Vision Research 123 (2016) 41–45

Contents lists available at ScienceDirect

Vision Research

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

Size matters: Perceived depth magnitude varies with stimulus height

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

[a] Centre for Vision Research, Department of Psychology, York University, 4700 Keele St., Toronto, Ontario, Canada
[b] Centre for Vision Research, Department of Electrical Engineering and Computer Science, York University, 4700 Keele St., Toronto, Ontario, Canada
[c] Eye Movement and Neuroscience Laboratory, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 10 June 2015
Received in revised form 9 April 2016
Accepted 10 April 2016
Available online 19 May 2016

Keywords:
Stereopsis
Depth perception
Disparity
Size–disparity correlation
Energy model

ABSTRACT

Both the upper and lower disparity limits for stereopsis vary with the size of the targets. Recently, Tsirlin, Wilcox, and Allison (2012) suggested that perceived depth magnitude from stereopsis might also depend on the vertical extent of a stimulus. To test this hypothesis we compared apparent depth in small discs to depth in long bars with equivalent width and disparity. We used three estimation techniques: a virtual ruler, a touch-sensor (for haptic estimates) and a disparity probe. We found that depth estimates were significantly larger for the bar stimuli than for the disc stimuli for all methods of estimation and different configurations. In a second experiment, we measured perceived depth as a function of the height of the bar and the radius of the disc. Perceived depth increased with increasing bar height and disc radius suggesting that disparity is integrated along the vertical edges. We discuss size–disparity correlation and inter-neural excitatory connections as potential mechanisms that could account for these results.

1. Introduction

It is well-documented that several aspects of stereoscopic depth perception vary with the scale of the stimulus. Schor and Badcock (1985) and Heckmann and Schor (1989) showed that stereoacuity decreased with decreasing spatial frequency of difference of Gaussians or sinusoidal luminance gratings. Using bars of different widths and the same height, Richards and Kaye (1974) showed that the maximum disparity that resulted in depth perception, and the disparity that produced the greatest depth percept, increased with increasing line width. In a related study, Tyler and Julesz (1980) used planar RDS displays composed of a central target in front of a larger background. They found that the maximum disparity that supported stereopsis increased as the central rectangle increased in size (both width and height).

In other studies, Tyler (1973, 1975) used vertical line stereograms with sinusoidal and square wave depth modulations of varying frequency to show that larger stereoscopic thresholds, fusional limits and upper disparity limits for stereopsis were obtained for lower frequency modulations compared to higher frequency modulations. He also found that for a particular spatial frequency, increasing the number of cycles visible to the observer, and thus the line height, increased the upper disparity limit. Tyler proposed that these effects were due to the existence of a size–disparity correlation, where neurons tuned to large scales encode large disparities and neurons tuned to small scales encode small disparities. According to this account, large disparity detectors require large objects (both in height and width) to fire optimally, but small disparity detectors respond best to fine features. As a result, both the upper disparity limits and discrimination thresholds are lower for smaller objects. Other support for the relationship between disparity selectivity and stimulus size was provided by Felton, Richards, and Smith (1972) who adapted observers to sinusoidal gratings presented with disparity relative to a fixation stimulus. Adaptation to large disparities occurred only for low-frequency gratings (large component width), while adaptation to small disparities occurred only for high-frequency gratings (small component width). Smallman and MacLeod (1994) reported that this size–disparity correlation was also evident at contrast threshold for filtered RDS stimuli, even when vergence was carefully monitored.

In a recent article, Tsirlin, Wilcox, and Allison (2012) reported a new relationship between depth from stereopsis and object size. They found that the *magnitude* of perceived depth between a filled rectangle and a small disc appeared smaller than the depth between the rectangle and a long bar, even when both the disc and the bar had the same relative disparity and width. To our knowledge, this was the first demonstration that object height can affect suprathreshold depth percepts for objects with equivalent disparity.

