three-POTS configurations (55.5, 56.0, and 55.1 cm for 300, 600, and 900 dots, respectively) is smaller than that (62.5 cm) of the two-POTS configuration for the 6.5-arcmin disparity condition; each of the magnitudes of perceived depth for the three-POTS configurations (86.9, 85.3, and 87.0 cm for 300, 600, and 900 dots, respectively) is smaller than that (91.7 cm) of the two-POTS configuration for the 12.9-arcmin disparity condition. These results are consistent with the results of Experiments 1 and 2, and they indicate that the depth reduction phenomenon is robust. Furthermore, as in Experiment 1, we calculated the ratios of the magnitude of perceived depth of the two-POTS configuration to that of the three-POTS configuration, which were averaged among the three dot-density conditions; the ratios were 0.89 and 0.94 for the 6.5- and 12.9-arcmin disparities, respectively. This result also indicates that the depth reduction phenomenon did occur in the present experiment.

The significant main effect of disparity can be seen in Fig. 4b, in which the magnitude of perceived depth increases as a function of disparity for both the two- and three-POTS configurations. The significant interaction of the two main effects can also be seen in Fig. 4b, in which the difference

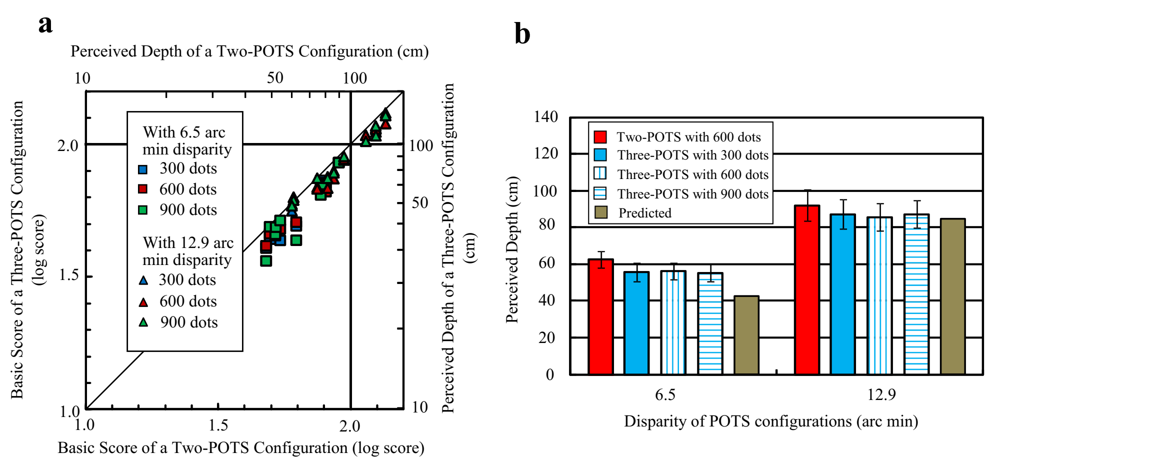

Fig. 4 Experiment 3. (**a**) Basic score of a three-POTS configuration as a function of that of a two-POTS configuration for each observer, as a function of disparity and dot density. The diagonal line is a reference line indicating the case when the scores for both POTS configurations are the same. The right *y*-axis indicates the magnitude of perceived depth in centimeters, which was transformed back from each basic score. (**b**) Mean perceived depths of two- and three-POTS configurations and the depth magnitude predicted from Eq. 1, in centimeters, as a function of disparity. The mean for each disparity, each POTS configuration, and each dot density was calculated by averaging the basic scores over ten observers and then transforming the average back by using the equation mean = $10 \wedge (\sum_{n=1}^{10} \log_{10}(\text{Reported Depth})_n/10)$. The vertical lines attached to the bars indicate 95 % confidence intervals that were transformed back from those for the basic scores

(5.3 cm) of the mean depth magnitude between the two-POTS configuration and the three-POTS configuration for the large disparity is smaller than that (7.0 cm) for the small disparity. The significant interaction is consistent with the fact that the ratio described above for the large disparity is larger than that for the small disparity.

More importantly, Fig. 4b shows that the perceived depth for the three-POTS configuration was relatively constant among the dot-density conditions used. This result suggests that the difference in the dot densities between two- and three-POTS configurations is not a critical factor for the depth reduction phenomenon reported in Experiment 1. In Experiment 3, the dot density in each surface of the three-POTS configuration was 0.4, 0.7, and 1.1 dots/deg^2 and that of the two-POTS configuration was kept at 1.1 dots/deg^2. In Experiment 1, the dot density in each surface of the three-POTS configuration was 0.7 dots/deg^2 and that of the two-POTS configuration was 1.1 dots/deg^2. Thus, the result that the perceived depth of the three-POTS configuration was similar to each other among the three dot density conditions in Experiment 3 implies that the perceived depth of the three-POTS configuration in Experiment 1 was not affected by the difference in dot densities between the two- and three-POTS configurations. That is, this suggests that dot density does not play an important role in the depth reduction phenomenon, at least within the range of dot densities used in this study. In the same vein, the disparity gradient is most likely not an important factor in contributing to the phenomenon.

On the surface the above discussion seems to be inconsistent with the report of Akerstrom and Todd (1988) in which observers reported difficulty in perceiving separate parallel planes as dot density increased. They argued that the difficulty can be due to the inhibitory interaction between disparity detectors or pools of detectors, which tune to each of the two different disparities of a two-POTS configuration; as dot density is increased, the inhibitory interaction might have been strengthened, thereby, leading to a degraded perception of stereo transparency. Note, however, that the range of dot densities in each surface of the two-POTS configuration (10.4–31.3 dots/deg^2) used in Akerstrom and Todd (1988) was much higher than that (0.7–1.1 dots/deg^2) used in Experiments 1–3 of the present study. With this consideration, one might resolve the difference in findings by arguing that the depth reduction phenomenon reported in the present study was observed only when the dot density of a multi-POTS configuration was sparse.

