PERGAMON

Vision Research 40 (2000) 2645–2651

Vision Research

www.elsevier.com/locate/visres

First- and second-order processing in transient stereopsis

Mark Edwards *, David R. Pope, Clifton M. Schor

School of Optometry, University of California, Berkeley, CA 94720, USA

Received 5 September 1998; received in revised form 23 March 2000

Abstract

Large-field stimuli were used to investigate the interaction of first- and second-order pathways in transient-stereo processing. Stimuli consisted of sinewave modulations in either the mean luminance (first-order stimulus) or the contrast (second-order stimulus) of a dynamic-random-dot field. The main results of the present study are that: (1) Depth could be extracted with both the first-order and second-order stimuli; (2) Depth could be extracted from dichoptically mixed first- and second-order stimuli, however, the same stimuli, when presented as a motion sequence, did not result in a motion percept. Based upon these findings we conclude that the transient-stereo system processes both first- and second-order signals, and that these two signals are pooled prior to the extraction of transient depth. This finding of interaction between first- and second-order stereoscopic processing is different from the independence that has been found with the motion system. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: First order; Second order; Stereopsis; Transient

1. Introduction

Two distinct mechanisms appear to be involved in the processing of stereoscopic-depth information. These systems differ in terms of their temporal and disparity tuning, and also in the nature of the perceived depth they generate. One system can process highly-diplopic and briefly-presented stimuli. The percept of depth generated by this system fades with long presentations, (Ogle, 1952). The other system processes dichoptic stimuli that are substantially within Panum's fusional area and requires longer stimulus durations. The percept of depth generated by this system is sustained in nature (Ogle, 1952; Westheimer & Tanzman, 1956). Also, observers can show asymmetries in their sensitivity to crossed and uncrossed depths with briefly presented stimuli that they do not show for long duration stimuli (Richards, 1971; Richards & Foley, 1971). Ogle labeled these systems qualitative and quantitative, respectively, due to his observation that the perceived depth mediated by the quantitative system varies with the magnitude of the disparity, while the qualitative system merely gives the sign of the depth. However, Richards and Kaye (1974) observed quantitative variations in stereo depth stimulated with brief duration disparities up to 4° in magnitude. Also, the qualitative system can process stimuli that are within Panum's area (Pope, Edwards & Schor, 1998). Based upon these observations, it would appear that the defining differences between these two systems are their respective temporal sensitivities and upper-disparity limits (D_{max}). Accordingly, we describe them in a way analogous to the description of the transient and sustained components of the disparity-vergence system (Jones, 1980).

In visual processing, a useful distinction can be made between first-order and second-order stimuli. First-order stimuli are defined by differences in either luminance or colour while second-order stimuli are defined by variations in these properties, e.g. contrast and texture variations (Badcock & Derrington, 1985; Chubb & Sperling, 1988, 1989; Cavanagh & Mather, 1989; Scott-Samuel & Georgeson, 1999). This distinction was initially made in the domain of motion perception (Cavanagh & Mather, 1989) however, a number of authors have provided evidence of second-order processing in stereo-perception (e.g. Edwards, Pope & Schor, 1998; Sato & Nishida, 1993; Hess & Wilcox, 1994). The extraction of first-order information can be achieved by using standard linear mechanisms while the

* Corresponding author. Fax: + 1-510-643-5109.
E-mail address: mark@hering.berkeley.edu (M. Edwards).

PII: S0042-6989(00)00126-7

extraction of second-order information requires a non-linear stage (Chubb & Sperling, 1988; Wilson, Ferrera & Yo, 1991). While it is theoretically possible to extract both first- and second-order information via a common pathway (e.g. Johnston, McOwan & Buxton, 1992) a number of studies have provided good evidence for the existence of separate and parallel first-order and second-order motion pathways (e.g. Derrington & Badcock, 1985; Nishida & Sato, 1993; Ledgeway & Smith, 1994; Edwards & Badcock, 1995; Nishida, Ledgeway & Edwards, 1997).

