Vision Research 54 (2012) 1–11

Contents lists available at SciVerse ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Perceptual asymmetry reveals neural substrates underlying stereoscopic transparency

Inna Tsirlin [a,*], Robert S. Allison [a,b], Laurie M. Wilcox [a]

[a] Centre for Vision Research, Department of Psychology, CSEB 0001, York University, 4700 Keele St., Toronto, Ontario, Canada
[b] Centre for Vision Research, Department of Computer Science & Engineering, CSEB 3051, York University, 4700 Keele St., Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 27 June 2011
Received in revised form 8 November 2011
Available online 13 December 2011

Keywords:
Stereoscopic transparency
Perceptual asymmetry
Bias
Neural network
Disparity interpolation
Signal strength

ABSTRACT

We describe a perceptual asymmetry found in stereoscopic perception of overlaid random-dot surfaces. Specifically, the minimum separation in depth needed to perceptually segregate two overlaid surfaces depended on the distribution of dots across the surfaces. With the total dot density fixed, significantly larger inter-plane disparities were required for perceptual segregation of the surfaces when the front surface had fewer dots than the back surface compared to when the back surface was the one with fewer dots. We propose that our results reflect an asymmetry in the signal strength of the front and back surfaces due to the assignment of the spaces between the dots to the back surface by disparity interpolation. This hypothesis was supported by the results of two experiments designed to reduce the imbalance in the neuronal response to the two surfaces. We modeled the psychophysical data with a network of inter-neural connections: excitatory within-disparity and inhibitory across disparity, where the spread of disparity was modulated according to figure-ground assignment. These psychophysical and computational findings suggest that stereoscopic transparency depends on both inter-neural interactions of disparity-tuned cells and higher-level processes governing figure ground segregation.

1. Introduction

Stereopsis allows us to reconstruct the three-dimensional structure of the world using the differences between the positions and appearance of the images of objects on the two retinae. An intriguing phenomenon in stereoscopic depth perception is stereo-transparency. Stereo-transparency refers to the percept of pseudo-transparent, or lacy, surfaces where depth is defined solely by binocular disparity. Pseudo-transparency is one of the three types of transparency as outlined by Tsirlin, Allison, and Wilcox (2008):

1. Glass-transparency—light passes through objects made of clear transparent materials such as glass.
2. Translucency—translucent materials allow light to pass through them only diffusely and cannot be clearly seen through. Examples of such materials are frosted glass and certain types of cloth.
3. Pseudo-transparency—light passes through gaps in non-transparent lacy objects such as wire fences or tree branches.

Pseudo-transparent surfaces in the real world can be simulated in the laboratory using transparent random-dot stereograms (RDSs) with overlaid planes, known as stereo-transparent RDS. Stereo-transparency is an example of the thin structure problem as the disparity field is discontinuous nearly everywhere. That is, at many locations in a stereo-transparent stimulus the disparity gradients (disparity differences between neighboring points) are very steep (Burt & Julesz, 1980; McKee & Verghese, 2002). Thus, stereo-transparency is not only ecologically relevant but it also poses a challenge for computational theories of stereopsis, which often assume that surfaces are generally smooth and continuous (Hayashi et al., 2004; Marr & Poggio, 1976; Prazdny, 1985; Watanabe & Fukushima, 1999).

Previous studies have shown that observers are capable of perceiving multiple overlaid stereo-transparent planes in RDS (Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998; Lankheet & Palmen, 1998; McKee & Verghese, 2002; Parker & Yang, 1989; Tsirlin, Allison, & Wilcox, 2008; Wallace & Mamassian, 2004; Weinshall, 1993). Compared with the perception of opaque surfaces in an RDS, the perception of transparent surfaces requires longer presentation times and is more difficult to establish reliably (Akerstrom & Todd, 1988; Tsirlin, Allison, & Wilcox, 2008; Wallace & Mamassian, 2004). It has also been reported that increases in total dot density and inter-plane disparity beyond a critical value have a detrimental effect on observers' ability to resolve stereo-transparency (Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998; Tsirlin, Allison, & Wilcox, 2008; Wallace & Mamassian, 2004).

* Corresponding author. Fax: +1 416 736 5857.
E-mail address: itsirlin@yorku.ca (I. Tsirlin).

0042-6989/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2011.11.013

Researchers studying the effects of dot density on stereo-transparency have consistently distributed the dots equally across the random-dot planes. Pilot experiments on stereo-transparency in our laboratory revealed that an interesting phenomenon occurs when the distribution of dots is not balanced. We measured the amount of disparity required to perceive stereo-transparency when the two planes used to generate a transparent RDS were composed of different dot densities. We found that a larger inter-plane separation was required to perceive stereo-transparency when the front (closer) plane was sparser than the back (farther) plane than in the converse situation. To investigate this phenomenon further we conducted a series of experiments in which we systematically varied the distribution of dots in stereo-transparent RDS. The results show there is a reliable asymmetry, and that it is linked to a difference in the signal strength generated by the transparent (front) and opaque (back) surfaces. To help us explore the possible neural basis for this phenomenon, and to link it to previous research on this topic we developed a computational model. This model, based on a network of excitatory and inhibitory inter-neural connections and figure-ground processes, was consistent both with our psychophysical data and results reported by others.

2. Experiment 1

In this experiment we document the asymmetry in pseudo-transparency described above. We varied the distribution of dots on two overlaid planes of an RDS from 10:90 percent (front:back) to 90:10 and asked observers to adjust the depth separation between the planes until they could just perceive the planes as two well-segregated surfaces.

2.1. Apparatus

Scripts for stimulus generation and presentation were created and executed on a G4 Power Macintosh using Python 2.3 and OpenGL libraries for Python, under Mac OS X 10.3. Stimuli were presented on a pair of CRT monitors (Clinton DS2000HB, 14.25″ × 10.7″) arranged in a mirror stereoscope at a viewing distance of 0.6 m. The monitors were calibrated to compensate for curvature. The resolution of the monitors was set to 1024 × 768 pixels and the refresh rate to 100 Hz. At this resolution and viewing distance, each pixel subtended 1.9′ of visual angle. Observers used a chin rest to stabilize head position during testing. The experimental room was completely dark except for the stereoscopic display.

2.2. Observers

Five experienced observers participated in the experiment. IT and LW are authors, while the other three KF, DS and AS, were naive as to the purpose of the study. All observers had normal or corrected-to-normal visual acuity and good stereo acuity as assessed by the Randot™ test.

2.3. Stimuli

Stimuli were 12.6° × 12.6° RDS, composed of 7.6 × 7.6 arc min dots. The average luminance of the stimuli was 10 cd/m^2 and the Michelson contrast was 99%. Each stereogram, when fused, depicted two overlaid planes of dots (see Fig. 1A). The plane closest to the observer was presented at the screen depth and the other plane was presented with uncrossed disparity with respect to the screen. Antialiasing was used to achieve subpixel positioning. When the disparity between the planes was adjusted, the back plane appeared to move closer to or farther away from the front plane. The algorithm used to position the dots in the stereogram is described elsewhere and ensured that there was no overlap of dots on the two planes (Tsirlin, Wilcox, & Allison, 2010). The dots were black on a gray background (the rest of the screen was the same color as the background).

The overall density of dots was held constant at 18.9 dots/deg^2 across all conditions. The relative density of the overlaid planes was varied from 10% of the total dots on the front plane, with 90% on the back plane to 90% of the dots on the front plane and 10% on the back plane, in steps of 20%. This resulted in five test conditions, namely front to back ratio of 10:90, 30:70, 50:50, 70:30 and 90:10. See Figs. 1A and 2 for an illustration of the stimulus.

