PERGAMON

Vision Research 40 (2000) 3501–3506

Vision Research

www.elsevier.com/locate/visres

Depth cue integration: stereopsis and image blur[☆]

George Mather *, David R.R. Smith

Laboratory of Experimental Psychology, Biology School, University of Sussex, Brighton BNI 9QG, UK

Received 25 June 1999; received in revised form 18 October 1999

Abstract

Depth-of-focus limitations introduce spatial blur in images of three-dimensional scenes. It is not clear how the visual system combines depth information derived from image blur with information from other depth cues. Stereoscopic disparity is the pre-eminent depth cue, so experiments were conducted to investigate interactions between image blur and stereoscopic disparity. Observers viewed two random dot stereograms (RDSs) in a 2AFC task, and were required to identify the RDS depicting the greatest depth. In control observations, all dots in both RDSs were sharply defined. In experimental observations, one RDS contained only sharply defined dots, but the other contained differential spatial blur to introduce an additional depth cue. Results showed that the addition of differential blur made only a marginal difference to apparent depth separation, and only when the blur difference was consistent with the sign of disparity. Cue combination between blur and disparity cues is thus weighted very heavily in favour of the latter. It is shown that blur and disparity cues co-vary according to geometric optics. Since the two cues are effective over different distances, the visual system is not normally called upon to integrate them, and is most likely to make use of blur cues over distances beyond the range of disparity mechanisms.

Keywords: Depth perception; Image blur; Stereoscopic disparity; Cue interaction

1. Introduction

Photographic and retinal images of three-dimensional scenes contain regions that are spatially blurred by differing amounts, due to depth-of-focus limitations in the imaging system. This blur variation offers a quantitative cue to the relative distances of points in the scene (Pentland, 1987). Natural images frequently contain a number of other cues to depth, such as motion parallax, interposition, and perspective. Pre-eminent among these other cues is stereoscopic disparity, since it is always present when viewing a three-dimensional scene binocularly (assuming the observer has normal binocular vision). A number of previous studies have explored interactions between multiple depth cues. Two commonly encountered forms of interaction are cue summation or averaging, and cue dominance or vetoing (Bülthoff & Mallot, 1988; Howard & Rogers, 1995). In cue summation, the two cues are combined in a weighted algebraic sum or mean. In cue dominance or vetoing, one cue completely dominates the perceptual judgement, so that the other cue is disregarded. The present paper investigates which form of interaction occurs when blur and stereo cues are available.

Observers viewed two random dot stereograms (RDSs) in a 2AFC task. Both stereograms contained a disparate central square of dots against a random-dot background. In control observations, all dots in both stereograms were sharply defined, or all were blurred by the same amount. In experimental observations, one stereogram (comparison) contained only sharply defined dots, but the other (reference) contained differential spatial blur between the central square and the background to introduce a second depth cue. In both control and experimental observations we measured the comparison disparity required to produce a depth match with a specific reference disparity.

We show in Appendix A that both the disparity and the blur produced by an object at a given distance from fixation are inversely proportional to the square of fixation distance. However, above-threshold blur values occur only at very large disparities, whereas disparities

[☆] Part of this work was presented at the European Conference on Visual Perception in Trieste, 1999 (Mather & Smith, 1999).

* Corresponding author. Fax: +44-1273-678433.

E-mail address: georgem@biols.susx.ac.uk (G. Mather).

