

# Co-operative interactions between first- and second-order mechanisms in the processing of structure from motion

**Craig R. Aaen-Stockdale**

McGill Vision Research, McGill University,  
Montreal, Canada, &  
School of Optometry and Visual Sciences,  
University of Bradford, Bradford, UK



**Reza Farivar**

McGill Vision Research, McGill University,  
Montreal, Canada



**Robert F. Hess**

McGill Vision Research, McGill University,  
Montreal, Canada



Structure from motion (SFM) is the ability to perceive three-dimensional structure from stimuli containing only two-dimensional motion signals and this ability seems to be a result of high-level cortical processes. It has long been thought that local motion signals defined by second-order cues only weakly contribute to perception of SFM since performance on purely second-order SFM tasks is poor, relative to first-order stimuli. We hypothesized that the mechanisms responsible for deriving SFM were insensitive to low-level stimulus attributes such as the first- or second-order nature of the dots composing the stimulus, in other words: that they were “cue-invariant”, but that large differences in sensitivity to local first- and second-order motions were responsible for previous findings. By manipulating the relative strength of first-order dots in an SFM stimulus that combines first- and second-order dots, we show that the two types of motion can separately support SFM and co-operatively interact to produce vivid three-dimensional percepts. This provides strong support that the mechanisms underlying SFM are cue-invariant.

Citation: Aaen-Stockdale, C. R., Farivar, R., & Hess, R. F. (2010). Co-operative interactions between first- and second-order mechanisms in the processing of structure from motion. *Journal of Vision*, 10(13):6, 1–9, <http://www.journalofvision.org/content/10/13/6>, doi:10.1167/10.13.6.

## Introduction

The extraction of three-dimensional (3D) structure from moving two-dimensional stimuli, often referred to as the “kinetic depth effect” or structure from motion (SFM), has a long history in the study of vision. Several models have been developed to explain how the visual system achieves this remarkable feat, some based on the tracking of local positional cues (Grzywacz & Hildreth, 1987; Shariat & Price, 1990; Snowden, Treue, Erickson, & Andersen, 1991; Ullman, 1984) and others based on the use of local motion information (Clocksin, 1980; Hildreth, Ando, Andersen, & Treue, 1995; Husain, Treue, & Andersen, 1989; Koenderink & van Doorn, 1986; Longuet-Higgins & Prazdny, 1980). The weight of evidence currently seems to suggest that motion rather than positional information is crucial to extracting SFM (Andersen & Bradley, 1998; Farivar, 2009; Farivar, Blanke, & Chaudhuri, 2009).

Provided that the dot density is high enough, SFM can be perceived with very short (two frame) displays (Lappin, Doner, & Kottas, 1980). SFM is not perceived

if a small number of dots is used. However, periodically repositioning a similarly small number of dots across the stimulus allows the surface of the object to be reconstructed via interpolation (Husain et al., 1989; Treue, Andersen, Ando, & Hildreth, 1995).

Local motions can be defined by changes in luminance (first-order motion) or other characteristics such as flicker, texture, and contrast. Motion that is defined by modulation of some property other than luminance is referred to as “second-order” motion (Cavanagh & Mather, 1989) and local second-order motion is invisible to first-order motion sensors (Chubb & Sperling, 1988).

An outstanding issue has been whether, or to what extent, the SFM system receives input from local motion defined by second-order characteristics. Wire frame shapes defined by flicker modulation of a noise background were reported to produce a 3D percept (Prazdny, 1986). However, in that study 2D cues may have played a role in the shape identification task required of the observer (Doshier, Landy, & Sperling, 1989). Subsequent work has suggested that the second-order input to the SFM system is weak or non-existent. Mather (1989)

showed that the kinetic depth effect was only supported by “short-range” or what are now referred to as first-order local motion signals. Doshier et al. (1989) found that the first-order (or “Fourier”) system was dominant in the extraction of 3D shape from motion but conceded that their stimuli (composed of small dots) may not have been suitable to stimulate the second-order (or “non-Fourier”) pathway, which operates over larger spatial scales. Landy, Doshier, Sperling, and Perkins (1991) improved greatly on Doshier et al.’s experiment by using a variety of stimuli, tailored toward activating the second-order pathway. Their results show near perfect performance for first-order conditions, mixed performance for reversed-phi stimuli, and poor or chance performance for second-order stimuli. Hess and Ziegler (2000) obtained similar results with a simpler task that involved discriminating the orientation of a depth-modulated grating defined by 2D motion. The contrast of the first- and second-order elements in this study was equated, but this manipulation may not have equated their visibility to first- and second-order motion detectors, as the second-order motion system has weaker directional selectivity (Ledgeway & Hess, 2002) and poorer temporal sensitivity (Derrington, Badcock, & Henning, 1993).