We have previously shown that the transient stereo-system exhibits broadband dichoptic-tuning to both the orientation and spatial-frequency of gabor stimuli and that it can extract depth from opposite-polarity gaussian stimuli (Edwards et al., 1998; Pope et al., 1998; Schor, Edwards & Pope, 1998). This pattern of results is indicative of performance being mediated by a second-order pathway. There is, however, no compelling evidence for the involvement of a first-order pathway in transient stereo-processing. The stimuli used in our previous studies consisted of spatially-localised stimuli. By their nature, spatially-localised stimuli result in a local variation in contrast. Since all putative second-order models essentially extract local-contrast variations, they will be sensitive to such stimuli. Specifically, popular second-order models incorporate three distinct processing levels: an initial filtering stage, some form of non-linearity (typically half-or full-wave rectification) and then a second filtering stage. The centre spatial-frequencies of the two filtering stages differ, with the second-stage typically being lower than the first by one octave (Wilson et al., 1991; Zhou & Baker, 1993). Consequently, such second-order models are sensitive to stimuli that have broadband energy in the spatial-frequency domain, e.g. spatially-localised stimuli, while being insensitive to stimuli having narrowband energy in the frequency domain, e.g. large-field luminance sinewaves (Edwards & Badcock, 1995).

The aims of the present study are 2-fold. The first aim is to determine whether the transient-stereopsis system is sensitive to both (pure) first-order and second-order stimuli. Assuming that the system is sensitive to both types of stimuli, the second aim is to determine whether the first- and second-order signals are pooled prior to the extraction of transient depth.

2. Experiment 1: first- and second-order sensitivity?

The aim of this experiment was to determine whether observers could extract depth with first-order and second-order stimuli. Spatially-extended stimuli were employed so that a pure first-order stimulus could be generated. This was possible since spatially-extended stimuli do not lead to a local variation in contrast, and hence are not second-order stimuli. In other words, the narrow-band spatial-frequency content of such stimuli cannot pass through both second-order filtering stages. The stimuli were large-field luminance-modulated random-dot fields (a pure first-order stimulus) and large-field contrast-modulated random-dot fields (a pure second-order stimulus). Note that the latter stimulus is a pure second-order stimulus only if the luminance increments and decrements are correctly matched (Nishida, Edwards & Sato, 1997).

2.1. Method

2.1.1. Observers

Three male observers were used, the two authors and an observer who was naive with respect to the aims of the experiment. All had normal (ZZ) or corrected to normal (CS & ME) visual acuity, normal stereopsis (as measured by a Randot Stereotest) and no history of any binocular visual-disorders.

2.1.2. Stimuli and procedure

First-order and a second-order vertical-grating stimuli were used. The starting point for both stimuli was a dynamic-random-dot field. That is, a field completely filled by rectangular elements ($0.08 \times 0.13°$) that were randomly assigned to be either light or dark. The contrast was set to 25%. The first-order stimulus consisted of a 25% sinewave modulation in the luminance of the random-dot field. The second-order stimulus consisted of a 25% sinewave modulation in the contrast of the random-dot field. Note that any non-symmetry either in the output of the monitor to the luminance increments and decrements, or in the response of the visual system to them, will introduce a first-order artefact (i.e. systematic luminance information) to the second-order stimulus. The use of a low second-order contrast reduces the chances of this occurring (Henning, Hertz & Broadbent, 1975). This issue is addressed more fully in Experiment 2. The mean luminance of the display was 20 cd/m^2. As viewed through the stereo-goggles, this luminance was reduced to about 3 cd/m^2. The horizontal spatial-extent of all stimuli was 37°, the largest field size we could achieve with our viewing distance of 0.5 m, and the spatial frequency of both modulations (luminance and contrast) was 0.25 cycles/deg. Each grating had a height of 16° (and a width of 37°). The temporal frequency of the dynamic carrier (dynamic random-dots) used was 60 Hz — half of the monitor frequency of 120 Hz. A disparity offset of 1° (90° phase offset of the modulator) was used. The disparity of the edges of the envelope was 0° (i.e. fixation depth). These combinations of disparity and spatial frequency were used since pilot studies indicated that good performance with the current stimuli was achieved with them.