2.4. Procedure

Observers were shown stimuli depicting two overlaid random-dot planes with different relative densities (see Fig. 2). At the beginning of each trial the planes had either a small disparity of 22.8 arc sec between them and appeared to all observers as a single plane, or a relatively large disparity of 7.6 arc min that for all observers created a percept of two well-segregated planes. The observers were asked to adjust the depth separation between the planes, in steps of 22.8 arc sec, until a coherent percept of two well-segregated surfaces was achieved or was just lost, depending on the starting point of the trial. The observers were instructed to adjust the relative depth until they could clearly distinguish two separate surfaces as opposed to being able to discriminate differences in the depth of the stimulus dots (which can be perceived at very small depth separations). The RDS dots were re-positioned (i.e. a different random sample of dot locations was used) every time an observer adjusted the depth of the planes. To assist the observers in their judgments, before testing, the observers were shown several examples of well-segregated transparent RDS.

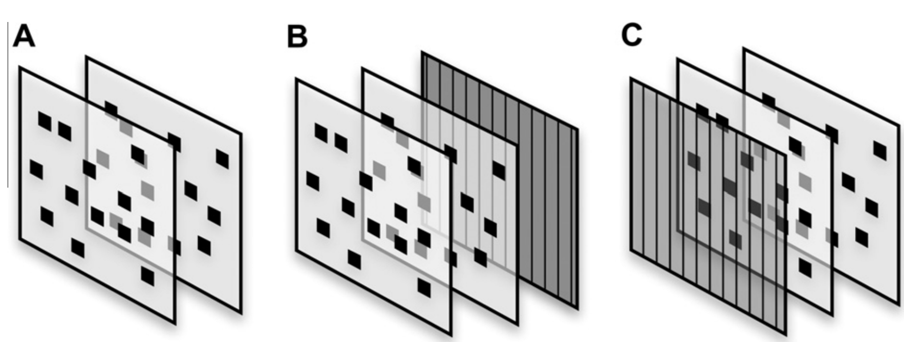

Fig. 1. A schematic representation of stimuli used in our experiments. (A) Experiment 1 – pseudo-transparent stereogram with two planes. (B) Experiment 2 – pseudo-transparent stereogram with an opaque striped surface added behind the far random-dot plane. (C) Experiment 3 – pseudo-transparent stereogram with a transparent striped surface added in front of the near random-dot surface.

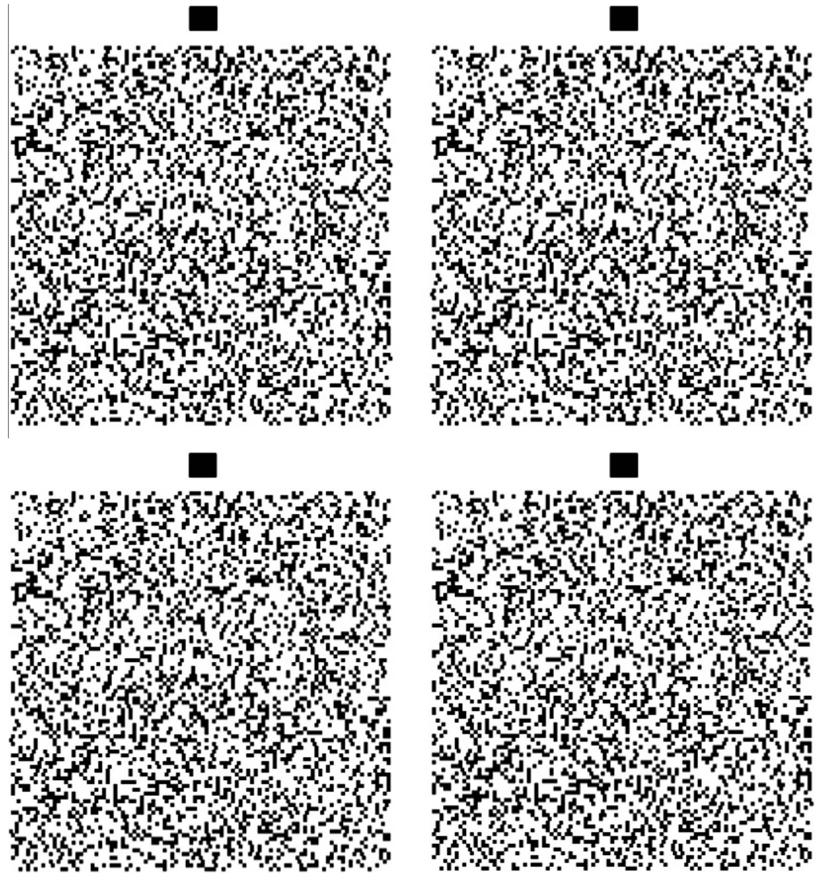

Fig. 2. Example of stimuli used in Experiment 1. For crossed fusion the top panel shows the 90:10 case (front plane 90% of the dots) and the bottom panel shows the 10:90 case. For divergent fusion the depth order is reversed.

Observers were allowed to shift their gaze freely during a trial and had unlimited viewing time. Twelve measurements for each of the conditions were acquired from each observer.

2.5. Results and discussion

The mean responses of the five observers in Experiment 1 are shown in Fig. 3 (blue[1] line, square symbols). There was an asymmetry between the amount of disparity required to perceive stereo-transparency when the front plane contained 10% of the dots (mean = 5.7 arc min) and when the front plane contained 90% of the dots (mean = 3.5 arc min). For stimuli with sparse front planes, more disparity between the overlaid planes was required to create a coherent percept of stereo-transparency. A repeated measures *t*-test showed that the disparity required to see stereo-transparency in the 10:90 condition was significantly greater than that in the 90:10 condition ($t(4) = 2.84$, $p < 0.05$). All *t*-tests in the article were repeated measures and one-tailed (unless specified otherwise) and the alpha levels were 0.05. The results of all *t*-tests were confirmed with two non-parametric rank-based tests (Wilcoxon and BDM).

In this experiment the total number of dots was held constant while the distribution of dots across planes was manipulated. It is possible that the density of the sparser plane was not sufficient to support a percept of a surface. This could have been a potential confound in the observed perceptual asymmetry. We tested this by repeating Experiment 1 with two observers using an RDS with an overall dot density of 28.3 dots/deg^2, which is 50% higher than the original stimulus. The data shown in Fig. 4 were very similar to the original results (Fig. 3) and clearly showed the same perceptual asymmetry.

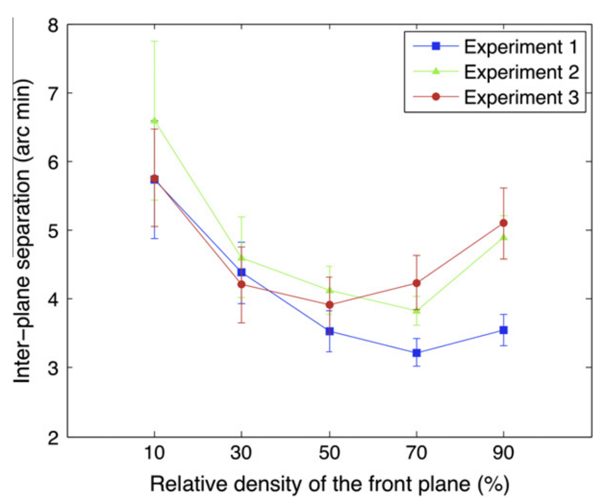

Fig. 3. Mean results of Experiments 1, 2 and 3 for five observers. The inter-plane separation required to perceive stereo-transparency is plotted as a function of the relative density of the front plane of the stereogram. The blue line with square marks represents the results of Experiment 1, the green line with triangular marks represents the results of Experiment 2 and the red line with disc marks represents the results of Experiment 3. Error bars show ±1 standard error of the mean.

While the occurrence of this asymmetry may be initially surprising, the phenomenology of stereo-transparency in these

[1] For interpretation of color in Figs. 3, 4 and 8–10, the reader is referred to the web version of this article.