Our result that dot density had no effect on the depth reduction phenomenon may have a theoretical implication for stereopsis. In some of the computational models, the visual system is assumed to find correct matches between the right and left retinal images from many possible matches to perceive three-dimensional objects that produce the retinal images. Because incorrect matches (often called ghost images) are assumed to hinder the visual system in finding correct matches, an increase of dot density could deteriorate stereopsis (e.g., Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998). The fact that the dot density of the surface does not affect the reduction of the perceived depth in the three-POTS configuration suggests that the depth reduction phenomenon may not be due to the increase of ghost images. In other words, given the present results, it is difficult to explain the phenomenon in terms of inhibitory interaction between correct matches and incorrect matches only.

Might any other possible factor(s) play an important role in the depth reduction phenomenon, besides dot density? One possible candidate is the number of stereo surfaces that make up a POTS configuration. In the literature, some studies have suggested an inhibitory interaction at the level of a surface representation in stereo transparency (e.g., Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998). For example, Akerstrom and Todd proposed a hypothesis that assumes that stereo matching consists of both local and global processing. In the first step, locally detected disparity information is assumed to propagate to form a stereoscopically transparent surface, and in the second, the propagation is assumed to be inhibited by other overlapping transparent surfaces having different disparities. If the inhibitory interaction between stereo surfaces is strengthened in some way by the number of stereo surfaces, so that the magnitude of perceived depth decreases as the number increases, the present results can be explained. If this is the case, the number of stereo surfaces can have an effect on the depth reduction phenomenon. Thus, in Experiments 4 and 5 we manipulated the number of stereo surfaces to examine its possible effect on the perceived depth of multi-POTS configurations.

As in Experiment 1, for the next experiment we also examined whether the magnitude of perceived depth was consistent with that predicted from Eq. 1. Figure 4b shows the predicted depth magnitudes of 42.4 and 84.8 cm for the 6.5- and 12.9-arcmin disparity conditions, respectively. As can be seen from Fig. 4b, the magnitude of the perceived depth does not necessarily correspond to the one predicted; the perceived depth magnitudes of both the two- and three-POTS configurations are larger than those predicted for the smaller disparity, and are similar to that predicted for the larger-disparity conditions. Follow-up *t*-tests performed on the difference between the basic score and the logarithmically transformed predicted value from Eq. 1 indicated that the difference was statistically significant for both the two-POTS configuration, $t(9) = 5.19$, $p < .01$, and the three-POTS configurations with 300 dots, $t(9) = 3.08$, $p < .05$, 600 dots, $t(9) = 3.54$, $p < .01$, and 900 dots, $t(9) = 3.11$, $p < .05$, in the smaller-disparity condition. The results were not statistically significant for the larger-disparity condition. We will discuss this result in the General Discussion.

Experiment 4: Effect of number of stereo surfaces on perceived depth magnitude

The aim of this experiment was to examine the effect of the number of surfaces of a POTS configuration on the depth reduction phenomenon. We varied the number of surfaces from two to four while the disparity of the closest and the farthest surfaces was kept constant. Observers were asked to reproduce the magnitude of perceived depth, as in Experiment 1.

Method

Stimuli The disparity of the surfaces of the two-POTS configuration, that of the outer surfaces of the three-POTS configuration, and that of the outermost surfaces of the four-POTS configuration were the same and, depending on the experimental condition, were either 9.7 or 19.4 arcmin. The disparity pairs used to generate the stimuli were 4.8 with –4.8 arcmin, and 9.7 with –9.7 arcmin. A positive or a negative value represents a crossed or an uncrossed disparity, respectively. The surfaces for the four-POTS configuration were spaced 3.2 arcmin apart for the configuration having the 9.7-arcmin disparity and were spaced 6.5 arcmin apart for the configuration having the 19.4-arcmin disparity. Each stereogram consisted of 720 dots (2.7 dots/deg^2). The dot density refers to the total dot density of the stimulus. In the two-POTS configuration, each surface consisted of 360 dots (1.3 dots/deg^2). In the three-POTS configuration, each surface consisted of 240 dots (0.9 dots/deg^2). In the four-POTS configuration, each surface consisted of 180 dots (0.7 dots/deg^2).

Procedure The procedure of Experiment 4 was similar to that of Experiments 1 and 3. Observers performed in six experimental sessions, and the disparities and POTS configurations for each session were randomly selected from the two different disparities and three different POTS configurations, with one repetition. Thus, each observer performed 36 trials in total. Before the experimental sessions, observers were trained with trials that were randomly selected from those used in the experimental sessions till the experimenter confirmed that observers understood the task. Observers' tasks were the same as those in Experiments 1 and 3; they were asked to report the total number of surfaces in the RDS and to reproduce the magnitude of the depth between the closest and farthest surfaces of the two-, three-, and four-POTS configurations. The RDS was visible until the observers finished responding with their estimates.

Observers Eleven students from the university community, ranging in age from 20 to 24 years, participated. Seven were newly recruited and four had participated in Experiment 3. They had normal or corrected-to-normal visual acuity and good stereopsis. All of them were naive as to the purpose of the experiment.