0042-6989/00/$ - see front matter
PII: S0042-6989(00)00178-4

within Panum's fusional area tend to occur with below-threshold blur values (see below for a more detailed discussion of this point). In Experiment 1 we selected an easily detectable blur value (Gaussian space constant 4.5 arcmin), paired with disparities within the range of fusion. Hence, the depth intervals signalled by the two cues were very different. Stereoscopic disparities ranged from -1.76 to $+1.76$ arcmin (negative values denote far disparities, and positive values denote near disparities). At the viewing distance used (114 cm), these disparities corresponded to depth intervals within ± 1.1 cm from the screen. At our viewing distance and pupil diameter (6 mm according to digital photography), a Gaussian blur space constant of 4.5 arcmin corresponded to a depth interval of 113 cm from the screen. If depth judgements involve cue summation or averaging even across inconsistent cues, then we expect that apparent depth (as measured by matching disparity) will be greater in RDSs containing both disparity and blur cues than in RDSs containing only the disparity cue, due to the greater depth signalled by the blur cue. If stereo cues dominate or veto blur cues in this situation, we expect little or no difference in apparent depth between control and experimental stimuli.

2. Experiment 1

2.1. Methods

2.1.1. Subjects

Six observers took part, one of the authors and five naïve subjects. Appropriate optical corrections were worn.

2.1.2. Apparatus and stimuli

Stimuli were generated by a Silicon Graphics O2™ workstation and displayed on a GDM-17E21 colour monitor (96 Hz frame rate). Stereograms were created using a field-sequential stereo display synchronised to a pair of electro-optical shutters (CrystalEyes 2™) via an infrared link. All images had a fixed Michelson contrast of 0.68 (measured from the lowest and highest luminances in the stimuli). Each stereogram contained a 384×384 element array of random binary black-white noise (11.8 and 61.9 cd/m^2). A 256×256 central region in one stereo half-image was given a horizontal offset relative to the other half-image, to create a disparity cue. The magnitude and direction of this offset varied during the experiment to supply a variety of near and far disparities. At the 114 cm viewing distance, the stimulus array subtended $5.5l \times 5.51$ arcdeg, and the disparate central square subtended 3.68×3.68 arcdeg. In between trials a central fixation cross was displayed at zero disparity relative to the computer screen. Each screen pixel subtended 0.87 arcmin. Each random element of the stereogram was a 2×2 pixel area 'dot' that subtended $1.72 \times l.72$ arcmin. When required, either the central region, the surround, or both were spatially blurred by the application of a Gaussian blurring function (space constant 4.5 arcmin), and quantised to 256 grey levels (via a linearised look-up table).

Note that half-occlusion regions were given the same degree of blur as the surface to which they were attached (Shimojo & Nakayama, 1990; Anderson, 1994).

2.1.3. Design and procedure

One stereogram (comparison) always contained sharply defined dots in both the central square and in the background. The other stereogram (reference) contained different combinations of blur in the square and background, defining four stimulus conditions:

- sharp central square on a sharp background (S/S);
- blurred square on a blurred background (B/B);
- sharp square on a blurred background (S/B); and
- blurred square on a sharp background (B/S).

An example stimulus is shown in Fig. 1. Conditions S/S and B/B constitute controls, since there is no difference in blur between the central square and the background. Conditions S/B and B/S do contain differential blur to provide a depth cue. The border between the central square and the background was always sharp, consistent with the square being nearer than the background in the S/B stimulus, and farther in the B/S stimulus (Marshall, Burbeck, Ariely, Rolland, & Martin, 1996; Mather, 1996). The central square in each reference stimulus could contain one of five disparities relative to the zero disparity background: -103, -52, 0, 52, and 103 arcsec (negative values denote far disparities, and positive values denote near disparities). Consequently, S/B stimuli contained an ecologically valid combination of cues only at positive disparities (subjects fixate in the plane of the background, so the sharp dots in the central square have near disparity). The method of constant stimuli was used in a self-paced temporal two alternative forced choice procedure. Each reference stimulus could be paired with one of five comparison stimuli, containing disparities bracketing the reference disparity of x arcsec, i.e. $(x-103)$, $(x-52)$, $(x+0)$, $(x+52)$, or $(x+103)$ arcsec, making a total of 100 stimulus pairs (four types of reference stimulus, five reference disparities, and five comparison disparities). A pseudo-randomly selected pair from this set was presented in each trial (initiated by a button press). Each stimulus in the pair was displayed for 250 ms, separated by an interval of 1000 ms. In between stimuli the display was evenly illuminated at mean luminance (37.8 cd/m^2), apart from the zero-disparity fixation mark. The pseudorandom sequence of stimulus pairs was constrained so that no stimulus would be presented for the $(n+1)$th time until all stimuli had been presented n times. Each stimulus pair was presented 20 times (in ten trials the comparison stimulus appeared first, and in ten trials it appeared second). After each presentation the observer pressed one of two keys to identify the stimulus in which the central square appeared furthest away. Stimuli were viewed without head restraint and with natural pupils. Observations took place in a darkened room. Observers fixated a central zero-disparity fixation mark which was removed from the display during stimulus presentation.