---

* Corresponding author.
E-mail address: inna.tsirlin@sickkids.ca (I. Tsirlin).

http://dx.doi.org/10.1016/j.visres.2016.04.006
0042-6989/

The aim of the current work is twofold. In Experiment 1, we evaluate the generalizability of Tsirlin et al.'s results to other stimulus configurations and depth estimation methods. In Experiment 2, we assess whether the increase in perceived depth with increasing object height depends on the spatial integration of disparity signals. We find that perceived depth depends on stimulus height regardless of the configuration and the depth estimation method, and that disparity is integrated along the vertical contours of the objects. Finally, we suggest size–disparity correlation and excitatory inter-neural connections as potential mechanisms underlying the integration of disparity signals along vertical stimulus edges.

2. Experiment 1

2.1. Methods

2.1.1. Observers

Six volunteers participated in the experiment. They all had normal (or corrected to normal) visual acuity and stereoacuity of 40 arcsec or less as assessed with the Randot stereoacuity test. All the observers, except for one (IT), were naive to the purposes of the experiment. Informed consent was obtained from each observer before the experiment. This work has been carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

2.1.2. Apparatus

Stimuli were shown using a mirror stereoscope built with a pair of ViewSonic G225f CRT monitors with resolution of 1920 × 1200 pixels and refresh rate of 100 Hz. At the viewing distance of 0.6 each pixel subtended 1.46 arcmin. Stimuli were generated using Psychtoolbox package (v. 3.0.8) (Brainard, 1997) for MATLAB (v 7.4).

2.1.3. Stimuli

Four types of stimuli were used (see Fig. 1):

(a) Rectangle-Bar – a rectangle with zero disparity (all disparities are specified with respect to initial fixation) next to a bar with uncrossed disparity (Fig. 1A).
(b) Rectangle-Disc – a rectangle with zero disparity next to a disc with uncrossed disparity (Fig. 1B).
(c) Two-Bars – two bars side by side, one with zero disparity and the other with uncrossed disparity (Fig. 1C).
(d) Two-Discs – two discs side by side, one with zero disparity and the other with uncrossed disparity (Fig. 1D).

The bars in all stimuli had a width of 5.8 and length of 146 arcmin and the discs had a diameter of 5.8 arcmin. The relative disparities between the pairs of objects were one of 2.9, 5.8 and 8.76 arcmin (theoretical depth of 0.56, 1.13 and 1.72 cm for an IOD of 6.5 cm). Stimuli were black on a grey background (10 cd/m^2). The Rectangle-Bar and the Rectangle-Disc stimuli were used by Tsirlin et al. (2012). The two new stimuli, Two-Bars and Two-Discs were used to test the generality of the Tsirlin et al. (2012) results. Prior to testing observers' interocular separation was measured using a pupilometer.

2.1.4. Procedure

Observers were asked to estimate the depth (or disparity) between the two objects on the screen (rectangle and bar, rectangle and disc, two bars or two discs) using three methods of estimation:

(1) Disparity probe (DP) – a square subtending 17.5 × 17.5 arcmin was presented to the left of the stimulus objects. The square could be moved in depth in 0.7 arcmin steps using a gamepad. Observers were asked to adjust the disparity probe to the perceived depth of the object with the uncrossed disparity.
(2) Virtual ruler (VR) – a virtual vertical ruler (Tsirlin et al., 2012) was presented on the screen to the left of the zero disparity object (see Fig. 2A). The ruler consisted of a vertical line subtending 3 × 496 arcmin bisected by a horizontal line subtending 30 × 3 arcmin and another, moveable, horizontal cursor line (30 × 3 arcmin). Observers were asked to position the cursor (using the mouse) such that the distance between the bisection mark and the cursor matched the perceived depth between the two objects.
(3) Physical ruler (PR) – a purpose-built touch sensitive sensor. A rectilinear SoftPot membrane potentiometer (SpectraSymbol) was mounted to an aluminium bar. The sensor strip was 200 mm long and 7 mm wide with a total resistance of 10 kOhm. The potentiometer allowed linear measurements across the 200 mm length, with a resolution of approximately 0.2 mm. Responses were read using an analog to digital converter and a 16-bit micro controller. The recorded voltage was converted to millimetres using a MATLAB script, and calibrated prior to testing. The observers were required