Results and discussion

The procedure for data analysis was essentially the same as that for Experiments 1 and 3. Different from the previous three experiments, the results indicate that some observers reported a few "wrong" surface numbers; the averages with respect to the percentage of wrong responses were 0.03 %, 0.03 %, and 0.15 % for the two-, three-, and four-POTS configurations, respectively. We excluded the wrong responses in order to calculate the average for each condition. We also excluded one observer's data because he reported the wrong responses more than four times for the four-POTS configuration. (Thus, the data from ten observers were processed further.) Note that for stereo transparency, as the number of surfaces in a POTS configuration is increased, it gets more difficult to report correctly the number of stereo surfaces in the configuration (Tsirlin et al., 2008). After screening, we transformed the "correctly" reported depth values logarithmically for every trial, and the transformed values for each disparity and each POTS condition were averaged for each observer and used as basic scores for the analysis, as in Experiments 1 and 3. The validity of this procedure is reflected in the data shown in Fig. 5a, which shows observers' mean transformed values for the three- and four-POTS configurations against those for the two-POTS configuration on linear–linear axes. As can be seen in Fig. 5a, there was no unusual or skewed distribution of the transformed mean values at each disparity condition.

We conducted a two-way repeated measures ANOVA (3 POTS × 2 disparities) on the scores, and found that the main effects of POTS [$F(2, 18) = 14.24$, $p < .001$] and disparity [$F(1, 9) = 130.49$, $p < .001$] were statistically significant, whereas their interaction was not. Post hoc analyses (Tukey test) showed that each of the mean scores for the three-POTS and four-POTS configurations was significantly different from that for the two-POTS configuration (ps < .05), whereas no statistically significant difference was apparent between the mean score for the three-POTS configuration and that for the four-POTS configuration.

The significant main effect of POTS can be seen in Fig. 5a, in which we have drawn a diagonal line as a reference to indicate where the score for the three- or the four-POTS configuration and that for the two-POTS configuration would be the same. Figure 5a shows that most scores lie below the diagonal, which suggests that the perceived depths for the three- and four-POTS configurations are smaller than those for the two-POTS configuration. As in Experiments 1 and 3, we estimated the difference of the magnitude of perceived depth among all three POTS configurations by transforming the averaged score over ten observers back. Figure 5b shows the transformed-back averaged score or the magnitude of perceived depth as a function of disparity with the number of stereo surfaces as the parameter; for the 9.7-arcmin disparity condition, the mean (63.2 cm) for the three-POTS

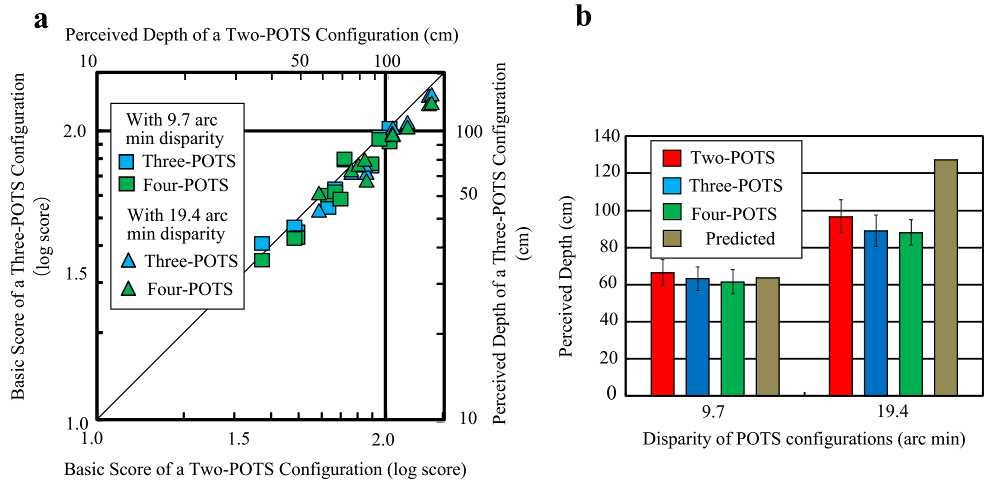

Fig. 5 Experiment 4. (**a**) Basic score of a three- or a four-POTS configuration as a function of that of a two-POTS configuration for each observer, as a function of disparity. The diagonal line is a reference line indicating the case when the score of the three- or the four-POTS configuration is the same as that of the two-POTS configuration. The right *y*-axis indicates the magnitude of perceived depth in centimeters, which was transformed back from each mean basic score. (**b**) Mean perceived depths of two-, three-, and four-POTS configurations and the depth magnitude predicted from Eq. 1, in centimeters, as a function of disparity. The mean for each disparity condition and each POTS configuration was calculated by averaging the basic scores over ten observers and then transforming the average back by using the equation mean = 10 ^ $(\sum_{n=1}^{10} \log_{10}(\text{Reported Depth})_n/10)$. The vertical lines attached to the bars indicate 95 % confidence intervals that were transformed back from those for the basic scores

configuration and the mean (61.3 cm) for the four-POTS configuration were smaller from that (66.3 cm) for the two-POTS configuration; for the 19.4-arcmin disparity condition, the mean (88.9 cm) for the three-POTS configuration and that (88.1 cm) for the four-POTS configuration were smaller than the mean (96.7 cm) for the two-POTS configuration. The ratios of the magnitude of perceived depth of the two-POTS configuration to that of the three-POTS configuration, when averaged, were 0.95 and 0.92 for 9.7- and 19.4-arcmin disparities, respectively. The ratios of magnitude of perceived depth of the two-POTS configuration to that of the four-POTS configuration, when averaged, was 0.92 and 0.91 for the 9.7- and 19.4-arcmin disparity conditions, respectively.

These results are consistent with the idea that the number of stereo surfaces plays a role in the depth reduction phenomenon. However, the magnitude of perceived depth for the three-POTS configuration was similar to that for the four-POTS configuration. Does this mean that the depth reduction phenomenon is not observed when comparing these two configurations? Before making a conclusion on this issue, in a follow-up experiment (Exp. 5) we reexamined the question whether or not the depth reduction phenomenon can be observed between the two configurations if a different method was used, namely, with the method of depth discrimination.