2.2. Results and discussion

Cumulative Gaussian functions were fitted to the psychometric functions from each subject using Probit analysis (Foster & Bischof, 1997), modified for 2AFC. Fig. 2 (left) plots mean points of subjective equality (PSE) in control conditions (no blur differences), in terms of comparison disparity required to achieve a depth match with the reference stimulus, as a function

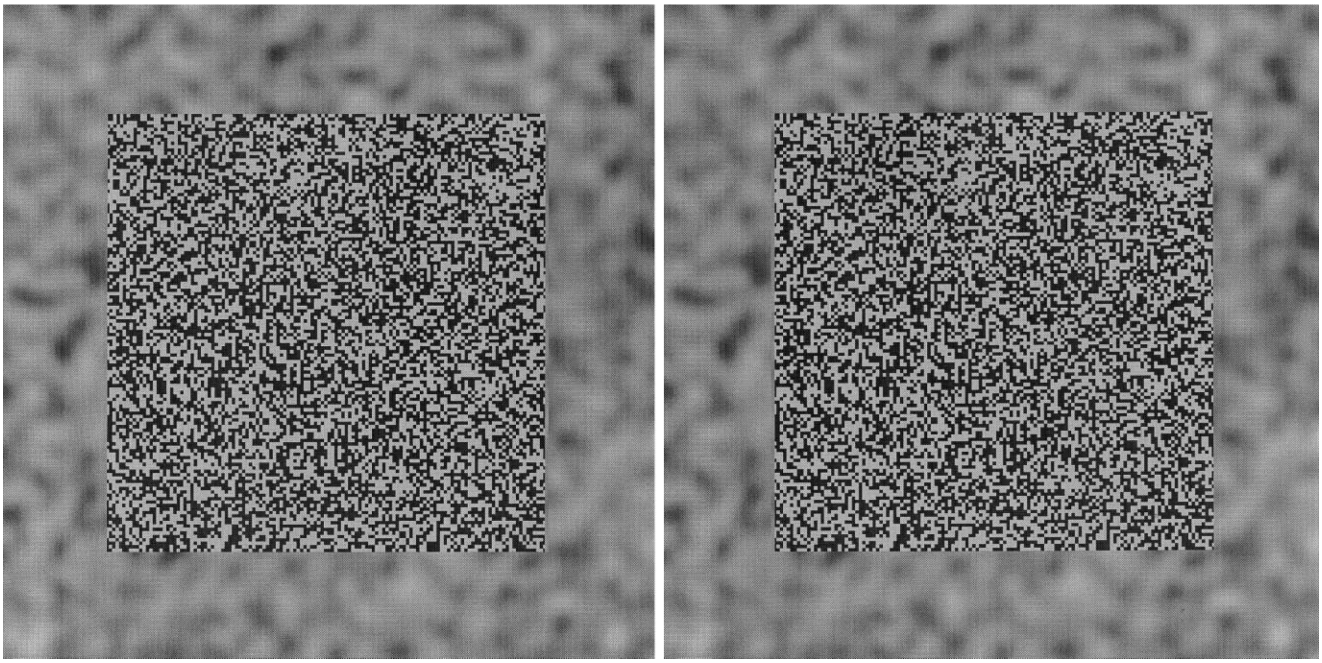

Fig. 1. Reproduction of a stereogram used in the experiments (from the S/B condition), arranged for crossed free fusion (actual images were displayed using electro-optical shutters). The central square is shifted horizontally by two elements in one array relative to the other, to create a near disparity signal. Background blur space constant matches that used in Experiment 1.