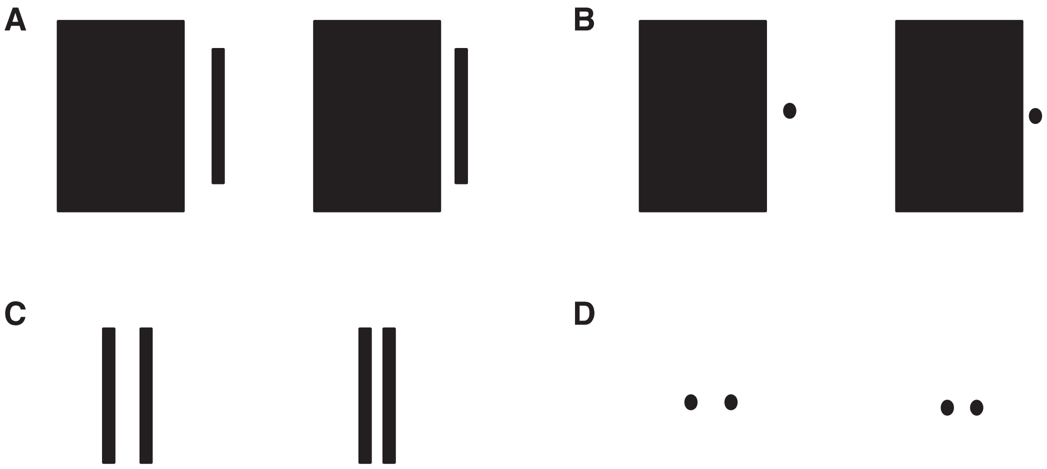


Fig. 1. The four types of stimuli of Experiment 1 – (A) Rectangle-Bar, (B) Rectangle-Disc, (C) Two-Bars and (D) Two-Discs. When cross-fused, the object to the right should appear further away than the object to the left. For divergent fusion the depth order is reversed.

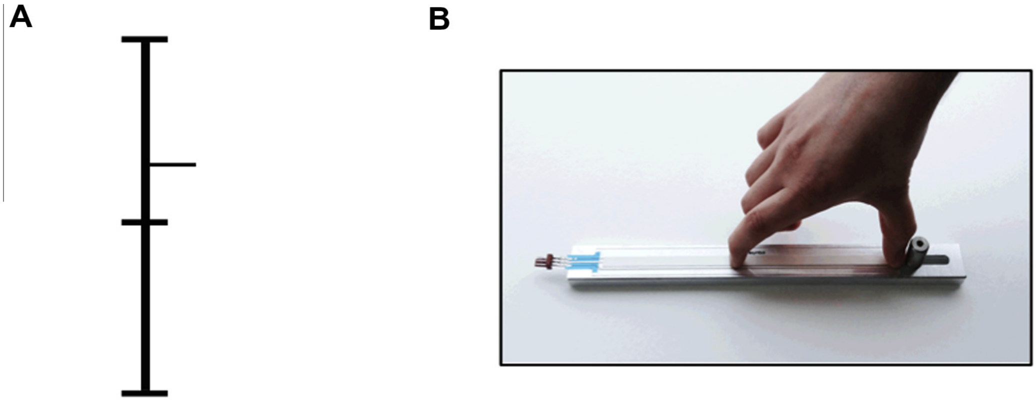


Fig. 2. (A) The virtual ruler, (B) the physical ruler.

to indicate the perceived depth between the two objects by first resting their thumb at the base of the ruler and then placing their index finger on the strip. A small red LED positioned 10.8 deg below the centre of the display signalled when sufficient pressure was applied to the sensor strip. When this light was on, and observers were satisfied with their thumb-finger separation, they pressed a button the computer keyboard with their other hand to enter their response (see Fig. 2). Since the experiment was conducted in the dark and the observers' hand was placed on the desk, below the stereoscope mirrors, it was not visible during testing. The LED disappeared when the response was entered, and the next trial was initiated.