The main effect of disparity can be seen in Fig. 5b, in which the magnitude of perceived depth for the 19.4-arcmin disparity is larger than that for the 9.7-arc min disparity each for the two-, three-, and four-POTS configurations. Furthermore, Fig. 5b shows the magnitudes predicted from Eq. 1, which are 63.6 and 127.2 cm for the 9.7- and 19.4-arcmin disparity conditions, respectively. As can be seen from Fig. 5, the magnitude of the perceived depth does not necessarily correspond to the predicted value; each of the perceived depth magnitude of the two-, three-, and four-POTS configurations are similar to the predicted for the smaller disparity and are smaller than the predicted for the larger disparity. Follow-up *t*-tests performed on the differences between the basic score and the logarithmically transformed predicted value from Eq. 1 indicated that the differences were statistically significant for the two-POTS configuration, $t(9) = 3.01$, $p < .05$, the three-POTS configuration, $t(9) = 4.06$, $p < .01$, and the four-POTS configuration, $t(9) = 5.13$ $p < .001$, in the larger-disparity condition, but not in the smaller-disparity condition. We will discuss these results together with those from Experiments 1 and 3 in the General Discussion.

Experiment 5: Effects of number of stereo surfaces on depth discrimination

The aim of this experiment was to examine whether the depth of a three-POTS configuration would appear to be smaller than that for a four-POTS configuration if a different task—that is, the method of depth discrimination—were used instead of the method of reproduction, as in Experiment 4.

Method

Stimuli The three-POTS and four-POTS configurations were displayed laterally on both sides of the midsagittal plane, and the horizontal separation between the laterally displaced POTS was 3.8 arcdeg. The disparity of the outermost surfaces of the four-POTS configuration was 7.7 or 19.4 arcmin, and the interplane disparities for the four-POTS configurations were the same as each other: Those for the 7.7-arcmin disparity conditions were 3.9, 1.3, –1.3, and –3.9 arcmin, and those for the 19.4-arcmin conditions were 9.7, 3.2, –3.2, and –9.7 arcmin. For the two experimental conditions in which the four-POTS configuration had 7.7- and 19.4-arcmin disparities, the disparity of the three-POTS configuration was varied from 5.2 to 10.3 and from 16.2 to 22.6 arcmin, respectively, in five steps within each range. A positive or a negative value represents a crossed or an uncrossed disparity, respectively. Each stereogram consisted of 600 dots, and its dot densities were 2.7 and 3.0 dots/deg^2 of visual angle for the disparity conditions of 7.7 and 19.4 arcmin, respectively. The dot density refers to the total dot density of the stimulus. Thus, in the three- and four-POTS configurations, each surface consisted of 200 dots (0.9 and 1.0 dots/deg^2 of visual angle for the 7.7 and 19.4 arcmin, respectively) and of 150 dots (0.67 and 0.74 dots/deg^2 of visual angle for 7.7 and 19.4 arcmin, respectively).

Procedure The procedure for Experiment 5 was the same as that for Experiment 2. The experiment consisted of two blocks of trials, and for each block the disparity of the four-POTS configuration was different. Each block consisted of ten sessions, in each of which disparities of the three-POTS configuration and its location were randomly selected from the five disparities and two locations (right and left), respectively, with one repetition. Six observers were assigned to each one of the two blocks and, thus, each observer performed 100 trials in total in each experimental session.

Observers Twelve students from the university community, ranging in age from 20 to 25 years, participated. Eight of them had participated in the previous experiments; four had participated in Experiment 1, one in Experiment 2, and three in Experiments 3 and 4. They had normal or corrected-to-normal visual acuity and good stereopsis. All of them were naive as to the purpose of the experiment.

Results and discussion

The procedure for the data analysis was the same as that for Experiment 2. After confirming that observers reported the correct number of surfaces (three or four) for the POTS configuration presented on their left side, we calculated a "matched disparity" for each observer and disparity condition (7.7 or 19.4 arcmin). The "matched disparity" was defined as the disparity of the two outer surfaces of the three-POTS configuration that produced the same magnitude of perceived depth as that of the two outermost surfaces of the four-POTS configuration. It was calculated from the psychometric function of the percentage of the responses, in which the depth of the surfaces obtained with the three-POTS configuration was perceived to be larger than that of the four-POTS configuration, as a function of the disparity of the three-POTS configuration. The psychometric function was fitted to the sigmoid function, and the disparity value of the three-POTS configuration that produced 50 % response in the fitted function was taken to be the value for the matched disparity.

To test whether the perceived depths were the same when the two outermost surfaces of the three- and four-POTS configurations had the same disparity, we compared the matched disparity in the three-POTS configuration and the disparity of the four-POTS configuration. Additional *t*-tests showed that the matched disparity for the three-POTS configuration was significantly different from that for the four-POTS configuration, $t(5) = 3.22$, $p < .05$, and $t(5) = 4.01$, $p < .05$, in the 7.7- and 19.4-arcmin disparity conditions, respectively.