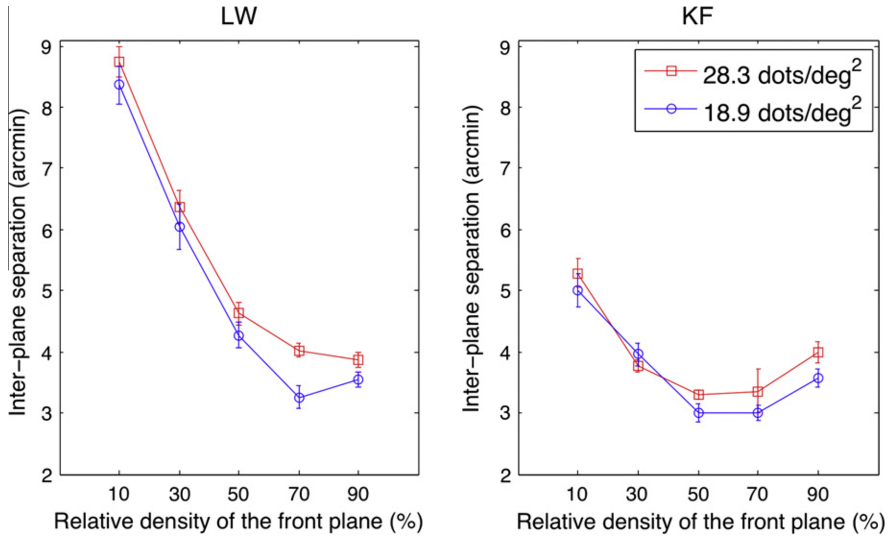

Fig. 4. Experiment 1 repeated with stimuli with larger density for two observers. Inter-plane separation required to perceive stereo-transparency is plotted as a function of the relative density of the front plane of the stereogram. The blue line with disc symbols represents the results of Experiment 1 and the red line with square symbols represents the results of the control experiment with larger density. Error bars show ±1 standard error of the mean.

stimuli provides a potential explanation. In stereo-transparent stimuli like the ones used in Experiment 1, only the plane closest to the observer is perceived as transparent. The back plane appears opaque because the blank regions between the dots defining the two planes, which have no explicit disparity signal, are perceived to lie at the depth of the back plane. We have confirmed this percept in a control experiment where observers reported the depth of inter-dot (blank) regions in randomly selected patches in the stimuli of Experiment 1 (as well as those of Experiments 2 and 3). The patches were specified to the observer using a white outline, which appeared on the screen for 5 s and then disappeared. The observers then evaluated the perceived depth (or depths) of the target area and responded verbally. Three observers, who were naïve as to the purpose of the experiment, participated in the study. Each observer completed 10 trials for each type of stimulus. All observers perceived the blank regions at the depth of the back random-dot plane in stimuli of Experiment 1.

It is conceivable that a stronger response is produced in the population of disparity selective neurons tuned to the depth of the back plane due to this 'default' perceptual assignment of the blank regions, when both planes have equal density. When the front plane has a smaller dot density than the back plane, the signal-strength of the back plane is intensified even more in comparison to the front plane. The increasing difference in the strength of the signals generated by the front and back planes could degrade the stereo-transparency percept, since the back plane might mask (or pull) the front plane, requiring larger separations between the planes to perceive them as two separate surfaces. To balance the signal strengths of the two planes and thus reduce the required inter-plane separation, a larger proportion of dots need to be assigned to the front plane. Consequently, the inter-plane disparity required to see stereo-transparency with dense front planes is smaller than that required with sparse front planes, producing the perceptual asymmetry.

3. Experiment 2

We tested the above hypothesis, which we refer to as the signal-strength hypothesis, by introducing a third, striped, surface behind the two random-dot planes (see Figs. 1B and 5). The logic was that the blank areas between the dots would be assigned to the striped background plane, not to the back random-dot plane. According to our signal-strength hypothesis this manipulation should reduce the strength of the signal for the back random-dot plane, thus reducing the asymmetry. Therefore we predicted that for these stimuli the inter-plane disparity required to perceive stereo-transparency should be similar for 90:10 and 10:90 cases.

In a control experiment (described in detail in Experiment 1) we confirmed that the perceived position of the blank regions (between the dots) in the stimuli of Experiment 2 was at the depth of the additional striped surface.

3.1. Methods

We used the same apparatus, observers and procedure as in Experiment 1. The stimuli were modified in the following fashion. An additional surface was added at an uncrossed disparity of 7.6 arc min with respect to the back random-dot plane (see Fig. 1B). The surface was composed of a gray square (darker than the RDS background) $13.3^\circ \times 13.3^\circ$ with 10 vertical stripes randomly positioned in the square (the positions of the stripes were varied for each instance of the stimulus, that is at the beginning of each trial). The stripes, which were 19 arc min wide, 13.3° high and colored light gray, provided a strong horizontal disparity signal for the additional surface. The disparity between the back random-dot plane and the striped surface remained fixed at 7.6 arc min as the observers adjusted the depth of the back dot plane (so the striped surface was re-positioned in depth as the back RDS surface was adjusted).

3.2. Results and discussion

The results of Experiment 2 are shown in Fig. 3 (green line, triangular symbols) along with those of Experiment 1. One important difference between the results of Experiments 1 and 2 is the increase in the inter-plane depth separation required to perceive stereo-transparency in stimuli with sparse back planes (90:10) in the presence of the background surface as predicted by the signal-strength hypothesis. The disparity required to see stereo-transparency in the 90:10 condition in the stimuli of Experiment 2 (mean = 4.9 arc min) was significantly larger than that in Experiment 1 (mean 3.5 arc min) ($t(4) = 6.94$, $p < 0.01$). Importantly when the background surface was present, there was no significant difference between the disparity required to see stereo-transparency

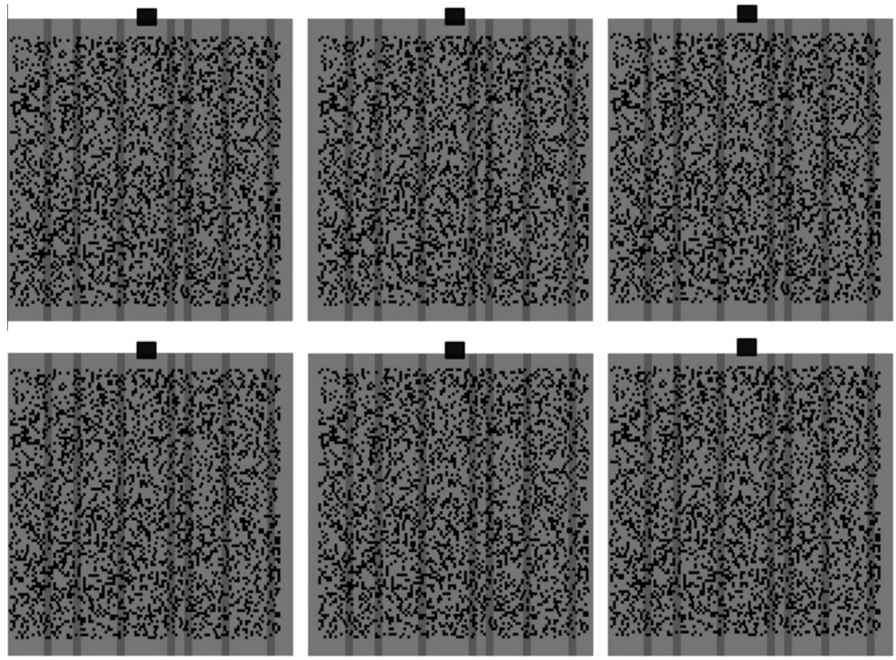

Fig. 5. Example of stimuli used in Experiment 2. The left and the central columns are arranged for crossed fusion and the central and the right columns for divergent fusion. The top panel shows the 90:10 case (front plane 90% of the dots) and the bottom panel shows the 10:90 case.

in the 10:90 (mean = 6.6 arc min) and the 90:10 (mean = 4.9 arc min) conditions ($t(4) = 1.63$, $p = 0.089$).