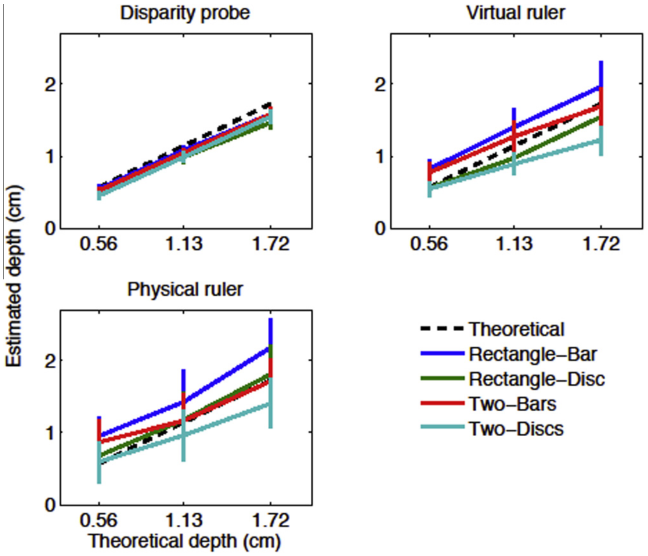
For all three methods, once the desired measurement was made, the observer pressed a button on a gamepad or on a mouse to record their response.

The three estimation methods were assessed in separate blocks and observers were given the opportunity to practice the respective technique before the session. The session order was randomized and counterbalanced across observers. Twenty repetitions of each of the 12 conditions (4 stimuli × 3 disparities) were presented in random order within a session (240 trials in total per session). The observer's head was stabilized with a chin rest, viewing time was not limited and fixation, after the start of the trial, was not controlled.

2.2. Results

Fig. 3 shows the results of Experiment 1 averaged across the six observers for each of the three estimation methods. Relative disparities between the two stimulus objects were converted to equivalent depth values in centimetres using the geometric relation between depth and screen disparity, and the average IOD of the group of observers. For clarity, relative disparities in arcmin are indicated in the caption. Disparity estimates made with the disparity probe were converted to equivalent depth using the same equation.

As can be seen in Fig. 3, there were only small differences between depth estimates in the four stimulus configurations when the disparity probe was used, and larger differences between the configurations for the other two estimation methods. To evaluate these differences, we applied repeated-measures two-way ANOVAs (using the average individual estimates) for each estimation method with Configuration and Disparity as factors and an alpha level of 0.05.


Disparity probe

Virtual ruler

Physical ruler

Fig. 3. Results of Experiment 1. The three graphs correspond to the three depth estimation methods. In each plot, the mean ($n = 6$) estimated depth is plotted as a function of depth corresponding to the relative disparity in the stimulus (2.9, 5.8 and 8.76 arcmin accordingly). The differently colored lines correspond the four different stimulus configurations. Error bars show ±1 standard error (SE). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

For all methods of estimation, Disparity (DP: $F_{(2,10)} = 784$, $p < 0.001$; VR: $F_{(2,10)} = 32$, $p < 0.001$; PR: $F_{(2,10)} = 31.6$, $p < 0.001$) had a significant main effect. As expected, increasing disparity resulted in increased depth estimates. Configuration also had a significant main effect on depth estimates for all methods (DP: $F_{(3,15)} = 3.7$, $p = 0.034$; VR: $F_{(3,15)} = 14.3$, $p < 0.001$; PR: $F_{(3,15)} = 17.5$, $p < 0.001$). To better understand the effect of Configuration we compared all pairs of configurations within each of the depth estimation methods. This posthoc analysis with Bonferroni adjustment showed that for all estimation methods, significantly more depth was perceived by the observers in the Rectangle-Bar than in the Rectangle-Disc condition (DP: $t_{(17)} = 5.37$, $p < 0.001$; VR: $t_{(17)} = 6.9$, $p < 0.001$; PR: $t_{(17)} = 5.5$, $p < 0.001$). In addition, significantly more depth was perceived in Two-Bars than in Two-Discs with the virtual ruler and the physical ruler methods (DP: $t_{(17)} = 3.27$, $p = 0.018$; VR: $t_{(17)} = 5.8$, $p < 0.001$; PR: $t_{(17)} = 5.3$, $p < 0.001$). These results confirm the findings of Tsirlin et al. (2012) and demonstrate that the differences in depth percepts between discs and bars generalize to different stimuli and different methods of estimation.