It is likely that first-order and second-order motions are initially analyzed in parallel by separate processing streams (Ledgeway & Smith, 1994; Nishida, Ledgeway, & Edwards, 1997; Pavan, Campana, Guerreschi, Manassi, & Casco, 2009; Smith & Ledgeway, 1997), but models of motion processing usually integrate these two streams at, or before, the level of global motion analysis (Lu & Sperling, 1995, 2001; Wilson, Ferrera, & Yo, 1992) and corresponding interactions between the first- and second-order pathways have been demonstrated for higher level motion processes such as global motion (Edwards & Nishida, 2004; Ledgeway, Hess, & McGraw, 2002) and biological motion (Aaen-Stockdale, Thompson, Hess, & Troje, 2008) suggesting that the two pathways are fully integrated at these stages.

Previous work (Badcock & Khuu, 2001; Edwards & Badcock, 1995) had suggested that the two pathways were fully independent, as luminance-modulated noise dots (visible to luminance-based mechanisms, but also to second-order mechanisms) could mask a signal carried by contrast-modulated dots, but contrast-modulated noise dots (invisible to luminance-based mechanisms, and only visible to second-order mechanisms) could not mask a signal carried by luminance-modulated dots. However, more recently, Ledgeway et al. (2002) demonstrated that when the luminance modulation of the first-order dots was reduced, contrast-modulated dots could mask luminance-modulated dots, and conversely, masking of contrast-modulated (CM) dots by luminance-modulated (LM) dots was relieved by the same manipulation. This suggests that the first- and second-order pathways were integrated at the level of global motion, but the local second-order signals were weaker and thus carried less weight in subsequent

global motion computations. The relative strength of the two types of local elements in the stimulus dictates whether they appear to interact or not, highlighting the importance of ensuring equal excitation of local first-order and second-order inputs when comparing the two systems at higher levels. This technique has since been used to demonstrate that the second-order system can fully support the perception of biological motion, and that the mechanisms responsible for analysis of biological motion signals are cue-invariant (Aaen-Stockdale et al., 2008), i.e., they do not discriminate between first- and second-order inputs as long as they are of equal local strength. The strength of the two stimulus types can be equated by varying the relative luminance or contrast modulation (Aaen-Stockdale et al., 2008) or the size of the dots (Gurnsey & Troje, 2010).

In the light of this work, it is possible that there may be no qualitative difference between the processing of SFM defined by first- or second-order motion and it may be possible to unify the findings regarding second-order SFM with a model in which the mechanism or mechanisms responsible for extraction of SFM are “cue-invariant,” but in which the local motions are weighted by their visibility to local motion mechanisms. Previous work, which has suggested that second-order input to SFM mechanisms is weak or non-existent, may have used a methodology that simply did not give the second-order motion system an equal chance.

In this paper, we show, in contrast to previous reports, that second-order motion signals can drive a vivid percept of depth via SFM. If the visibility of the first-order elements is reduced, performance drops to be equal to or worse than the second-order system. Furthermore, we demonstrate, by combining first- and second-order dots in a single SFM display, that when the performance of the two systems is equated, constant performance is obtained regardless of the relative proportion of first- or second-order dots in the stimulus. SFM thresholds for a display composed of both first- and second-order dots are entirely predictable from the relative sensitivity of the two systems and the proportion of first- to second-order dots. Both types of motion contribute to the percept of 3D SFM, but the second-order motion pathway is limited by the weaker local motion signals it generates. This finding strongly implicates a cue-invariant mechanism in the processing of SFM.

## Methods

### Observers

The observers were two of the authors (CAS and RF) and two observers who were naive to the hypothesis (RD and AM).

## Apparatus

Stimuli were presented on a CRT monitor (Mitsubishi Diamond Pro 2070SB) with a screen resolution of  $1024 \times 768$  and a frame rate of 75 Hz. The monitor screen subtended  $39.5^\circ$  at the viewing distance of 57 cm. Stimulus presentation was controlled by a 2.33-GHz MacBook Pro, on which the color look-up table was gamma-corrected. Stimuli were programmed in Matlab 7.2 using the Psychophysics Toolbox routines (Brainard, 1997; Pelli, 1997).