4. Experiment 3

Another way to reduce the strength of the response to the more distant random-dot plane is to introduce a tinted glass-transparent surface with a strong disparity signal in front of the random-dot surfaces (see Figs. 1C and 6). Tinted glass-transparent surfaces allow light to pass through them clearly but attenuate the luminance or filter the color of the reflected light from the background. Consequently, the blank areas between the dots defining the RDS surfaces should be assigned to two depths simultaneously: that of the far random-dot surface and that of the transparent overlay. According to the signal strength-hypothesis the perceptual asymmetry should be reduced. We tested this proposal in Experiment 3.

In a control experiment (described in detail in Experiment 1) we confirmed that for the stimuli of Experiment 3 the blank areas (between the dots) appeared to be located simultaneously at the depth of the back dot surface and the transparent overlay.

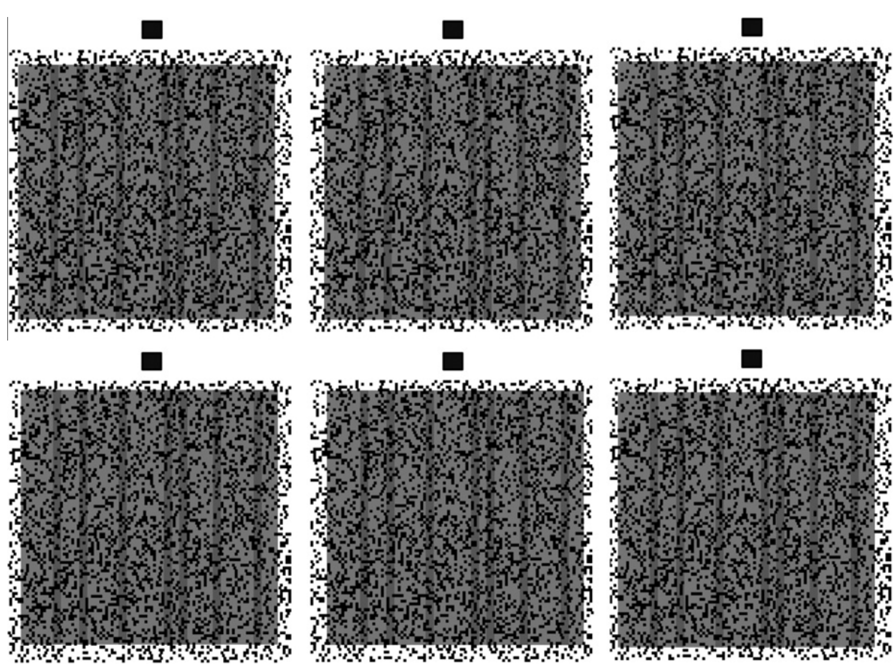

Fig. 6. Example of stimuli used in Experiment 3. The left and the central columns are arranged for crossed fusion and the central and the right columns for divergent fusion. The top panel shows the 90:10 case (front plane 90% of the dots) and the bottom panel shows the 10:90 case.

4.1. Methods

The apparatus, observers and procedure were the same as in Experiment 1. The stimuli of Experiment 1 were modified in the following fashion. An additional surface was added at a crossed disparity of 7.6 arc min relative to the front plane (see Figs. 1C and 6). The surface was composed of a gray square 12° × 12° with 10 randomly positioned vertical stripes (the positions of the stripes were varied for each instance of a stimulus). The stripes were 19 arc min wide by 12° high and colored light gray. The alpha-channel of this surface was set to 0.6 so that the surface appeared transparent. Observers reported that the third surface looked like smoked glass superimposed on the random-dot surfaces. The disparity between the front dot plane and the striped surface remained fixed at 7.6 arc min as the observers adjusted the depth of the back dot plane.

4.2. Results and discussion

The results of Experiment 3 are shown in Fig. 3 (red line, square symbols) along with the results of the previous two experiments. Again, the data show the reduction in perceptual asymmetry predicted by the signal-strength hypothesis. The disparity required to achieve the stereo-transparency percept in the 90:10 condition of Experiment 3 (mean = 5.1 arc min) was significantly larger than that in Experiment 1 (mean = 3.5 arc min) ($t(4) = 2.69$, $p < 0.05$). There was no significant difference found between the disparity required to see stereo-transparency in the 10:90 (mean = 5.76 arc min) and that in the 90:10 condition (mean = 5.1 arc min) ($t(4) = 1.56$, $p = 0.096$).

Since the blank areas in the stereograms of Experiment 3 were perceived both at the depth of the front striped surface and the back random-dot surface, it might be expected that the reduction in the asymmetry would be smaller than in Experiment 2 where the blank areas are assigned only to the back striped surface. However, other factors might have affected performance in this experiment, for instance, inhibitory interactions between neurons tuned to different disparities. Such inhibitory interactions have been proposed in many models of stereopsis (Hayashi et al., 2004; Marr & Poggio, 1976; Prazdny, 1985; Watanabe & Fukushima, 1999) and in studies examining stereo-transparency (Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998). The intensity of inhibition from these connections usually increases with increasing difference in disparity tuning. In the case of the stimuli used in Experiment 3, the transparent overlay would have a much stronger inhibitory influence on the back random-dot plane than on the front random-dot plane since the back plane has a larger relative disparity with respect to the overlay. Hence, this inhibitory effect might weaken the strength of the signal from the back random-dot plane, contributing to the reduction of the asymmetry.

Following the same reasoning, in the stimuli used in Experiment 2, the back striped surface should inhibit the front random-dot plane more than the back random-dot plane. Thus, we might expect to see elevated thresholds for smaller relative densities of the front random-dot plane. There is some evidence for this in the data of Experiment 2 (see Fig. 3). However, in the stimuli of Experiment 2 the blank areas between the dots are only assigned to the disparity of the striped surface and thus the signal of the back random-dot plane is weaker than in Experiment 1 or Experiment 3. Hence even with the inhibitory influence on the front random-dot plane the asymmetry is largely eliminated.

Consequently, the combination of inhibitory influences and the assignment of blank areas to specific disparities could account for the similarity in reduction of the perceptual asymmetry in Experiments 2 and 3. In Section 6 we describe a computational model that combines these mechanisms and is able to predict the psychophysical data.

5. Experiment 4

Given the well-established link between stereo-transparency perception and RDS density (Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998; Tsirlin, Allison, & Wilcox, 2008; Weinshall, 1993), we would expect the difference in signal strength of the two planes in transparent RDS to also affect the perception of relative density. That is, the surface with the stronger signal might be perceived as being denser than the one with the weaker signal. We conducted Experiment 4 to test this hypothesis and in doing so evaluate our signal-strength proposal using a different paradigm.

5.1. Methods

In this experiment we used the stimuli from Experiments 1, 2 and 3 and the disparity between the dot planes was kept constant at 7.6 arc min. The relative density of the front surface was varied from 30% to 70% in steps of 2% resulting in 63 different stimuli (3 configurations × 21 relative densities). Observers (the same as in previous experiments) made 20 responses for each type of stimulus. For each stimulus the observers were asked to indicate which random-dot plane, the front or the back, appeared to contain more dots (i.e. was more dense). If the asymmetry manifests itself in these density judgments then the two planes should appear equally dense when in fact the front plane has more dots than the back plane.

5.2. Results and discussion

For each observer and stimulus type we fitted the data with a Weibull function with percent 'front plane denser' responses as a function of the relative density of the front plane. From this function we then estimated the point of subjective equality (PSE), which was defined as the point where the observer was equally likely to say the front plane was denser or that the back plane was denser (50% 'front plane denser responses'). Fig. 7 shows the mean PSE's for the three conditions.