The significant statistical results can be seen in Fig. 6, in which the left vertical bar in each pair represents the matched disparity averaged over six observers for each disparity condition of the four-POTS configuration. As can be seen in Fig. 6, the matched disparity of the four-POTS configuration is smaller than the disparity used in the three-POTS configuration. From the mean value of the matched disparity, we estimated the magnitudes of the perceived depths of the three- and four-POTS configurations (see Appendix A), as in Experiment 2. The ratios of the magnitude of perceived depth of the three-POTS configuration to that of the four-POTS configuration, when averaged, were 0.89 and 0.96 for

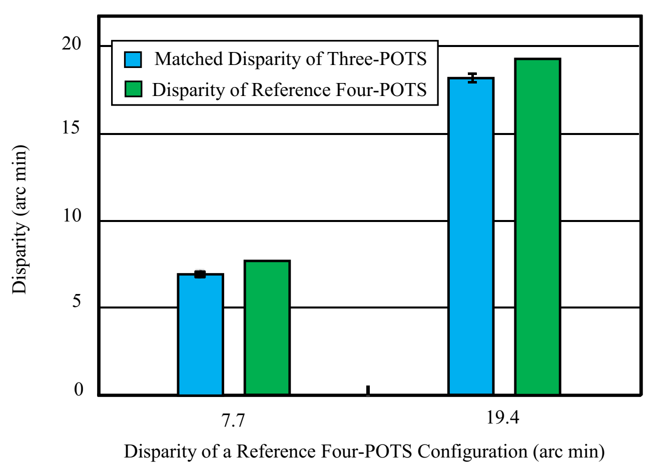

Fig. 6 Matched disparities of a three-POTS configuration as a function of the disparity of a reference four-POTS configuration in Experiment 5. The matched disparity is indicated by the bar on the left of each pair, and the disparity of the reference configuration is indicated by the right bar. The vertical lines attached to the left bars indicate the *SE*s

the 7.7- and 19.4-arcmin disparity conditions, respectively. These results are consistent with the idea that the number of stereo surfaces plays a role in the depth reduction phenomenon.

Note, however, a slight difference between the results from the depth reproduction task and those from the depth discrimination task. The depth reduction phenomenon between the three-POTS and four-POTS configurations was observed for the depth discrimination task in the present experiment, but not for the depth reproduction task in Experiment 4. This difference suggests that the first (discrimination) task is more sensitive than the second (reproduction) task in being able to reveal the phenomenon.

General discussion

Five experiments showed that the perceived depth for random-dot stereograms that depict parallel, overlapping, transparent stereoscopic surfaces can be different, depending on the number of stereoscopic surfaces depicted, even if the relevant surfaces depict the same disparity (from 5.2 to 19.4 arcmin) and are viewed at the same distance (380 cm). This phenomenon was observed with the perceived depth obtained using two different tasks—a depth reproduction task and a depth discrimination task. In Experiments 1, 3, and 4, the magnitude of perceived depth between the two outer stereo surfaces of the three- or the four-POTS configuration was smaller than that between the two outer stereo surfaces of the two-POTS configuration. In Experiment 2, viewer-matched disparity of the two surfaces of the two-POTS configuration was smaller than that of the two outer surfaces of the three-POTS configuration even though the magnitudes of perceived depth between the two surfaces of both configurations were the same. In Experiment 5, the matched disparity of the three-POTS configuration, giving rise to the same amount of perceived depth as the two outer surfaces of the four-POTS, was smaller than that of the four-POTS. Overall, the results of Experiments 1–4 showed that the depth of the three-POTS configuration was reduced 8 %, compared with that of the two-POTS configuration. The results of Experiments 4 and 5 showed that the depth of the four-POTS configuration was reduced 5 %, compared with that of the three-POTS configuration. These results suggest that the perceived depth for the POTS configuration depends on the number of stereo surfaces present in the configuration.

We conducted a further analysis to compare the data obtained from our two different tasks. As is depicted in Fig. 7, we replotted the data from Experiments 1, 3, and 4, which were obtained with the depth reproduction task, as well as the data from Experiments 2 and 5, which were obtained with the depth discrimination task. We assumed that the magnitude of perceived depth would increase approximately as a linear function of disparity, as in Appendix A, and fitted two regression lines with zero intercepts to the two-POTS and three-POTS data in Fig. 7a, except at the largest disparity. These regression lines are depicted in Fig. 7b. In Fig. 7b, we also plotted the depth values of the matched disparities of the two-POTS configuration. The depth values of the two-POTS configuration have the same perceived depth values (y values) as those of the reference three-POTS configuration, but have disparities (x values) as measured in Experiment 2 (i.e., matched disparities). For comparison, the relevant depth values for the reference three-POTS configuration are those located on the regression line fitted for the three-POTS data at disparities of 6.5, 9.7, and 19.4 arcmin (not marked). Furthermore, the same procedure was repeated for the matched disparities (7.7 and 19.4 arcmin) of the three-POTS configuration obtained in Experiment 5. As can be seen in Fig. 7b, the depth values indicated by the symbols for the matched disparities are located relatively close to the fitted lines. Since the symbols were calculated from the data obtained from the depth discrimination task, and the lines from the depth reproduction task, this result indicates that the data from the depth reproduction task are consistent with those from the depth discrimination task.

The depth reduction phenomenon clearly shows that the magnitude of perceived depth between the two outermost surfaces of the multi-POTS configurations is not always the same even when the surfaces have the same binocular disparity and are viewed at the same viewing distance. This finding is inconsistent with the prediction from Eq. 1. Nevertheless, the results of Experiments 1, 3, and 4 also indicated that the magnitude of perceived depth agreed well with that predicted from Eq. 1 in five of the seven disparity conditions, as we discussed previously in the Results section under each of the experiments. Furthermore, the perceived depth covaried well with the disparity predicted from Eq. 1, although the magnitude for the largest disparity used (19.4 arcmin) was smaller than predicted[1] (see Fig. 7a). The results that both the depth reduction phenomenon and covariation between the perceived depth and the disparity for POTS configurations occur at the same time may be explained by the hypothesis that the visual system calculates perceived depth according to the geometrical relationship represented in Eq. 1 and, furthermore, that the "registered" (or perceived) distance information and/or "registered" (or perceived) disparity information can be affected by one or more properties of the stimulus.

[1] Although we do not yet have a clear explanation of why the magnitude of perceived depth was underestimated from the equation only for the largest disparity, such underestimation has been reported in the literature (e.g., Patterson & Martin, 1992; Richards, 1971; Ritter, 1977, 1979; Shimono et al., 2002).