## Stimuli

The observers viewed an SFM display composed of dots of a constant size, the 2D velocity of which varied according to their position on a 3D “object” rotating around a vertical axis of  $90^\circ$  from  $\pm 45^\circ$ , through the frontoparallel plane, to  $\pm 45^\circ$ . The object rotated  $1.5^\circ$  every frame, giving a rotation speed of  $111^\circ/\text{s}$ , and the animation consisted of 60 frames producing a total presentation duration of 800 ms. The object was a left- or right-leaning, depth-modulated Gabor patch of one cycle in sine phase, rotating in depth around the vertical axis. The observer’s task was to report in which direction the Gabor was oriented (leftward or rightward oblique). The peaks and troughs of the Gabor stimulus may be perceived as depth reversed, but this should not affect the judgment that the observer is required to make. Figure 1 shows a schematic of the stimulus.

To create the SFM stimulus, 2048 dots were projected along the  $z$ -axis (front-to-back) onto the surface of a

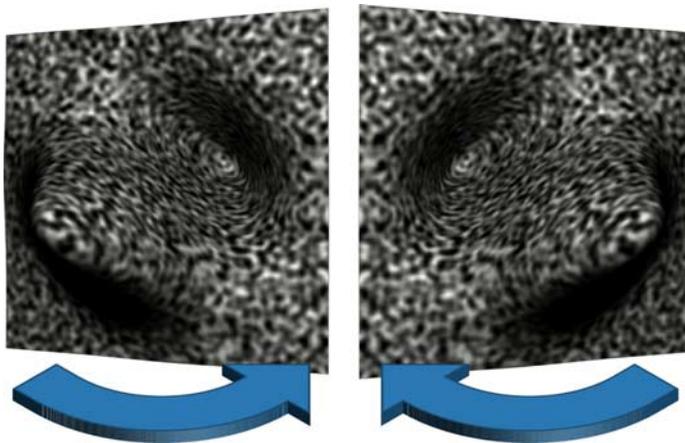


Figure 1. The “object” rotating in depth was a right- or left-leaning Gabor (here rendered visible with the use of a textured surface). On each trial, the object was randomly tilted left or right and the start position (+ or  $-45^\circ$ ) was randomized, resulting in four possible stimulus configurations (only two of which are shown here). The object was rotated through the frontoparallel plane to the opposite position. The task of the subject was to report whether the Gabor was leaning left or right.

depth-modulated Gabor patch (rendered in 3D Studio Max and saved as an .obj file). The size of the dots remained constant, regardless of their changing position in depth. The Gabor patch was randomly rotated either left–right or right–left around its vertical axis and the dots moved with the object. This obviously created areas of higher and lower dot densities, so every 15 frames (200 ms) the 200 most crowded dots were redistributed to a part of the stimulus with the lowest density. A Delauney triangulation, implemented in Matlab, is used to connect each dot location to its neighbors, and the 200 points with the smallest total distance to their neighbors were redistributed to the centers of the 200 triangles with the largest area. The duration of each trial was 800 ms. Each stimulus movie was 60 frames long presented at 75 Hz.

Apart from this replotting to control for dot density, we did not use the limited lifetime technique used in some other SFM studies (Husain et al., 1989; Treue et al., 1995) and dots remained on screen for the full stimulus presentation duration. Due to the much poorer temporal sensitivity of the second-order system (Derrington et al., 1993), limiting the lifetime of the second-order dots would almost certainly have eradicated any percept due simply to the failure of local second-order motion (Ledgeway & Hess, 2002). The main utility of limited lifetime dots is to prevent transient 2D density cues or dot tracking. The former was controlled by our dot repositioning method. Randomizing rotation direction in the stimulus should make the motion direction of any single dot uninformative, but observers could theoretically track and correlate the motion of multiple dots. This possibility is discussed later.

The background of the stimulus subtended  $19.7^\circ$  at the center of the screen and was composed of binary noise of 10% contrast. The rest of the screen was mean luminance gray. The background noise was dynamic; the noise matrix was replaced with a new sample every frame. This minimized, and hopefully eliminated, the presence of any luminance artifacts in the second-order dots. The Gabor object occupied approximately the central third of the stimulus area, having a spatial frequency of  $\sim 0.15$  cycle/degree; therefore, some dots belonged to the “object” and others belonged to the background.

The dots subtended  $0.2^\circ$  and could be either luminance modulations (LMs) or contrast modulations (CMs) of the background noise. The magnitude of the luminance modulation of the LM dots was defined as

$$\text{Luminance modulation} = (L_{\text{dot}} - L_{\text{bg}}) / (L_{\text{dot}} + L_{\text{bg}}), \quad (1)$$

where  $L_{\text{dot}}$  is the mean luminance of the noise within the dots and  $L_{\text{bg}}$  is the mean luminance of the background noise.