The mean PSE corresponding to stimulus used in Experiment 1 was 58.6% (dots on the front plane) indicating that the front dot plane needed more dots than the back dot plane in order to be perceived as equally dense. This PSE was significantly greater than 50% (one-sample $t(4) = 5.48$, $p < 0.01$). Thus a perceptual asymmetry similar to that reported in Experiment 1 is seen in density judgments for RDS planes. This asymmetry was reduced by using either the stimulus from Experiment 2 that had an additional striped

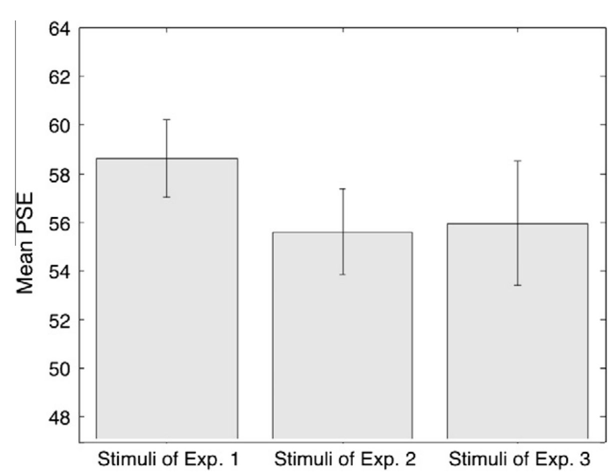

Fig. 7. Results of Experiment 4 for five observers and three types of stimuli. Each bar depicts the mean of the observers' PSEs for each type of stimulus. The error bars show ±1 standard error.

surface added behind the random-dot planes or the stimulus from Experiment 3 that had an additional translucent surface in front of the random-dot planes. The PSE for stimuli from Experiment 1 was significantly higher than the PSE for stimuli from Experiment 2 (58.6 and 55.6, respectively) ($t(4) = 2.47$, $p < 0.05$) but there was no significant difference between PSEs for stimuli from Experiments 1 and 3 (58.6 and 55.9, respectively) ($t(4) = 1.36$, $p = 0.12$). The PSE for stimuli from Experiment 2 was still significantly different from 50% (one-sample $t(4) = 3.21$, $p < 0.05$) but the PSE of Experiment 3 was not significantly different from 50% (one-sample $t(4) = 2.33$, $p = 0.079$).

These results show that the perceptual asymmetry observed in Experiment 1 is also present in the subjective judgment of density in transparent RDS. As in Experiments 2 and 3, adding surfaces either in the front or in the back reduced the asymmetry in density judgments to a certain degree. These findings suggest that similar mechanisms underlie the asymmetries observed with relative density and stereo-transparency percepts.

6. Simulation experiments

In this section, we present a simple computational model that encapsulates an architecture of neural connectivity that can support the disparity spreading processes described by the signal-strength hypothesis. The goal is to assess whether such a neural network can produce the psychophysical results obtained in Experiments 1–3 and thus verify the plausibility of the signal-strength hypothesis. Note that the model is not concerned with the earliest stages of stereopsis including solving the correspondence problem, but rather with the later stages of disparity propagation and surface creation. Recall that the signal strength hypothesis states that the blank areas between the dots of the RDS are assigned to the disparity of the surface that is considered to be the background (and/or the transparent foreground) through the process of disparity interpolation. Accordingly, the model contains three neuronal mechanisms that are capable of propagating the disparity signal in agreement with the figure-ground interpretation of the scene:

1. *Excitatory connections*: The assignment of the blank regions between the dots of the RDS—which are devoid of features that could provide a disparity signal—to a particular depth, suggests that the disparity signal spreads across the blank areas. Disparity interpolation and extrapolation across areas of uniform luminance has been demonstrated in numerous experiments (e.g. Julesz, 1971; Takeichi, Watanabe, & Shimojo, 1992; Wilcox & Duke, 2003, 2005; Yang & Blake, 1995). The propagation of disparity can be mediated by excitatory connections between neurons tuned to similar disparities. This type of connection has been used in models of stereopsis to implement the smoothness constraint (Hayashi et al., 2004; Marr & Poggio, 1976; Prazdny, 1985; Watanabe & Fukushima, 1999). The excitatory connections used in these models are normally short-range (between nearest neighbors) since long-range excitatory connections are likely to span regions that belong to different objects located at different depths. The relatively short extent of lateral excitatory connections between disparity detectors is also supported by psychophysical data (McKee & Mitchison, 1988; Mitchison & McKee, 1987). Earlier models showed that short-range excitatory connections also support the perception of stereo-transparent RDS (Pollard, Mayhew, & Frisby, 1985; Prazdny, 1985) and can successfully account for disparity interpolation and disparity pulling (Lehky & Sejnowski, 1990). In our model the excitatory connections extend from each neuron to its immediate neighbors both on the same and on adjacent disparity planes.

2. *Inhibitory connections*: Theoretically, the disparity signal could spread to the blank areas in our stimuli from the dots of both the front and the back planes (as well as the additional surfaces). However, since we tend to perceive the blank areas only at one depth in transparent random-dot stereograms (in absence of an extra transparent plane as in Experiment 3), it is likely that there is a competition between the neurons tuned to the disparities of the different planes (see also the discussion in Section 4.2). Competitive interactions can be implemented as long-range inhibitory connections between neurons tuned to sufficiently different disparities but with similar position tuning. This type of inhibitory connection was used in classical models of stereopsis to implement the uniqueness constraint (Hayashi et al., 2004; Marr & Poggio, 1976; Prazdny, 1985; Watanabe & Fukushima, 1999). They have also been proposed in models of stereo-transparency to account for the deterioration of the stereo-transparency percept as the inter-plane disparity increases beyond a certain value (Akerstrom & Todd, 1988; Gepshtein & Cooperman, 1998). In these models, the intensity of inhibition increases with increase in the difference in disparity tuning.[2] We adopted this type of inhibitory connection in our model.

3. *Figure-ground processes*: The assignment of blank areas to the depth of a particular plane (or two planes) reflects the visual system's resolution of figure ground relationships. The effect of figure ground segregation processes on disparity interpolation has been demonstrated well by Wu et al. (1998). These authors found that the perceived depth of texture dots with ambiguous disparity depended on the context imposed by the interpretation of the figure-ground configuration (Fig. 5 in their article). The importance of a high-level feedback process in perception of stereo-transparency might explain the relatively large latencies required to perceive stereo-transparency compared to depth in opaque stimuli (Akerstrom & Todd, 1988; Tsirlin, Allison, & Wilcox, 2008) as well as in the deterioration of the stereo-transparency percept with an increasing number of overlaid planes as suggested by Tsirlin, Allison, and Wilcox (2008). In our model, this top-down influence is implemented as an excitatory input provided by a neuron higher in the hierarchy to the neurons of the plane that is perceptually assigned as the background (and/or transparent foreground). This extra input is a function of processing time, to reflect the latency associated with higher-level feedback processes.

The implementation details of our model and all the parameters are provided in Appendix A. To evaluate our model we simulated the psychophysical experiments described in this paper using Matlab 7.9 for OS X. For a given relative dot density (e.g. 10% front 90% back) a proportional number of neurons at each of the two disparity planes z_i and z_j were assigned an initial value of 1 (firing neurons) and all the other neurons in the network were initialized with 0 (resting neurons). The positions of the dots on the disparity planes were chosen randomly. Then the network was allowed to run n iterations and the average normalized response of the network at each disparity was recorded. Next, this procedure was repeated with a larger inter-plane separation (with the back plane moving away from the front plane). This was done to simulate an observer moving the planes away from each other in depth as

[2] Note that in a network that combines such inhibitory connections with short-range excitatory connections, pulling two stereo-transparent surfaces apart initially would free them of the averaging effect of the excitatory connections across planes with similar disparity. At small inter-plane disparities, the weights on inhibitory connections are relatively low and hence the perception of stereo-transparency should be facilitated. As the surfaces continue to be pulled apart, the inhibitory influence becomes stronger, eventually disrupting the percept of stereo-transparency. This is consistent with the data of Tsirlin, Allison, and Wilcox (2008) where as the inter-plane disparity initially increased there was an improvement in performance. Then, after a peak disparity was reached, performance declined with further increases in inter-plane disparity. Thus long-range inhibitory connections spanning different disparities coupled with short-range excitatory connections are a suitable neural substrate for stereo-transparency perception.