We also compared estimated depth between the configurations containing bars – Rectangle-Bar and Two-Bars and between configurations containing discs – Rectangle-Disc and Two-Discs, since it appears from Fig. 3 that there might be a difference between these conditions. The only significant differences were found between Rectangle-Disc and Two-Discs ($t_{(17)} = 4.8$, $p < 0.001$) and Rectangle-bar and Two-bars ($t_{(17)} = 4$, $p = 0.003$) for the physical ruler estimation method with more depth is seen in the configurations containing the rectangle.

Finally, we compared the estimated depths to the theoretically predicted depth using one sample *t*-tests with *p*-values adjusted for the number of comparisons in each condition/method. We found that, for all conditions, estimates for the bar stimuli did not differ significantly from the geometric predictions. For the disc stimuli, both in DP and VR methods of estimation, we found that estimates either differed significantly or tended towards significance (DP Rectangle-Disc [$p = 0.08$, $p = 0.08$, $p = 0.05$]; DP Two-Discs [$p = 0.06$, $p = 0.06$, $p = 0.08$]; VR Two-Discs [disp 1.72 cm $p = 0.1$])

Taken together, these results demonstrate that the perceived relative depth of the small disc was underestimated in comparison to the long bar and that the bar estimates were more veridical. This is especially interesting considering that according to linear perspective the larger bar should be perceived as lying closer to the rectangle than the small dot (such a percept is typical when the stimulus is viewed monocularly). Our results also underscore the often-overlooked difference between measurements of depth and disparity. There was a large difference between depth estimates for both stimuli configurations using the virtual and the physical ruler methods. However, using the probe method in which observers are able to match disparity rather than depth, estimates were significantly different only in the stimuli with the rectangle.

3. Experiment 2

3.1. Motivation

Experiment 1 showed that there is a difference in the perceived depth of a small disc and a long bar that have the same relative disparity. This was true for all types of stimuli and all methods of estimation. In Experiment 2 we test the hypothesis that this difference is rooted in the relative size of the features as proposed by Tsirlin et al. (2012). That is, they argued that because the vertical extent of the disc is much smaller than that of the bar, depth estimates for the bar are more reliable because information is integrated along the vertical contours. To test this hypothesis we used the virtual ruler estimation method and asked observers to estimate the depth using the Rectangle-Bar and Rectangle-Disc stimuli; we varied the length of the bar and diameter of the disc.

3.2. Methods

3.2.1. Observers and apparatus

Observers and apparatus were the same as in Experiment 1.

3.2.2. Stimuli

We used the Rectangle-Bar and the Rectangle-Disc stimuli described in Experiment 1 and disparities 5.8 and 8.76 arcmin (theoretical depth of 1.13 and 1.72 cm, respectively). The bar was 5.8 arcmin wide and either 5.8, 11.7, 17.5 or 58.4 arcmin long. The disc had a diameter of 5.8, 11.7, 17.5 or 58.4 arcmin and the gap between the nearest edge of the disc and the rectangle was fixed at 29 arcmin.

3.2.3. Procedure

Depth estimation was performed using the virtual ruler method (see Experiment 1) in two separate sessions, one for the Rectangle-Bar and one for the Rectangle-Disc configuration. For each configuration there were 8 stimuli (4 sizes × 2 disparities) that were repeated 20 times in random order for a total of 160 trials per configuration.

3.3. Results

Fig. 4 shows the mean results of Experiment 2 for six observers; it is apparent that perceived depth increases with increasing stimulus height. A repeated-measures three-way ANOVA with Configuration, Size and Disparity as factors, showed no significant effect of Configuration ($F_{(1,5)} = 1.8$, $p = 0.23$). This indicates that when the height of the disc and the bar are the same, the amount of depth perceived in the two configurations is not significantly different. In subsequent analyses, the data from the two configurations (Rectangle-Bar and Rectangle-Disc) were collapsed. A two-way repeated-measures ANOVA showed a significant main effect of Size ($F_{(3,15)} = 9.4$, $p < 0.001$) and Disparity ($F_{(1,5)} = 50.16$, $p < 0.001$) but no significant interaction between the factors. These results indicate that there is a significant increase in perceived depth between the rectangle and the bar/disc as the height of the bar/disc increases, and this effect generalizes across disparities.