The observer first maintained binocular fixation on a pair of crosses and vertical nonius lines that were located at the centre of the screen. Once the observer had established binocular fixation, while perceiving the nonius lines aligned, he initiated the presentation of the test stimulus. The test stimulus replaced the fixation cross and nonius lines and was presented for 200 ms. The stimulus configuration was tailored to the particular observer in order to achieve optimal performance. For observers ZZ and CS the test stimulus consisted of a single grating (37 × 16°) whose upper edge was 0.7° below the fixation point at either a crossed or uncrossed disparity. For ME, the test stimulus was a pair of gratings (each 37 × 16°) one above and the other below the fixation point with one at a crossed and the other at an uncrossed disparity. A vertical gap of 1.4° was maintained between the upper and lower stimuli. Different stimulus configurations were employed since we have found that crowding effects between vertically-offset stimuli at opposite depths can impair performance for some observers. This effect is probably linked to individual variations in the disparity pooling region, i.e. a form of depth crowding (Westheimer & Truong, 1988) employed by the transient-stereo system and is currently the focus of research in this laboratory. Stimuli were presented in blocks of 20, within which the type of stimulus used in each trial (first-order or second-order) was kept constant. Each data point reported represents the mean of ten blocks of trials.

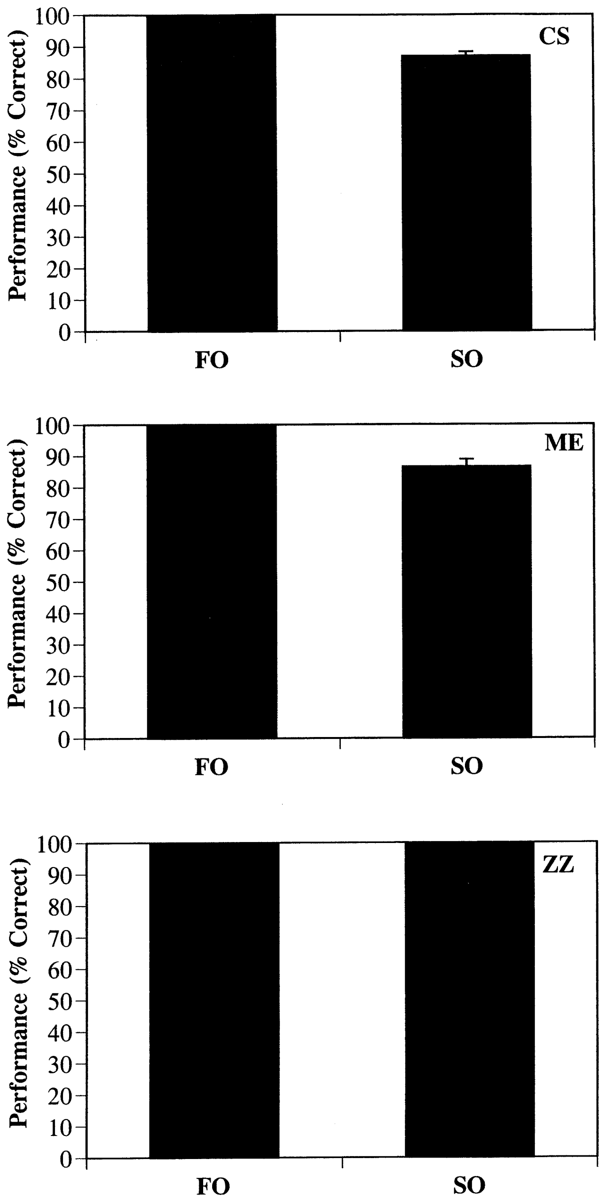

Fig. 1. The results for the three observers for experiment 1. Performance, measured as percentage of the responses that were correct, is plotted for both the first-order (FO) and second-order (SO) conditions. Error bars represent ± 1 SEM of the ten data blocks. All observers could extract depth with both types of stimuli at a perfect, or near perfect, performance level.

2.1.3. Apparatus

Stimuli were generated using a Cambridge Research Systems VSG 2/3 graphics card in a host Pentium computer and displayed on a custom Model 3 Vision Research Graphics monitor at a luminance resolution of pseudo 12 bits. The dichoptic half-images were selectively presented to each eye via Vision Research Graphics ferro-electric shutters. The fast decay rate of the monitors P46 phosphor (0.1 μs to 10% of the phosphor's initial luminance value) ensured that there was no interocular cross-talk via the shutters. The frame rate of the monitor was 120 Hz so that the effective frame rate to each eye was 60 Hz. At this frame rate there was no noticeable flicker of the images. The observer initiated each trial and responded via a button box. A chin rest was used to stabilise the observer's head.