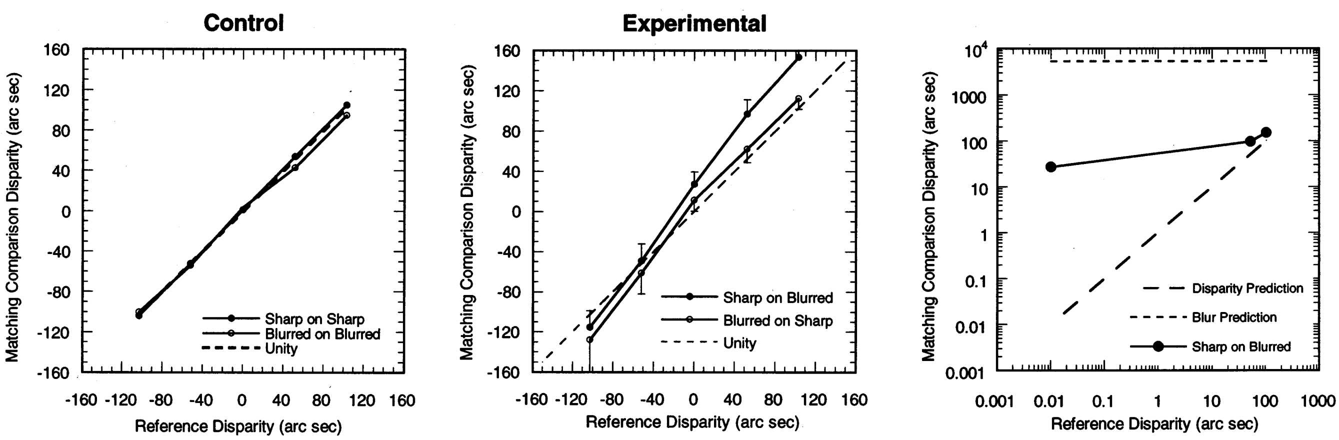

Fig. 2. Results from Experiment 1. The left-hand graph plots mean subjective matches between comparison RDSs containing sharp foreground and sharp background elements, and reference RDSs that were either sharp-on-sharp again (filled symbols), or blurred-on-blurred (open symbols, Gaussian space constant equal to 4.5 arcmin). The middle graph plots mean subjective matches between sharp-on-sharp comparison RDSs and reference RDSs that either contained a sharp foreground on a blurred background (sharp on blurred, filled symbols), or a blurred foreground on a sharp background (blurred on sharp, open symbols). Negative values denote far disparities, vertical lines show ± 1 SE of the mean (where bars are not visible SEs were too small to plot). The dashed line at unit slope identifies physically correct disparity matches. The right-hand graph re-plots S/B data at near disparities from the middle graph, and includes predictions on the basis of the blur cue (small dashes), and on the basis of the stereo cue (large dashes). Data at zero disparity have been plotted at a disparity of 0.01 s, to permit use of logarithmic axes.