4. Discussion

The perceived depth of the small discs in Experiment 1 was smaller than that of long bars even though they had the same relative disparity and width. This result was replicated for all stimuli configurations and measurement methods. In Experiment 2 we assessed the relationship between perceived depth and feature height. We hypothesized that integration of disparity along the vertical edges of the objects was responsible for the greater depth perceived in the bar stimuli in Experiment 1. Consistent with this hypothesis, as the radius of the disc and the height of the bar increased, the perceived depth also increased and became closer to the geometrically predicted depth. There was no difference in perceived depth for bars and discs of the same height.

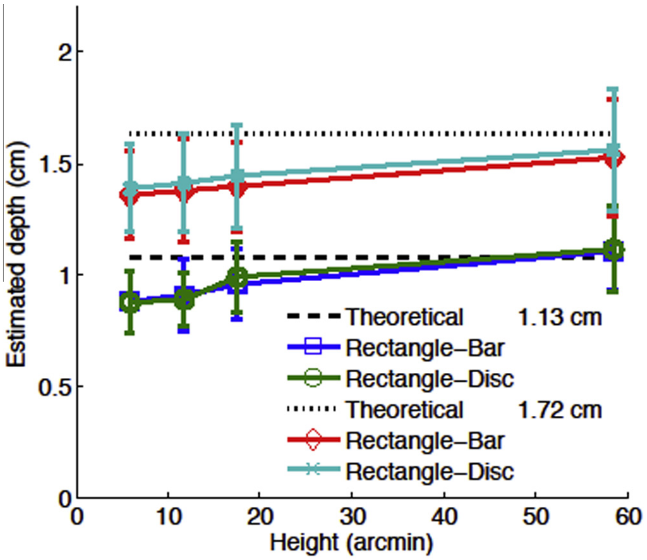


Fig. 4. Results of Experiment 2. Mean estimated depth ($n = 6$) is plotted as a function of bar and disc height (diameter). Different colored lines and symbols show the different stimuli configurations (Rectangle-Bar and Rectangle-Disc) and two disparities. The dashed and dotted lines show the predicted depth for the two test disparities. Error bars show ±1 SEM. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Why does integration of disparity along the vertical edges result in a greater and potentially more veridical depth estimate? The increased magnitude and accuracy of the depth signal could occur for at least two reasons. First, a longer stimulus might support a greater disparity range, resulting in greater perceived depth. Second, spatial integration could reduce noise leading to a more robust disparity signal and more accurate depth estimates. Moreover, this increased reliability of the disparity signal could result in an increased weighting of disparity and larger depth estimates in the presence of other cues indicating no depth (e.g., lack of focal blur) and an inherent bias towards small disparities (Banks & Vlaskamp, 2009; McKee, Verghese, Ma-Wyatt, & Petrov, 2007; Prince & Eagle, 2000). In support of this noise-reduction hypothesis, Tsirlin et al. (2012) provided evidence of a link between positional noise and stimulus height. Using stimuli similar to the Rectangle-Bar and Rectangle-Disc they showed that the just noticeable difference in size of the gap between the disc and the rectangle was larger than that between the bar and rectangle. This suggests that the disc produced noisier visual positional responses, and that integration of information along the vertical edge of the objects intensified the signal.