2.2. Results

The results for the three observers are shown in Fig. 1. Performance, measured as percentage of correct responses, is plotted for both the first-order (FO) and second-order (SO) conditions. Error bars represent plus and minus one standard error of the means. All observers could extract depth with both types of stimuli. It is important to note, however, that this was a relatively difficult task and that a number of people who were tested could not extract depth, especially with the second-order stimulus. That the second-order version of the task is more difficult than the first-order version is consistent with previous findings, e.g. Li-Ming & Wilson, 1996. Also note that 'correct' performance does not necessarily mean that the observer saw the depth in the direction that corresponded to the nearest-neighbour match — see Edwards and Schor (1999). The present results indicate that both first- and second-or-

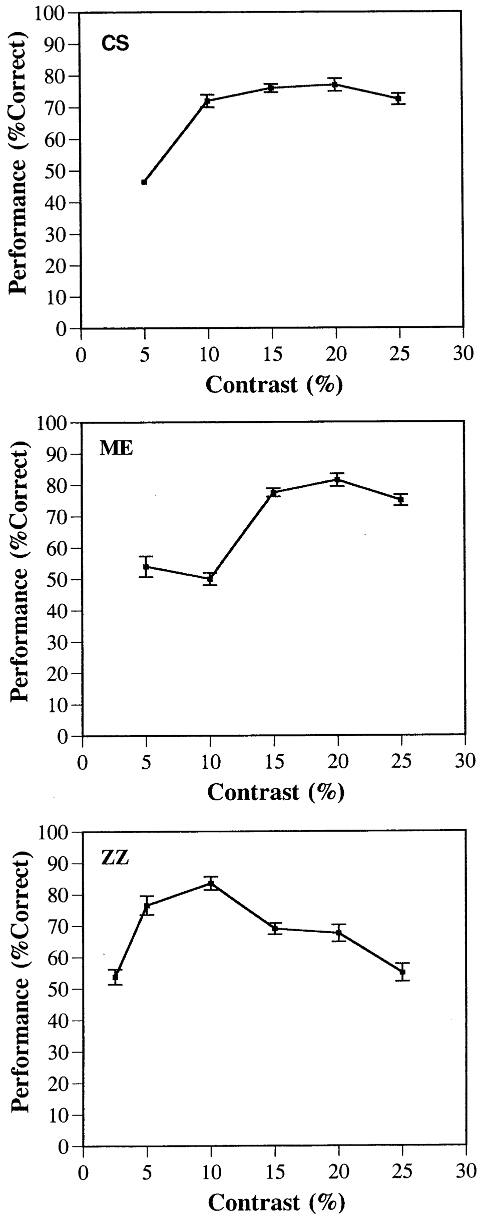

Fig. 2. The results for the three observers for matched-frequency condition in experiment 2. Performance (% correct) is plotted against the contrast of the first-order sinewave modulation.

der information are processed by the transient-stereo system.

3. Experiment 2: single or dual pathways?

This experiment addresses the issue of whether first- and second-order signals are processed by a single pathway or by two independent pathways. While it is theoretically possible to extract both first- and second-order information via a common (non-linear) pathway (e.g. Johnston et al., 1992) a number of studies have provided good evidence that, at least in the domain of motion perception, separate and parallel first-order and second-order pathways exist (e.g. Derrington & Badcock, 1985; Ledgeway & Smith, 1994; Edwards & Badcock, 1995; Nishida et al., 1997).

3.1. Stimuli and procedure

The stimuli used where the same as those used in the previous experiment (FO + SO condition). The experimental procedure consisted of dichoptically-paired first-order and second-order stimuli. This technique is a variation of the one employed by Ledgeway and Smith (1994) to investigate the same issue within the motion domain. With dichoptically-mixed first- and second-order stimuli, observers would only be able to extract depth if the first- and second-order signals are pooled prior to the extraction of transient-stereoscopic depth.

Two conditions were employed, a matched-frequency conditions in which the spatial frequency of both stimuli were the same (0.25 cpd) and a mixed-frequency condition in which the frequency of the second-order stimulus was twice that of the first-order stimulus. The second condition was employed since any single pathway that implements fullwave rectification would result in frequency doubling of the first-order stimulus but not of the second-order stimulus. Three observers were tested with the first condition, and two (CS & ME) on the second. Also, to test for any possible contrast tuning (Halpern & Blake, 1988; Legge & Gu, 1989; Schor & Heckman, 1989; Schor et al., 1998) a range of first-order contrasts were tested; 25, 20, 15, 10 and 5%. Second-order contrast was held constant at 25%.