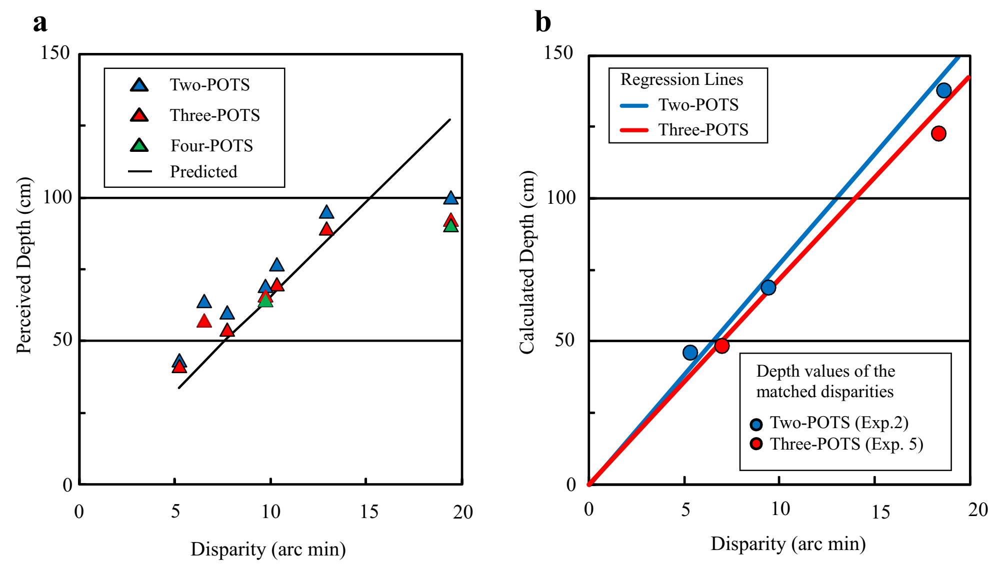

Fig. 7 Mean perceived depth (**a**) and predicted depth (**b**) of the two-, three-, and four-POTS configurations as a function of disparity. (**a**) The perceived depths of the two- and three-POTS configurations are the results from Experiments 1, 3, and 4 (depth reproduction task), and those of the four-POTS configuration are the results of Experiment 4. The solid line indicates the magnitude of depth calculated from Eq. 1. (**b**) Symbols indicate the depth values that were calculated from the depth discrimination data of Experiments 2 and 5, which are plotted along with regression lines that are fitted to the depth reproduction data shown in Fig. 7a, except for the largest disparity used (see the text for the computational details)

Consistent with the hypothesis, Stevenson et al. (1991) reported that the perceived location of the front surface of a two-POTS configuration depended on the disparity of the two-POTS configuration; the perceived location can be farther from observers than its physical location when the disparity of the POTS is relatively small (2–3 arcmin) and can be perceived to be closer to observers when its disparity is relatively large (6–8 arcmin). Because the perceived location of the farthest surface was not measured in that study, it is unclear whether the results were attributed to the perceived egocentric distance, the perceived depth or both. Whatever the case, their results suggest that the perceived egocentric distance or perceived disparity depends on a stimulus property of the POTS configuration.

Interestingly, Stevenson et al. (1991) proposed a model in which the perceived relative distance between two perceived surfaces, or perceived disparity, of a two-POTS configuration may deviate from a "physical" disparity, to explain the phenomenon they reported. They argued that the perceived surface locations of a two-POTS configuration can be simulated by the profiles of cross-correlation functions between the left and right retinal images. Following Stevenson et al. (1991), we applied the cross-correlation analysis to examine whether the depth reduction phenomenon can be explained by the profiles of the cross-correlation function. If the peaks of the profiles are squeezed closer to each other as the number of stereo surfaces is increased, we can argue that the simulated relative distance between the two outermost surfaces of the multi-POTS configurations (or the magnitude of perceived depth of the configurations) can be reduced as a function of their surface numbers. Our analysis on the cross correlation, calculated using the more contemporary version of the cross-correlation method (Filippini & Banks, 2009), indicates that the distance between the two outermost peaks of the profiles or simulated disparity corresponds well to the disparity of the stereo surfaces (see Appendix B for details). The analysis does not show a dependency on the number of stereo surfaces.

We interpret the fact that the results of the cross correlation analysis are not consistent with those of the present experiments as indicating that the phenomenon does not occur at an early stage of processing for disparity detection but do occur later at the level of visual processing at which representation of stereo surfaces is achieved. Here we assume that binocular matching by computing correlation between the right and left visual images, as used in the cross correlation analysis, represents an early disparity detection process (e.g., Banks, Gepshtein & Landy, 2004; Filippini & Banks, 2009). The fact that a cross-correlation model does not explain our data suggests that interactions between the representation of surfaces at a higher level are likely to be responsible for the depth reduction phenomenon.

The idea that an interaction between stereo surfaces plays a role in stereo transparency was proposed by Tsirlin et al. (2012). They reported that the minimum interplane separation (or disparity) between the two outermost surfaces of a two-POTS configuration to see two well-segregated surfaces depends on the distribution of dots to the two surfaces when the overall dots density of the configuration was constant. They explained the phenomenon in terms of a figure–ground process in addition to local (excitatory) and global (inhibitory) processes that are often assumed in models of stereo transparency (e.g., Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998; see the Results and Discussion section of Exp. 3 above). The figure–ground process is assumed to provide a bias to a back surface, such that it assigns the disparity of dots in the back surface to their surrounding "blank" areas to form an opaque ground surface. It also increases neural activities representing the back surface and determines the minimum interplane disparity to see transparency, via its excitatory feedback to the local and global processes. The depth reduction phenomenon can be explained if the degree of bias correlates with perceived depth[2] and if the bias associated with the back surface is "weakened" when one or more surfaces are introduced between the two outer surfaces. This idea can explain the fact shown in Fig. 7a that the magnitude of perceived depth is larger than those predicted by Eq. 1 at most disparity conditions, except for the largest, and is closer to those predicted by Eq. 1 as a function of the number of stereo surfaces; if the excitatory bias of a back surface of a two-POTS configuration increases the magnitude of perceived depth, and introducing other surface(s) weakens the bias, the perceived depth of multi-POTS configuration may change, as reflected in the data shown in Fig. 7a, except for the largest disparity. Although Tsirlin et al.'s model (2012) does not deal with the magnitude of perceived depth as it stands, we think that it can be applied to explain the depth reduction phenomenon.