of reference disparity. As expected, apparent depth matches were obtained at matching disparity values, since the functions have unit slope. Fig. 2 (middle) plots mean PSE as a function of reference disparity in experimental conditions. Filled symbols show results for condition S/B, and open symbols show results for condition B/S. There is a clear departure from unit slope, indicating that stimuli containing a combination of stereo and blur gave a greater impression of depth. However, this effect was apparent only for ecologically valid stimuli-S/B stereograms at near disparities (i.e. filled symbols and positive disparities in Fig. 2, middle). The addition of blur had only a marginal effect on matching disparity — the depth interval of 113 cm defined by a blur space constant of 4.5 arcmin corresponds to a binocular disparity of 1.5°. Fig. 2 (right) re-plots S/B data for near disparities from Fig. 2 (middle) along with predictions on the basis of the blur cue (small dashes), and on the basis of the disparity cue (large dashes). Data fall much closer to the disparity predictions than to the blur predictions (note the logarithmic axes). When there is a gross discrepancy between the depth intervals given by the two cues, as there was in our stimuli, observers' depth matches are dominated by the stereo cue. It may be that blur is always treated as a relatively weak depth cue by the visual system because: (i) it is relatively unreliable (blur magnitude varies with pupil diameter and refractive state as well as with depth); and (ii) it may be a rather imprecise metric for depth, given the JND for blur discrimination (Mather, 1997). A similar situation applies in the case of shape-from-texture cues compared to disparity cues to depth (Johnston, Cumming, & Parker, 1993). On the other hand, the predominance of stereo may largely be due to the marked inconsistency between the cues in our stimuli. Bülthoff and Mallot (1988) reported this kind of dominance when inconsistent stereo and shading cues were used.

3. Experiment 2

To distinguish between these two possibilities, we sought to devise stimuli in which the stereo and blur cues were consistent. If stereo still dominated, the explanation based on cue conflict could be rejected. According to geometrical optics, there is a direct relation between disparity (ϕ) and defocus blur (σ):

$$\tan \sigma = \tan \phi (p/2a) \tag{1}$$

where p is pupil diameter, and a is interpupillary distance (IPD). See Appendix A for the derivation of this equation. In the conditions of our experiment (pupil diameter 6 mm and mean IPD 59.8 mm) increasing disparity to a relatively large value produces only a small degree of defocus blur. For example, a disparity of 20 arcmin is in the region of the upper disparity limit for RDSs (though the issue is complicated by the dependence of the upper disparity limit on stimulus size; see Nielsen & Poggio, 1984; Wilcox & Hess, 1995; Glennerster, 1998) but equates to a blur space constant of only 1 arcmin (approximately three times threshold; Mather, 1997). Compromise stimuli were devised, containing larger disparities and a smaller amount of blur than in Experiment 1, to determine whether the relative weighting of blur in depth judgements increased when depth cues conflicted less.

3.1. Method

3.1.1. Subjects

Two subjects took part, one author and a naïve observer.

3.1.2. Apparatus and stimuli

Two types of reference stimulus were created (S/B and B/S) identical to those used in Experiment 1 except that blur space constants were reduced to 0.97 and 1.92 arcmin, and disparities were increased to ± 20.63 and ± 10.32 arcmin. Specifically, we paired disparity and blur as follows: 20.36′ disparity and 0.97′ blur, 10.32′ and 1.92′, $-10.32'$ and 1.92′, $-20.36'$ and 0.97′. The combination of 20.63′ disparity and 0.97′ blur space was ecologically valid (given the interpupillary distances and pupil sizes of our observers, and the fixation distance). The ecologically valid blur space constant for a 10.32′ disparity was below threshold (confirmed by observation for our particular stimuli and observers), so we employed a larger blur at this reference disparity that fell between ecologically valid blur and the large amount of blur used in Experiment 1. Comparison stimuli (no blur cue) covered a range of disparities from -24.07 to 24.07 arcmin. All other stimulus details were identical to those described earlier.

3.1.3. Design and procedure

A 2AFC task was used, using a procedure identical to that employed in Experiment 1.

3.2. Results and discussion

Both subjects reported problems with diplopia, especially for near disparities, making it impossible to calculate PSEs for these stimuli. Judgements were more reliable at far disparities. PSEs are plotted in Fig. 3. It is clear that matching comparison disparities show no effect of the blur present in the reference stimuli, since all 50% points fall close to the unit slope line. Reducing the conflict between blur and stereo cues thus failed to enhance the contribution of blur to stereo-matching judgements.