3.2. Results

The results for the matched-frequency condition are shown in Fig. 2 and those for the frequency-doubled condition in Fig. 3. For the matched-frequency condition, two of the observers (CS & ME) showed above chance performance at the highest contrast used, and then minimal variation in performance (slight improvement for ME) as the first-order contrast was decreased. Observer ZZ, on the other hand, showed chance performance at the lowest first-order contrast-level, and then improved performance as the contrast was reduced. Even at 5% contrast ZZ, was still showing above-chance performance. Further testing at 2.5% contrast resulted chance-level performance. For the frequency-doubled condition, neither of the observers could extract depth (Fig. 3).

3.3. First-order artefact?

It is possible that the above chance performance for the FO + SO condition was due to the presence of a first-order artefact in the second-order stimulus. Such an artefact could occur due to asymmetries in the generation of, or in the visual system's response to luminance increments and decrements. Observers would then have been able to extract depth from the FO + SO condition since the binocular stimulus would have essentially consisted of two first-order dichoptic components. To test for this possibility, we converted the stereo stimulus into a motion stimulus. One of the stimuli that formerly comprised half of the dichoptic pairing was presented continuously and binocularly for 75 ms. It was then followed by the other stimulus that formerly comprised the dichoptic pair. The spatial-phase offset between the two stimuli was the same as that used in the stereo experiments; 90° phase offset of the carrier. The stimuli were viewed through the shutters. With such a stimulus, pairing first-order or second-order stimuli lead to a percept of motion in the direction of displacement. However, when the first- and second-order stimuli were paired, unambiguous motion was not perceived. Instead, a flickering, jumping sense of motion was perceived; one that did not result in any consistent or stable sense of motion direction. Such a finding is consistent with the results of Ledgeway and Smith (1994) and indicates that any first-order artefacts present in the second-order stimulus were not functionally significant. Thus the present finding of above-chance stereo performance with dichoptically-mixed first- and second-order stimuli can be interpreted as indicating that those signals are pooled prior to binocular processing.

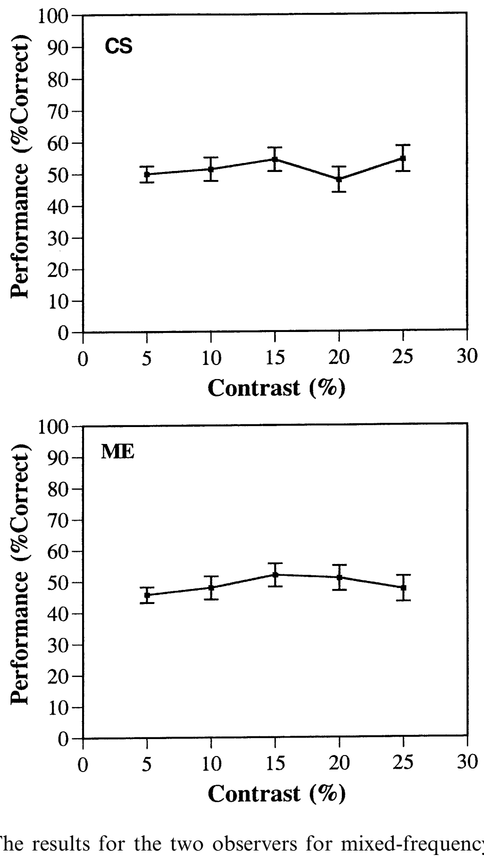

Fig. 3. The results for the two observers for mixed-frequency condition in experiment 2. Performance (% correct) is plotted against the contrast of the first-order sinewave modulation. For both observers, depth could not be extracted from either condition, regardless of the contrast of the first-order stimuli.

Note that this form of motion stimulus forms a powerful test as to the presence of a first-order artefact in the stimulus. This is because the stage in the motion system where signals are pooled appears to be multiplicative in nature. Hence any signal (first-order artefact) present in one motion frame can be amplified by pairing it with a stronger version of the (first-order) signal in the following motion frame. See Lu and Sperling (1999) for further details. The stereo system does not appear to use a simple multiplicative pooling stage and would therefore be less sensitive to such artefacts (Kontsevich & Tyler, 1994; Schor et al., 1998).