Finally, in the present study, we found the phenomenon in which the magnitude of perceived depth for random-dot stereograms depicting multiple POTS can be reduced as a function of the number of stereo surfaces. This finding indicates that the number of stereo surfaces is one of several factors that may affect perceived stereo transparency, such as disparity magnitude (Stevenson et al., 1989), dot density (Akerstrom & Todd, 1988; Tsirlin et al., 2008, 2012; Wallace & Mamassian, 2004), and the color (Akerstrom & Todd, 1988), and luminance (Gepshtein & Cooperman, 1998) of dots. We suggest that the depth reduction phenomenon observed in this study is probably the result of a higher-order mechanism mediating stereo transparency, with multiple contributing factors.

[2] Although Tsirlin et al. (2012) did not measure the magnitude of depth, their Fig. 3 shows that when a third plane was introduced, the minimum interplane separation for the two-POTS configuration had to be set larger for observers to perceive two separate stereo surfaces. If the different minimum disparities produced the same magnitude of perceived depth to see two separable stereo surfaces, the result can be interpreted to indicate that the magnitude of perceived depth for the two-POTS configuration is larger than that for the configuration containing the third surface, consistent with the depth reduction phenomenon.

Author note This work was partly supported by the Sasakawa Scientific Research (23-208) to the first author and by grant-in-aids for Scientific Research (23330215) provided by the Japanese Ministry of Education, Science, and Culture, to the second author. Part of this study was reported in SPIE (2012) held at San Francisco.

Appendix

A: Predicted ratios of magnitude of perceived depth for Experiments 2 and 5

To calculate the difference in perceived depth between the two-POTS and three-POTS configurations in Experiment 2 and that between the three-POTS and four-POTS configurations in Experiment 5, we assumed that the magnitude of perceived depth (d') is expressed by the following equation:

$$\mathrm{d}' \fallingdotseq Kn\delta, \tag{A1}$$

where K is constant for each POTS, n denotes the number of stereoscopic surfaces, and δ is the disparity. In Experiment 2, the perceived depth obtained with the two-POTS configuration with disparity (δ_2) was the same as that for the three-POTS configuration with disparity (δ_3). From Eq. A1, we obtain

$$K_2\delta_2 = K_3\delta_3, \tag{A2}$$

where δ_2 is the matched disparity of the two-POTS configuration and δ_3 is the corresponding disparity of the three-POTS configuration. From Eq. A2, we calculate K_2 over K_3 (K_2/K_3), which is the predicted ratio of the perceived depth for the three-POTS configuration to that for the two-POTS configuration when their two outer surfaces have the same disparity. Similarly, we calculated K_3 over K_4 (K_3/K_4), which is the predicted ratio of the perceived depth for the four-POTS configuration to that for the three-POTS configuration when their two outermost surfaces have the same disparity.

B: Details of the cross-correlation analysis

We simulated the psychophysical experiments described in this article to examine whether the cross-correlation analysis can explain the depth reduction phenomenon. The procedure used to calculate the cross-correlation values was the same as that used in Filippini and Banks (2009). After calculating the values, subpixel interpolation of the correlation function was carried out in order to obtain decimal values of disparity. The subpixel disparity was computed by fitting parabolas to three cross-correlation values around each of the local maxima in the correlation function (Shimizu & Okutomi, 2002). We regarded a maximum (mode) of the parabolas as the perceived location of a stereo surface, as in Stevenson et al. (1991), and estimated the perceived disparity of two outermost surfaces in a multi-POTS configuration.

In the simulation, we generated two-, three-, and four-POTS configurations that had five disparities (6.2, 9.4, 12.5, 15.6, and 18.7 arcmin), which were similar to those used in Experiments 1–5. The total dot density was 2.2 dots/deg^2, which was within the range of dot densities used in Experiments 1–5. The dots were distributed equally on each stereo surface, and there were eight trials for each of the five disparities and the three POTS configurations. Thus, 120 trials in total were presented for each dot-density condition.

We averaged the simulated disparities between the two outermost surfaces over eight trials, which was calculated from local modes in the function for each POTS configuration, and examined whether or not the depth reduction phenomenon occurred. Table 1 shows the ratios of the depth magnitudes, which were calculated from the averaged simulated disparities, between each two POTS configurations (i.e., the ratio of the depth magnitude for a three-POTS configuration to that for a two-POTS configuration, the ratio of the depth magnitude for a three-POTS configuration to that for a two-POTS configuration, and the ratio of the depth magnitude for a three-POTS configuration to that for a four-POTS configuration). We judged that the depth reduction phenomenon occurred when the ratio was less than unity. As can be seen from Table 1, the depth reduction phenomenon can be seen for smaller disparities (6.2, 9.4, and 12.5 arcmin), but not for the larger disparities (15.6 and 18.7 arcmin). Note, however, that the degree of depth reduction in the simulation was much smaller than that observed in the psychophysical experiments described in this article. Thus, the results of the simulated experiment using the cross-correlation methods are not necessarily consistent with those obtained in the psychophysical experiments.