It is possible that the different effects of near and far

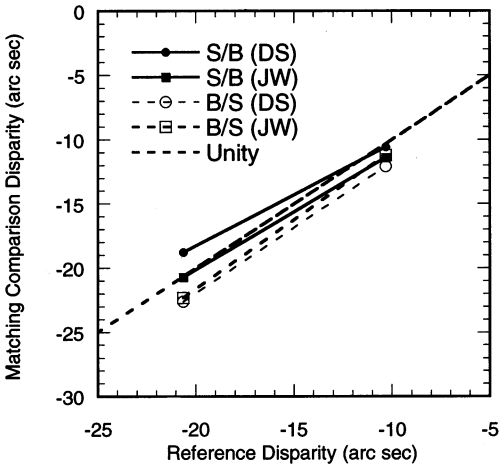

Fig. 3. Results for far disparity conditions in Experiment 2, for two observers. The graph plots subjective matches between comparison RDSs containing sharp foreground and sharp background elements, and reference RDSs that were either sharp-on-blurred (S/B; filled symbols), or blurred-on-sharp (B/S; open symbols). Gaussian space constant was equal to 0.97 arcmin (for -20.63 arcmin disparity) and 1.92 arcmin (for -10.32 arcmin disparity).

disparities in Experiment 2 were due to the presence of uncontrolled and unnoticed fixation disparities, provoked partly by the presence of blur. Stimulus exposure durations were kept very short (250 ms) to avoid vergence changes during stimulus presentation (Wilson, 1973, reported a mean latency for accommodation vergence of 249.5 ms), but it may be that the experimental arrangement resulted in the presence of fixation disparities at the onset of each trial, and these interfered with judgements. A control experiment measured fixation disparity in our subjects under the viewing conditions of the two experiments (i.e. lighting, viewing distance, observers). In each trial, nonius lines were briefly presented immediately above and below the inter-trial fixation marker, for the same duration as the experimental stimulus (250 ms). Following each presentation of the lines the observer pressed a response button to report whether the upper (left-eye) line was displaced to the left or to the right of the lower (right-eye) line. Results revealed the presence of a negligible near fixation disparity of 1 arcsec (standard error ± 15 secarc, $n=6$) for Experiment 1. This increased to a small far fixation disparity of 31 secarc (SE ± 30 arcsec, $n=2$) for Experiment 2. Previous studies (e.g. Jaschinski-Kruza, 1994) have reported much larger fixation disparities. We conclude that fixation disparity is unlikely to have contributed to the results of our experiments.

4. Conclusions

These two experiments indicate that differential image blur makes only a small contribution to the impression of depth seen in random dot stereograms, when measured against depth seen in RDSs without blur cues. The increased impression of depth found in stimuli containing a combination of stereo and blur was a very small fraction of that predicted by the magnitude of the blur cue, and only occurred in ecologically valid stimuli. The marginal effect of blur may possibly be due to its effect on the disparity matching process. In natural images the two cues are constrained to co-vary (Eq. (1)), and due to sensory limitations whenever one cue is perceptually effective, the other is not. So in natural conditions, there is no need to integrate them. When both cues are present, it seems that stereo is dominant. The visual system is most likely to make use of the blur cue to depth over distances beyond the range of disparity mechanisms. We are currently investigating the integration of the blur cue with pictorial depth cues.

Acknowledgements

This research was supported by a research grant from the Engineering and Physical Sciences Research Council, UK.

Appendix A. The relation between disparity and defocus blur

The relation between angular disparity (ϕ) and depth is given by the following equation (see Howard & Rogers, 1995):

$$\tan\phi = \frac{a(D-u)}{Du} \tag{A1}$$

where a is interpupillary distance (IPD), u is fixation distance, and D is the distance of the non-fixated object.