4. General discussion

The main results of the present study are that: (1) depth could be extracted with large-field first-order and second-order stimuli; and (2) depth could be extracted from stimuli that consisted of dichoptically mixed first- and second-order stimuli, but not when the spatial frequency of the second-order stimulus was twice that of the first-order one.

Spatially-localised 'first-order' stimuli are also effectively second-order stimuli, since they produce local variation in contrast. Thus, the use in the present study of large-field stimuli, with no disparity offset to the edges of their spatial envelopes, allowed us to construct stimuli that where either pure first-order or second-order in nature. Using these stimuli we were able to establish that the transient stereo-system is sensitive to both types of stimuli. Our earlier studies (Schor et al., 1998; Edwards, Pope & Schor, 1999; Pope, Edwards & Schor, 1999) had provided good evidence for the existence of a second-order transient stereo-system but until the present findings, there was no good evidence for the existence of a first-order transient stereo-system.

The finding that depth could be perceived when first- and second-order stimuli were dichoptically mixed supports the notion that these two pathways are pooled prior to the extraction of transient depth. This finding of interaction between first- and second-order processing is different from what has been found with the motion system (Derrington & Badcock, 1985; Nishida & Sato, 1993; Ledgeway & Smith, 1994; Edwards & Badcock, 1995). There are a number aspects of the performance level with the mixed condition that are worth noting, namely: (1) Performance was not as good for the mixed (first- plus second-order) condition as for the matched (first- or second-order) ones; (2) For observer ZZ, performance on the mixed condition improved markedly as the contrast of the first-order stimulus was initially decreased; and (3) Performance for the frequency-doubled condition was at chance. The first two points are probably, to an extent, related. The marked variation in ZZ's performance with contrast indicates the presence of contrast tuning (Halpern & Blake, 1988; Legge & Gu, 1989; Schor & Heckman, 1989; Schor et al., 1998) with this task. Evidence of such tuning is also evident, to a far lesser degree, in the results for ME (Fig. 2). Thus, the failure to match the performance obtained in the mixed (first- plus second-order) condition with the matched (first- or second-order) conditions may be in part due to a failure to use the same effective contrast for the two types of stimuli. The third point suggests that the transient system does not implement fullwave rectification. This is surprising since we have previously presented evidence that supports such a non-linear stage (Pope et al., 1999). So why was there a failure to perceive depth with the frequency-doubled condition in the present study? We have considered two possible reasons. The first is that the previous study used gaussian stimuli. It is possible that the putative fullwave-rectifying second-order pathway may not be able to process the more complex patterns used in the present system. Instead, performance may have been mediated by a halfwave-rectifying system. The other possibility is that the putative-fullwave system may have been able to process the current stimuli, but that its performance level for the mixed condition may have been too low. In the previous study, while performance with the mixed-contrast-polarity gaussians was above chance levels, it was not as good as the matched contrast-polarity condition (Pope et al., 1999).

Based upon our studies to date, it appears that, prior to binocular depth processing, the transient-stereopsis system can pool across filters tuned to many different features of the stimulus. These features include: spatial frequency, orientation, luminance polarity and the order of the stimulus feature, i.e. first- and second-order (Edwards et al., 1998; Pope et al., 1998; Schor et al., 1998).

Acknowledgements

This work was supported by NEI Grant EYO 8882.

References

Badcock, D. R., & Derrington, A. M. (1985). Detecting the displacement of periodic patterns. *Vision Research*, *25*, 1253–1258.

Cavanagh, P., & Mather, G. (1989). Motion: the long and the short of it. *Spatial Vision*, *4*(2/3), 103–129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America*, *A*, *5*, 1986–2006.

Chubb, C., & Sperling, G. (1989). *Second-order motion perception: Space/time seperable mechanisms. Proceedings: Workshop on visual motion* (pp. 126–138). Irvine, CA: IEEE Computer Society Press.

Derrington, A. M., & Badcock, D. R. (1985). Separate detectors fir simple and complex grating patterns? *Vision Research*, *12*, 1869–1878.