Table 1 Cross-correlation simulation results: Ratio of the magnitude of predicted depth of the two-POTS configuration to that of the three-POTS configuration (3P/2P), ratio of the magnitude of predicted depth of the two-POTS configuration to that of the four-POTS configuration (4P/2P), and ratio of the magnitude of predicted depth of the three-POTS configuration to that of the four-POTS configuration (4P/3P)

	Disparity (in arcmin) Between Two Outermost Surfaces				
Surface Combination	6.2	9.4	12.5	15.6	18.7
3P/2P	.98	.99	1.00	1.00	1.00
4P/2P	.95	.99	.99	1.00	1.00
4P/3P	.97	1.00	.98	1.00	1.00

References

Aida, S., & Shimono, K. (2010). Difference in perceived depth between two stereo-overlapping and three stereo-overlapping planes [Rittai to-meishi ni okeru mikakeno okuyuki: 2 men to 3 men no sa] (in Japanese with English abstract). *Vision*, *22*, 1–11. Retrieved from www.visionsociety.jp/VISION/vol22/no1/VISION220101.pdf

Akerstrom, R. A., & Todd, J. T. (1988). The perception of stereoscopic transparency. *Perception & Psychophysics, 44,* 421–432. doi:10.3758/BF03210426

Anderson, B. L. (1992). Hysteresis, cooperativity, and depth averaging in dynamic random-dot stereograms. *Perception & Psychophysics, 51,* 511–528. doi:10.3758/BF03211649

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience, 24,* 2077–2089.

Bülthoff, H., Fahle, M., & Wegmann, M. (1991). Perceived depth scales with disparity gradient. *Perception, 20,* 145–53. doi:10.1068/p200145

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science, 208,* 615–617. doi:10.1126/science.7367885

Filippini, H. R., & Banks, M. S. (2009). Limits of stereopsis explained by local cross-correlation. *Journal of Vision, 9*(1), 8.1–18. doi:10.1167/9.1.8

Gepshtein, S., & Cooperman, A. (1998). Stereoscopic transparency: A test for binocular vision's disambiguating power. *Vision Research, 38,* 2913–2932. doi:10.1016/S0042-6989(98)00091-1

Glennerster, A., Rogers, B. J., & Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. *Vision Research, 36,* 3441–3456. doi:10.1016/0042-6989(96)00090-9

Howard, I. P., & Rogers, B. (2002). *Depth perception: Vol. 2. Seeing in depth*. Toronto: Porteous.

Howard, I. P., & Rogers, B. (2012). *Stereoscopic vision: Vol. 2. Perceiving in depth*. New York: Oxford University Press.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research, 31,* 1351–1360. doi:10.1016/0042-6989(91)90056-B

McKee, S. P., & Verghese, P. (2002). Stereo transparency and the disparity gradient limit. *Vision Research, 42,* 1963–1977. doi:10.1016/S0042-6989(02)00073-1

Nakamizo, S., & Shimono, K. (2001). Depth scaling by the visual system with absolute distance information: A review [Shikakukei niyoru zettaikyori jyo-ho o mochiita okuyuki no suke-rinngu] (in Japanese). *Vision*, *13*, 163–180. Retrieved from www.visionsociety.jp/VISION/vol13/no3/VISION130303.pdf

Ono, H., & Comerford, J. (1977). Stereoscopic depth constancy. In W. Epstein (Ed.), *Stability and constancy in visual perception: Mechanisms and processes* (pp. 91–128). New York: Wiley.

Patterson, R., & Martin, W. L. (1992). Human stereopsis. *Human Factors, 34,* 669–692.

Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America, 61,* 410–414. doi:10.1364/JOSA.61.000410

Ritter, M. (1977). Effect of disparity and viewing distance on perceived depth. *Perception & Psychophysics, 22,* 400–407. doi:10.3758/BF03199707

Ritter, M. (1979). Perception of depth: processing of simple positional disparity as a function of viewing distance. *Perception & Psychophysics, 25,* 209–214. doi:10.3758/BF03202988

Shimizu, M., & Okutomi, M. (2002). Significance and attributes of subpixel estimation on area based matching [gazou no matching ni okeru sabupikuseru suitei no imi to seisitu], *The Institute of Electronics, Information and Communication Engineers* D-, J85-D-1791-1800. (in Japanese)

Shimono, K., & Nakamizo, S. (1990). Stereoscopic depth perception as a function of absolute distance in stereonormal and stereoanomalous observers [Abstract]. *Perception, 19,* 362–363.

Shimono, K., Tam, W. J., Stelmach, L., & Hildreth, E. (2002). Stereo-illusory motion concomitant with lateral head movements. *Perception & Psychophysics, 64,* 1218–1226. doi:10.3758/BF03194767

Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1989). Hyperacuity, superresolution and gap resolution in human stereopsis. *Vision Research, 29,* 1597–1605. doi:10.1016/0042-6989(89)90141-7

Stevenson, S. B., Cormack, L. K., & Schor, C. M. (1991). Depth attraction and repulsion in random-dot stereograms. *Vision Research, 31,* 805–813. doi:10.1016/0042-6989(91)90148-X

Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2008). Stereoscopic transparency: Constraints on the perception of multiple surfaces. *Journal of Vision, 8*(5), 5.1–10. doi:10.1167/8.5.5

Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2012). Perceptual asymmetry reveals neural substrates underlying stereoscopic transparency. *Vision Research, 54,* 1–11. doi:10.1016/j.visres.2011.11.013

Wade, N. J., & Ono, H. (2012). Early studies of binocular and stereoscopic vision. *Japanese Psychological Research, 54,* 54–70. doi:10.1111/j.1468-5884.2011.00505.x

Wallace, J. M., & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces. *Vision Research, 44,* 2253–2267. doi:10.1016/j.visres.2004.04.013