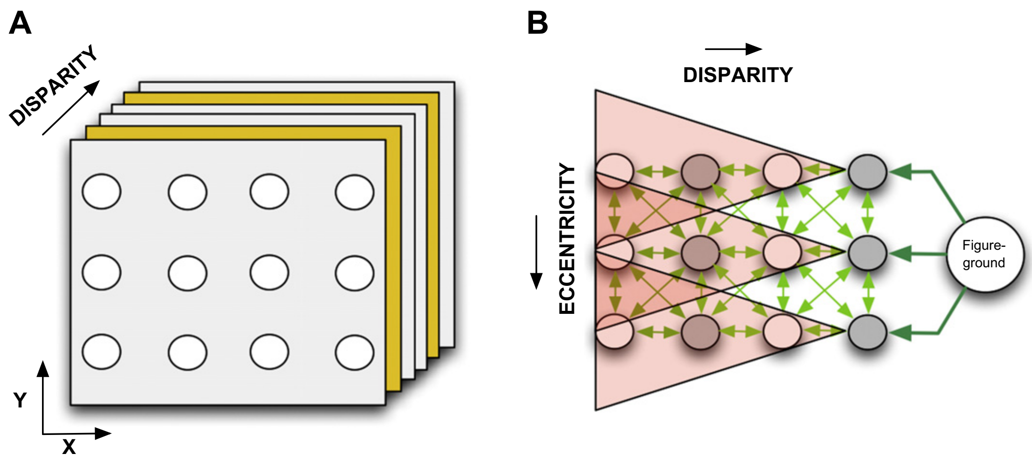

Fig. 8. The architecture of the neural model. (A) The neural network matrix. The white circles are the neurons and the different planes represent the populations of neurons tuned to different disparities. Planes marked in yellow represent the populations responding to the two random-dot surfaces. (B) Connectivity of the neural network. Green arrows show reciprocal excitatory connections. The red cones show the inhibitory connections (for only three neurons on the back plane) between neurons tuned to different disparities. Activated neurons on two disparity planes are shown in dark gray. The excitatory input from the higher-level figure-ground process that provides the bias to the back plane is shown in dark green arrows originating from a higher-level neuron.

in the psychophysical experiments.[3] The planes were considered well segregated and the percept of stereo-transparency established, when two (or three for simulations of Experiments 2 and 3) distinct local maxima emerged in the response function of the network, indicating the network detected two (or three) separate surfaces. The relative disparity of the planes at this point was taken as the inter-plane separation required to perceive stereo-transparency. Thirty estimates of this inter-plane separation were obtained by repeating the above procedure to avoid abnormal results due to outlier dot distributions. The experiment was repeated for relative densities 90:10 to 10:90 and the mean inter-plane separations were plotted as a function of relative density (see Fig. 9).

The model was evaluated under three conditions corresponding to Experiments 1–3. In the first condition, only the two random-dot planes were initialized and the bias (the excitatory input from the figure-ground process) was applied to the back random-dot plane. In the second condition, an additional plane was added at a fixed uncrossed disparity ($d = 3$) with respect to the back dot plane. On the additional plane, several columns of neurons were activated (given a value of 1) to simulate the stripes of the third plane in Experiment 2 (density of the stripes was matched to that in psychophysical stimuli). The bias was then applied to this additional plane. In the third condition, the additional plane was added with a fixed crossed disparity ($d = -3$) with respect to the front dot plane. The bias was applied both to the additional plane and to the back dot plane in accordance with the perceived depth of the blank areas in stimuli of Experiment 3. In this case the bias was distributed evenly across the two planes (50% on the additional plane and 50% on the back dot plane).

In these simulations it was assumed that the correspondence problem was resolved correctly, and the neurons with the appropriate disparity and location tuning are responding to the dots of the RDS. We have also performed simulations where the input to the above model was obtained by applying a cross-correlation algorithm to actual stimuli replicating those of Experiments 1, 2

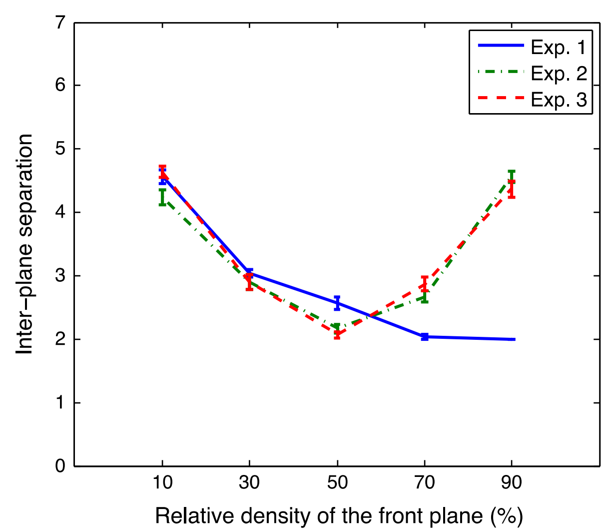

Fig. 9. Results of the computational experiments. The smallest inter-plane separations for which the response curves showed two peaks are plotted as a function of the relative density of the front plane. The blue solid line shows the results for a simulation of Experiment 1. The green dashed and dotted line shows the results for the simulation of Experiment 2. The red dashed line shows the results for the simulation of Experiment 3. Error bars show the ±1 standard error of the mean (the S.E. of the 30 simulation trials, see text for details).

and 3 (but smaller in size). Pre-filtered matches from the cross-correlation algorithm provided the initial activation matrix for the model. The results of these simulations were similar to those presented in Fig. 9.

Fig. 9 shows the smallest inter-plane separations for which the response curves of the neural population exhibited two/three distinct peaks. The separations are plotted as a function of the relative density of the front plane with respect to the back plane. The blue solid line shows the data from the simulation of Experiment 1. There is a clear asymmetry between the 90:10 and the 10:90 cases similar to the psychophysical results. The green dashed and dotted line shows the data from the simulation of Experiment 2 and the red dashed line shows the data from the simulation of Experiment 3. In both cases the asymmetry between the 10:90 and the 90:10 is eliminated. These data closely resemble the

[3] In the psychophysical experiments the observers were free to move their eyes and therefore the absolute disparities of the two planes were not fixed. However, the relative inter-plane disparity was not affected by changes in fixation. In the model the front plane remains at the same disparity for simplicity of implementation. Since the performance of the model relies on the relative disparity of the planes, the absolute disparity of the front plane should not matter.

psychophysical data demonstrating that the proposed neural architecture and thus the signal strength hypothesis can account, at least in part, for our psychophysical findings.

7. General discussion

The experiments presented in this manuscript demonstrated a perceptual asymmetry in stereo-transparency. Experiment 1 showed that a larger inter-plane disparity is required to perceive stereo-transparency in stimuli with a higher proportion of dots on the back plane than in stimuli with a higher proportion of dots on the front plane. We hypothesized that since the blank regions between the dots in the transparent RDS are perceived at the depth of the back random-dot plane, more neurons tuned to its disparity are activated than those tuned to the disparity of the front random-dot plane, creating a stronger signal for the back plane. A large difference in the strength of the signals for the two planes could degrade the stereo-transparency percept, since the back plane might mask or pull the front plane. This in turn necessitates larger inter-plane disparity to perceive the two surfaces as distinct. The difference in signal strength would be intensified when the front random-dot plane is sparser than the back random-dot plane and attenuated when the front plane is denser than the back plane yielding the observed perceptual asymmetry.

In support of this hypothesis, Experiments 2 and 3 showed that when the signal strength of the back random-dot surface was reduced (by adding an extra plane in front or behind the random-dot planes) the perceptual asymmetry was also reduced.

Gepshtein and Cooperman (1998) and Akerstrom and Todd (1988) proposed a combination of excitatory and (primarily) inhibitory connections to account for the effect of density and disparity on stereo-transparency. Tsirlin, Allison, and Wilcox (2008) suggested that some of the properties of stereo-transparency perception reported in the literature can be simply explained by a well-known characteristic of disparity detectors, namely the size-disparity correlation. However, the phenomena described in this article require a more complex neural mechanism such as a network of excitatory and inhibitory inter-neural connections and the involvement of higher-level processes responsible for figure-ground segregation. These higher-level processes can control the propagation of the disparity signal across blank areas in the RDS by the means of a network of excitatory and inhibitory connections. In Section 6 we presented a computational model incorporating these neural mechanisms that was able to account for our psychophysical data.