At least two physiological mechanisms could underlie integration of disparity along the vertical edges and provide both more robust disparity signals and a greater disparity range for longer vertical edges: size–disparity correlation and inter-neural connections. As outlined in the Introduction, previous research has suggested that the size–disparity correlation could account for the dependence of stereopsis on stimulus scale (Tyler, 1975). Size–disparity correlation posits that larger detectors encode large disparities but cannot discern high spatial frequencies. Conversely, small detectors are sensitive to high spatial frequencies, and encode relatively small disparities. So maximum disparity range is larger for large objects than for small objects. Theoretically, size–disparity correlation is consistent with phase-based coding of disparity information. That is, the maximum disparity that phase-shift disparity detectors, observed in the early visual areas, can discern is limited by their spatial frequency tuning and thus their size (Anzai, Ohzawa, & Freeman, 1997, 1999). The effects of integration of information along the vertical contours of the objects observed in our experiments could be explained by the size–disparity correlation property of disparity detectors. Longer objects excite larger cells more than shorter objects (integration of contours inside the receptive fields), thus supporting a greater disparity range. Larger cells are also less sensitive to noise, both disparity and positional, which can in turn result in more accurate estimates of disparity for longer objects.

Spatial integration of disparity could also take place through excitatory inter-neural connections of neurons with similar disparity tuning located in close proximity along the edge of an object. Such facilitatory interactions have been proposed in many computational models of stereopsis (Grimson, 1981; Marr & Poggio, 1976; Zitnick & Kanade, 2000). Excitation from neighbouring neurons can increase the firing rate of each individual unit, thus providing a more robust disparity signal for a given object. As discussed above, a more robust signal can result in both greater accuracy and a greater range.

The two mechanisms are not mutually exclusive and can work in tandem to produce the phenomena observed in Experiments 1 and 2. Further psychophysical and computational analysis is required to elucidate the exact physiological underpinnings of our findings.

Grant support

This research was supported by NSERC grants to L.M. Wilcox and R.S. Allison and a CGS NSERC scholarship to Inna Tsirlin.

References

Anzai, A., Ohzawa, I., & Freeman, R. D. (1997). Neural mechanisms underlying binocular fusion and stereopsis: Position vs. phase. *Proceedings of the National Academy of Sciences of the United States of America, 94*(10), 5438–5443.

Anzai, A., Ohzawa, I., & Freeman, R. D. (1999). Neural mechanisms for encoding binocular disparity: Receptive field position vs. phase. *Journal of Neurophysiology, 82*(2), 874–890.

Banks, M. S., & Vlaskamp, B. (2009). The Venetian-blind effect: A prior for zero slant or zero disparity? (Abstract). *Journal of Vision, 9*(8), 45.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Felton, T. B., Richards, W., & Smith, R. A. (1972). Disparity processing of spatial frequencies in man. *The Journal of Physiology, 225*(2), 349–362.

Grimson, W. E. (1981). A computer implementation of a theory of human stereo vision. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences, 292*(1058), 217–253.

Heckmann, T., & Schor, C. M. (1989). Is edge information for stereoacuity spatially channeled? *Vision Research, 29*(5), 593–607.

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

McKee, S. P., Verghese, P., Ma-Wyatt, A., & Petrov, Y. (2007). The wallpaper illusion explained. *Journal of Vision, 7*(14).

Prince, S. J., & Eagle, R. A. (2000). Weighted directional energy model of human stereo correspondence. *Vision Research, 40*(9), 1143–1155.

Richards, W., & Kaye, M. G. (1974). Local versus global stereopsis: Two mechanisms? *Vision Research, 14*(12), 1345–1347.

Schor, C. M., & Badcock, D. R. (1985). A comparison of stereo and vernier acuity within spatial channels as a function of distance from fixation. *Vision Research, 25*(8), 1113–1119.

Smallman, H. S., & MacLeod, D. I. (1994). Size–disparity correlation in stereopsis at contrast threshold. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 11*(8), 2169–2183.

Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2012). Da Vinci decoded: Does da Vinci stereopsis rely on disparity? *Journal of Vision, 12*(12).

Tyler, C. W. (1973). Stereoscopic vision – Cortical limitations and a disparity scaling effect. *Science, 181*(4096), 276–278.

Tyler, C. W. (1975). Spatial organization of binocular disparity sensitivity. *Vision Research, 15*(5), 583–590.

Tyler, C., & Julesz, B. (1980). On the depth of the cyclopean retina. *Experimental Brain Research, 40*(2), 196–202.

Zitnick, C. L., & Kanade, T. (2000). A cooperative algorithm for stereo matching and occlusion detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*, 675–684.