The relation between depth and defocus blur is given by the following equation (see Pentland, 1987):

$$D = \frac{Frv}{rv - F(r+\sigma)} \tag{A2}$$

where D is the distance of the non-fixated (blurred) object, F is the focal length of the lens, r is lens aperture radius, v is the distance between the lens and the image plane, and a is blur circle radius (in linear units). Solving for σ yields

$$\sigma = \frac{Drv - Frv - DFr}{DF} \tag{A3}$$

According to the well-known lens equation,

$$F = \frac{uv}{u+v} \tag{A4}$$

where u is the distance of a perfectly focussed image (equivalent to fixation distance). Substituting Eq. (A4) for F in Eq. (A3) and simplifying yields

$$\sigma = \frac{r(D-u)v}{Du} \tag{A5}$$

Converting σ in Eq. (A5) into angular units, and removing the sign of the depth difference (since blur cannot be negative),

$$\tan\sigma = \frac{r|D-u|}{Du} \tag{A6}$$

If the depth interval $\Delta d = (D-u)$ is very small, the small angle approximation can be used

$$\tan\sigma = \frac{r|\Delta d|}{d^2} \tag{A7}$$

where d is fixation distance. Eq. (A1) for disparity has the same form as Eq. (A6) for blur radius. Thus it is possible to calculate blur given only disparity, aperture radius and IPD:

$$\tan\sigma = |\tan\phi(r/a)| \tag{A8}$$

or

$$\tan\sigma = |\tan\phi(p/2a)| \tag{A9}$$

where p is pupil diameter.

All of these calculations are based on geometrical optics. Blurring of the retinal image can be modelled as a two-dimensional Gaussian function. To a first approximation, we can equate blur circle radius (σ) with Gaussian space constant.

References

Anderson, B. L. (1994). The role of partial occlusion in stereopsis. *Nature*, *367*, 365–368.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Journal of the Optical Society of America A*, *5*, 1749–1758.

Foster, D. H., & Bischof, W. F. (1997). Bootstrap estimates of the statistical accuracy of thresholds obtained from psychometric functions. *Spatial Vision*, *11*, 135–139.

Glennerster, A. (1998). D_{max} for stereopsis and motion in random dot displays. *Vision Research*, *38*, 925–935.

Howard I. P., & Rogers B. J. (1995). *Binocular vision and stereopsis*. Oxford: Oxford University Press.

Jaschinski-Kruza, W. (1994). Dark vergence in relation to fixation disparity at different luminance and blur levels. *Vision Research*, *34*, 1197–1204.

Johnston, E., Cumming, B., & Parker, A. (1993). Integration of depth modules: stereopsis and texture. *Vision Research*, *33*, 813–826.

Marshall, J., Burbeck, C., Ariely, D., Rolland, J., & Martin, K. (1996). Occlusion edge blur: a cue to relative visual depth. *Journal of the Optical Society of America A*, *13*, 681–688.

Mather, G. (1996). Image blur as a pictorial depth cue. *Proceedings of the Royal Society of London B*, *263*, 169–172.

Mather, G. (1997). The use of image blur as a depth cue. *Perception*, *26*, 1147–1158.

Mather, G., & Smith, D. R. R. (1999). Blur and stereoscopic interactions influence depth perception. *Perception*, *28*, B80.

Nielsen, K., & Poggio, T. (1984). Vertical image registration in stereopsis. *Vision Research*, *24*, 1133–1140.

Pentland, A. P. (1987). A new sense for depth of field. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *9*, 523–531.

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

Wilcox, M., & Hess, R. (1995). D_{max} for stereopsis depends on size, not spatial frequency content. *Vision Research*, *35*, 1061–1069.

Wilson, D. (1973). A centre for accommodative vergence motor control. *Vision Research*, *13*, 2491–2503.