To our knowledge, there have been no previous studies of the nature and cause of this density-based perceptual asymmetry in stereo-transparency. However, we believe that there is preliminary evidence of this asymmetry in the experiments of Gepshtein and Cooperman (1998). They presented observers with RDS depicting a sparse cylinder behind a pseudo-transparent plane and asked them to indicate the orientation of the cylinder (horizontal or vertical) as they varied the density of the plane. Among other things, they found that as the inter-plane separation increased the maximum density of the plane for which the orientation of the cylinder could be discriminated (limiting density) decreased. More importantly, when the sparse cylinder was placed in front of the plane the limiting density was much smaller (for all inter-plane disparities) than when the cylinder was presented behind the plane. The authors noted that when the cylinder was placed in front of the plane "the observers reported that discrimination of the cylinder orientation was more difficult" (p. 2929).

Gepshtein and Cooperman proposed that this asymmetry reflected a difference in connectivity of neurons tuned to crossed and uncrossed disparities resulting from a natural bias to fixate such as to bring more details within the region of uncrossed disparities (in their experiments the observer always fixated at the depth of the dense plane).

The signal-strength hypothesis provides an alternative account for their findings. As in our stimuli, the farther of the two surfaces of the RDS in Gepshtein and Cooperman's stimuli appeared opaque due the assignment of the blank regions to the disparity of the back plane. Hence, when the sparse cylinder was at the back, it provided a stronger signal than when it was in front of the dense plane. Consequently, when the cylinder was behind the plane, the maximum density of the plane at which stereo-transparency was still achievable was larger than when the cylinder was in front of the plane.

To test this hypothesis we simulated Gepshtein and Cooperman's experiment using the computational model presented in Section 6. In their experiment the density of the cylinder remained constant, while the density of the plane was changed to measure the limiting density. For simplicity, we represented the cylinder as a flat plane of a constant density (as in the original experiment it was set to 1%). We then increased the density of the second plane in steps of 1% and evaluated the response of the model for each density configuration. Gepshtein and Cooperman measured the limiting density of the plane for which stereo-transparency could not be achieved and the orientation of the cylinder was not discernable. In the case of the model, we assumed that the percept of stereo-transparency would be maintained while two distinct peaks existed in the response of the population of neurons. Thus, the limiting density was the smallest density for which the response of the population contained only one peak. The limiting densities were computed for several inter-plane separations (6–11) in order to confirm that the limiting density decreased with increase in inter-plane separation as in the original study. The above scenarios were tested for two surface configurations, "cylinder" behind of the plane and "cylinder" in front of the plane. The parameters in the model were the same as before. As can be seen in Fig. 10, the limiting densities are much smaller when the dense plane is behind the sparse "cylinder" than when it is in front at all inter-plane disparities. As in the original study, limiting densities decrease with increase in inter-plane separation. These data show that the proposed model also predicts the type of asymmetry observed by Gepshtein and Cooperman (1998).

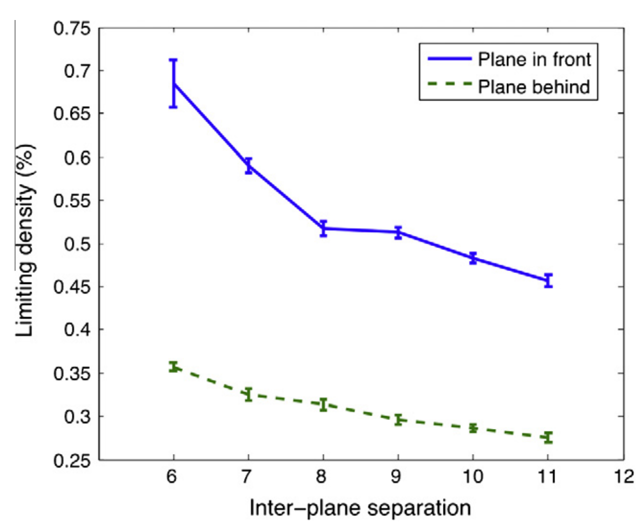

Fig. 10. Results of the computational experiments using Gepshtein and Cooperman's (1998) setup. The largest plane densities for which the response curves of the neural population showed two peaks are plotted as a function of the inter-plane separation between the plane and the cylinder. The solid blue line shows the results for a simulation where the plane was in front of the cylinder. The green dashed line shows the results for the simulation where the plane was behind of the cylinder.

8. Conclusions

This study demonstrates a perceptual asymmetry in stereo-transparency in which larger inter-plane separations are required to perceive transparency in an RDS when the front dot plane is sparse relative to the back than in the converse situation. We propose that interpolation of the blank areas at the depth of the back surface increases the signal strength of this surface, creating a bias towards the far plane. This degrades the percept of stereo-transparency and requires larger inter-plane separations to properly segregate the two dot surfaces. This hypothesis was supported by experiments designed to reduce the perceptual assignment of the blank areas to the rear surface.

It is plausible that these biases are mediated by a network of connections between disparity-tuned neurons. Short-range excitatory connections provide the necessary circuitry for depth propagation and disparity interpolation across surfaces. Inhibitory connections between neurons with different disparity tuning form the basis for competitive interactions between potential disparities in the blank areas. We also proposed that the bias towards the back plane is a result of a top-down influence based on figure-ground segregation. A computational model incorporating the proposed neural architecture produced results comparable to the psychophysical data from the present experiments and those reported in the literature.

Taken together our psychophysical and computational experiments suggest that the perceptual asymmetry described here, and stereo-transparency in general, depends not only on the properties of disparity detectors (Tsirlin, Allison, & Wilcox, 2008) but also on inter-neural interactions of disparity-tuned cells and higher-level processes governing figure ground segregation.

Acknowledgments

This work was supported by NSERC Discovery Grants to R. Allison and L. Wilcox and NSERC Canada Graduate Scholarship to I. Tsirlin.

Appendix A. Details of the computational model

A.1. Neural network

In our model, neurons are organized in a 3D matrix which represents the two spatial dimensions (x, y) and disparity (z) (see Fig. 8A). A neuron has its receptive field (RF) centered at a particular location (x_i, y_i) and is tuned to a particular disparity z_i. The neurons are interconnected with excitatory and inhibitory connections as described below. All neurons are initialized to state 0 or rest (effectively a neutral membrane potential), except for those that represent the hypothetical dots of the RDS surfaces. That is, the binocular matching stage is skipped entirely since the model focuses on the propagation of disparity signals across the network as a function of time (however, as described in Section 6, the model performed similarly when the output of a cross-correlation algorithm was used to initialize the network). It is assumed that binocular matching is performed successfully and the neurons with RF locations and preferred disparities that correspond to the locations and disparities of the random dots of the stimulus are assigned an initial activity state of 1 (firing neurons). Further calculation of neural activity is done in n iterations. At each iteration two matrices are used. One matrix reflects the state of the neural network at time $t-1$ and the other at time t. The state p_t of a neuron (x_i, y_i, z_i) at time t is computed in the following fashion:

$$p_t(x_i, y_i, z_i) = f(u) \tag{1}$$

where

$$f(u) = \begin{cases} 1 \rightarrow u > \delta \\ 0 \rightarrow \text{otherwise} \end{cases} \tag{2}$$

and

$$u(x_i, y_i, z_i) = p_{t-1}(x_i, y_i, z_i) + E(x_i, y_i, z_i) - I(x_i, y_i, z_i) \tag{3}$$

$E(x_i, y_i, z_i)$ and $I(x_i, y_i, z_i)$ are the excitatory and inhibitory inputs for neuron (x_i, y_i, z_i) respectively and δ is a constant specifying the activity threshold.