Edwards, M., & Badcock, D. R. (1995). Global motion perception: no interaction between the first- and second-order motion pathways. *Vision Research*, *35*, 2589–2602.

Edwards, M., & Schor, C. M. (1999). Depth aliasing by the transient stereo-system. *Vision Research*, *39*, 4333–4340.

Edwards, M., Pope, D. R., & Schor, C. M. (1999). Orientation tuning of the transient-stereopsis system. *Vision Research*, *39*, 2717–2727.

Halpern, D. L., & Blake, R. (1988). How contrast affects stereoacuity. *Perception*, *17*, 3–13.

Henning, G. B., Hertz, B. G., & Broadbent, D. E. (1975). Some experiments bearing on the hypothesis that the visual system analyses spatial patterns in independent bands of spatial frequency. *Vision Research*, *15*, 887–897.

Hess, R. F., & Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision Research*, *34*, 2431–2438.

Jones, R. (1980). Fusional vergence: sustained and transient components. *American Journal of Optometry and Physiological Optics*, *57*(9), 640–644.

Johnston, A., McOwan, P. W., & Buxton, H. (1992). A computation model of the analysis of some first-order and second-order motion patterns by simple and complex cells. *Proceedings of the Royal Society*, *250*, 297–306.

Kontsevich, L. L., & Tyler, C. W. (1994). Analysis of stereo thresholds for stimuli below 2.5 c/deg. *Vision Research*, *34*(17), 2317–2329.

Ledgeway, T., & Smith, A. T. (1994). Evidence for separate motion-detecting mechanisms for first- and second-order motion in human vision. *Vision Research*, *34*, 2727–2740.

Legge, G. E., & Gu, Y. (1989). Stereopsis and contrast. *Vision Research*, *29*(8), 989–1004.

Li-Ming, L., & Wilson, H. R. (1996). Fourier and non-fourier pattern discrimination compared. *Vision Research*, *36*, 1907–1918.

Lu, Z. L., & Sperling, G. (1999). The mechanism of isoluminant chromatic motion perception. *Proceedings of the National Academy of Science USA*, *96*(14), 8289–8294.

Nishida, S., Edwards, M., & Sato, T. (1997). Simultaneous motion contrast across space: involvement of second-order motion? *Vision Research*, *37*, 199–214.

Nishida, S., Ledgeway, T., & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Research*, *37*, 477–490.

Nishida, S., & Sato, T. (1993). Two kinds of motion after effect reveal different types of motion processing. *Investigative Ophthalmology and Visual Science*, *34*, 3262.

Ogle, K. (1952). On the limits of stereoscopic vision. *Journal of Experimental Psychology*, *44*, 253–259.

Pope, D. R., Edwards, M., & Schor, C. M. (1999). Extraction of depth from opposite-contrast stimuli: transient system can, sustained system can't. *Vision Research*, *39*, 4010–4017.

Richards, W. (1971). Anomalous stereoscopic depth perception. *Journal of the Optical Society of America*, *61*, 410–414.

Richards, W., & Foley, J. M. (1971). Interhemispheric processing of binocular disparity. *Journal of the Optical Society of America*, *61*, 419–421.

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

Sato, T., & Nishida, S. (1993). Second-order depth perception with texture-defined random-dot stereograms. *Investigative Ophthalmology and Visual Science*, *34*, 1438.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does an early non-linearity account for second-order motion? *Vision Research*, *39*, 2853–2865.

Schor, C. M., Edwards, M., & Pope, D. R. (1998). Spatial-frequency tuning of the transient-stereopsis system. *Vision Research*, *38*, 3057–3068.

Schor, C. M., & Heckman, T. (1989). Interocular differences in contrast and spatial frequency: effects on stereopsis and fusion. *Vision Research*, *29*, 837–847.

Westheimer, G., & Tanzman, I. J. (1956). Qualitative depth localization with diplopic images. *Journal of the Optical Society of America*, *46*(2), 116–117.

Westheimer, G., & Truong, T. T. (1988). Target crowding in foveal and peripheral stereoacuity. *American Journal of Optometry and Physio Optics*, *65*(5), 395–399.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1991). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, *9*, 79–97.

Zhou, Y., & Baker, C. L. (1993). A processing stream in mammalian visual cortex neurons for non-fourier responses. *Science* (*New York*), *261*, 98–101.