The excitatory inputs are collected from the local support neighborhood such that each neuron sends an excitatory input to its immediate neighbors (both across space and disparity). The excitatory connections propagate the disparity signal across space to account for disparity interpolation. The excitatory inputs are computed in following fashion:

$$E(x_i, y_i, z_i) = \sum_{x_j, y_j, z_j \in \Phi_E} \alpha * p_{t-1}(x_j, y_j, z_j) \tag{4}$$

where Φ_E is the support neighborhood (a cube of size 3×3 around neuron (x_i, y_i, z_i)) and α is the weight assigned to the excitatory input. To simulate the back plane bias, imposed by the figure-ground segregation process as described in the Computational Experiments section, an extra excitatory input k is added to the neurons tuned to the disparity of the plane being biased (see Fig. 8B). Since this extra excitatory input is provided by a higher-level feedback process we assume that there is a temporal latency in the assignment of the bias. We modeled this latency as an exponential function of the number of iterations:

$$E(x_i, y_i, z_i) = \left[\sum_{x_j, y_j, z_j \in \Phi_E} \alpha * p_{t-1}(x_j, y_j, z_j) \right] + k \left[\frac{\exp(ii)}{\exp(n)} \right] \tag{5}$$

where n is the total number of iterations and ii is the current iteration number. Thus the applied bias is very small in the first iteration and grows progressively larger with each iteration.

We chose to model the inhibitory connections according to the disparity gradient limit as suggested by Gepshtein and Cooperman (1998). In their model each neuron (x_i, y_i, z_i) sends inhibitory signals to neurons that are tuned to disparities and spatial locations that violate the disparity gradient limit of 1 with respect to (x_i, y_i, z_i) (Burt & Julesz, 1980). The inhibitory region is thus shaped like a cone extending along the disparity (z) dimension (see Fig. 8B). The inhibitory zones grow wider as disparity increases. The inhibitory weights are calculated as follows:

$$I(x_i, y_i, z_i) = \sum_{x_i, y_i, z_j \in \Phi_I} \beta * p_{t-1}(x_i, y_i, z_j) \tag{6}$$

where β is the weight on the inhibitory input and Φ_I is the inhibitory neighborhood computed in accordance with the disparity gradient limit. That is, the inhibitory neighborhood Φ_I for a neuron at position (x_i, y_i, z_i) includes all neurons for which the ratio of the disparity difference Δd to the cyclopean separation ΔS is larger than 1 (Burt & Julesz, 1980):

$$\frac{\Delta d}{\Delta S} = \frac{|z_i - z_j|}{\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} > 1 \tag{7}$$

A.2. Parameter estimation

The network described above requires the specification of several parameters including: excitatory weight α, inhibitory weight β, back plane bias k, action potential threshold δ and the number of iterations n. To establish the range of parameters that can predict the psychophysical data, a preliminary analysis of the model was performed in which we assessed a variety of parameter combinations. In particular, the requirements included a sufficiently large asymmetry between the inter-plane segregation required for two peaks to emerge at 10:90% and 90:10% distributions in the simulation of Experiment 1 and little or no asymmetry in the simulations of Experiments 2 and 3. This analysis showed that a wide range of parameters resulted in the pattern of results observed in Experiments 1–3. The bias parameter k and the inhibitory weight β affected the asymmetry magnitude and the excitatory weight α, inhibitory weight β, number of iterations n and the threshold δ affected the spread of the disparity signal. The values used in the final simulation were within the ranges outlined by the preliminary analysis:

- Number of repetitions $n = 5$.
- Excitatory weight $\alpha = 0.15$.
- Inhibitory weight $\beta = 0.001$.
- Bias $k = 0.9$.
- Action potential threshold $\delta = 0.8$.

Other parameters of the network were as follows:

- Network size – $20 \times 20 \times 19$ ($X \times Y \times Z$).
- The front random-dot plane was always at $Z = 6$ and the back plane was shifted during the evaluation process (see above) from $Z = 7$ to $Z = 14$. In Experiment 2 simulation the third (striped) plane was placed at uncrossed disparity of 3 with respect to the back random-dot plane (hence it moved from $Z = 10$ to $Z = 17$). In the simulation of Experiment 3 the third (striped) plane was place at crossed disparity of -3 with respect to the front random-dot plane (at $Z = 3$).
- Total dot density of the simulated RDS was 30%.

References

Akerstrom, R. A., & Todd, J. T. (1988). The perception of stereoscopic transparency. *Perception and Psychophysics, 44*(5), 421–432.

Burt, P., & Julesz, B. (1980). A disparity gradient limit for binocular fusion. *Science, 208*(4444), 615–617.

Gepshtein, S., & Cooperman, A. (1998). Stereoscopic transparency: A test for binocular vision's disambiguating power. *Vision Research, 38*(19), 2913–2932.

Hayashi, R., Maeda, T., Shimojo, S., & Tachi, S. (2004). An integrative model of binocular vision: A stereo model utilizing interocularly unpaired points produces both depth and binocular rivalry. *Vision Research, 44*, 2367–2380.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.

Lankheet, M. J., & Palmen, M. (1998). Stereoscopic segregation of transparent surfaces and the effect of motion contrast. *Vision Research, 38*(5), 659–668.

Lehky, S. R., & Sejnowski, T. J. (1990). Neural model of stereoacuity and depth interpolation based on a distributed representation of stereo disparity. *Journal of Neuroscience, 10*, 2281–2299.

Marr, D., & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science, 194*(4262), 283–287.

McKee, S. P., & Mitchison, G. J. (1988). The role of retinal correspondence in stereoscopic matching. *Vision Research, 28*, 1001–1012.

McKee, S. P., & Verghese, P. (2002). Stereo transparency and the disparity gradient limit. *Vision Research, 42*(16), 1963–1977.

Mitchison, G. J., & McKee, S. P. (1987). Interpolation and the detection of fine structure in stereoscopic matching. *Vision Research, 27*, 295–302.

Parker, A. J., & Yang, Y. (1989). Spatial properties of disparity pooling in human stereo vision. *Vision Research, 29*(11), 1525–1538.

Pollard, S. B., Mayhew, J. E., & Frisby, J. P. (1985). PMF: A stereo correspondence algorithm using a disparity gradient limit. *Perception, 14*(4), 449–470.

Prazdny, K. (1985). Detection of binocular disparities. *Biological Cybernetics, 52*(2), 93–99.

Takeichi, H., Watanabe, T., & Shimojo, S. (1992). Illusory occluding contours and surface formation by depth propagation. *Perception, 21*, 177–184.

Tsirlin, I., Allison, R. S., & Wilcox, L. M. (2008). Stereoscopic transparency: Constraints on the perception of multiple surfaces. *Journal of Vision, 8*(5), 1–10.

Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2010). Perceptual artifacts in random dot stereograms. *Perception, 39*(3), 349–355.

Wallace, J. M., & Mamassian, P. (2004). The efficiency of depth discrimination for non-transparent and transparent stereoscopic surfaces. *Vision Research, 44*(19), 2253–2267.

Watanabe, O., & Fukushima, K. (1999). Stereo algorithm that extracts a depth cue from interocularly unpaired points. *Neural Networks, 12*, 569–578.

Weinshall, D. (1993). The computation of multiple matching doubly ambiguous stereograms with transparent planes. *Spatial Vision, 7*(2), 183–198.

Wilcox, L. M., & Duke, P. A. (2003). Stereoscopic surface interpolation supports lightness constancy. *Psychological Science, 14*(6), 525–530.

Wilcox, L. M., & Duke, P. A. (2005). Spatial and temporal properties of stereoscopic surface interpolation. *Perception, 34*(11), 1325–1338.

Wu, X., Zhou, Q., Lin, X., & Wang, Y. J. (1998). Stereo capture: Local rematching driven by binocularly attended 3-D configuration rather than retinal images. *Vision Research, 38*, 2081–2085.

Yang, Y., & Blake, R. (1995). On the accuracy of surface reconstruction from disparity interpolation. *Vision Research, 35*(7), 949–960.