For the contrast-modulated stimuli, the Michelson contrast  $[(L_{\text{max}} - L_{\text{min}}) / (L_{\text{max}} + L_{\text{min}})]$  of the binary noise within the dots was increased above that of the

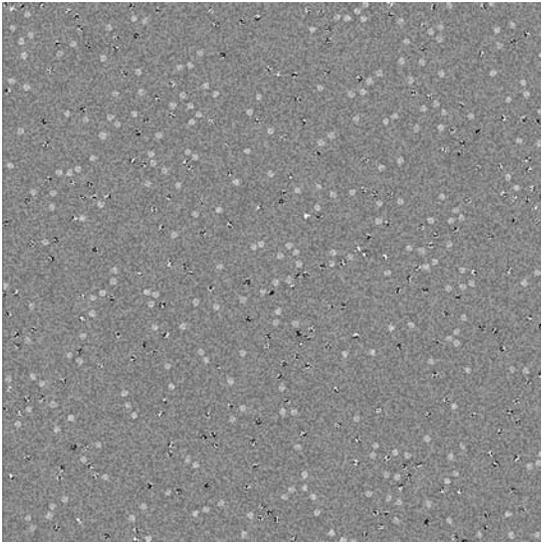


Figure 2. A movie showing a single instance of the stimulus.

background. The magnitude of the contrast modulation of the CM dots is calculated by

$$\text{Contrast modulation} = (C_{\text{dot}} - C_{\text{bg}}) / (C_{\text{dot}} + C_{\text{bg}}), \quad (2)$$

where  $C_{\text{dot}}$  is the Michelson contrast of the noise within the dots and  $C_{\text{bg}}$  is the contrast of the background noise. For the experiments discussed here, the modulation of the CM dots was always 0.8 (in other words, maximum contrast within the dots and 10% contrast outside the dots).

Figure 2 shows a movie of the stimulus; in this instance, a rightward-leaning Gabor rotating in depth composed of 50% first-order dots with a luminance modulation of 0.2 and 50% CM dots with a contrast modulation of 0.8.

## Procedure

Each trial was preceded by a variable duration fixation point, which was jittered on each trial within a  $1^\circ \times 1^\circ$  square, so that the position of the stimulus relative to fixation was varied on each trial, although not the position relative to the screen edges.

As explained above, there were 2048 dots projected onto the Gabor object, but the number of dots actually presented on each trial was varied according to a QUEST algorithm (Watson & Pelli, 1983) implemented in the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). This allowed us to obtain a threshold of the minimum number of dots in the stimulus necessary to correctly discriminate the orientation of the Gabor patch. The dots displayed were selected randomly each trial from the total cohort. Observers responded via a key press. Responding initiated the next trial.

The proportion of first- to second-order dots and the luminance of the first-order dots were varied between blocks and the order of blocks was randomized. Dots were randomly assigned to be either luminance- or contrast-modulated. The contrast-modulated dots were always a maximum modulation (0.8). The magnitude of the luminance modulation was varied and chosen by the observer before each staircase.

## Results

The results for four observers are shown in Figure 3. The threshold number of dots necessary to correctly identify the orientation of the 3D Gabor is plotted on the y-axis while the proportion of first-order dots is plotted on the x-axis. At the highest luminance contrast (black lines and markers), thresholds are very low when the proportion of first-order dots is high, as might be expected from such a highly visible stimulus. As the proportion of first-order dots relative to second-order dots decreases, the task gets progressively more difficult until it reaches a level similar to that obtained with a purely second-order stimulus (shown by the open marker).

As the luminance contrast of the first-order dots is reduced (encoded in the plots as lighter grayscale values), the sensitivity of the observer to the luminance-modulated dots is reduced and thresholds obtained with the purely first-order stimulus increase to the point where they equal (at a luminance contrast of approximately 0.025), and then exceed (at values approximately  $<0.25$ ), thresholds obtained with a purely second-order stimulus.

The data show that the first- and second-order systems not only perform similarly if the relative strength of the local motion signals is equated but also show that, as the relative number of first- and second-order dots is varied, thresholds fall predictably between the two extreme thresholds, i.e., that obtained with a stimulus that is totally first order or totally second order.

To further investigate this relationship, we normalized the thresholds obtained with an intermediate proportion of dots from each subject relative to the thresholds for an entirely FO or entirely SO stimulus. This analysis is shown in Figure 4. A stimulus composed of 50% FO dots, if it is analyzed by an SFM mechanism that is blind to SO dots, should require twice as many *total* dots to obtain a threshold. If this is the case, then that data point should fall on the right border of the square box marked “no summation” in Figure 4.

Conversely, if the mechanism extracting SFM is cue-invariant and can use first- and second-order local motions, then thresholds for a stimulus composed of 50% FO dots will simply be 50% of the FO threshold and 50% of the SO threshold. In other words, there would be perfect linear summation between the two mechanisms/

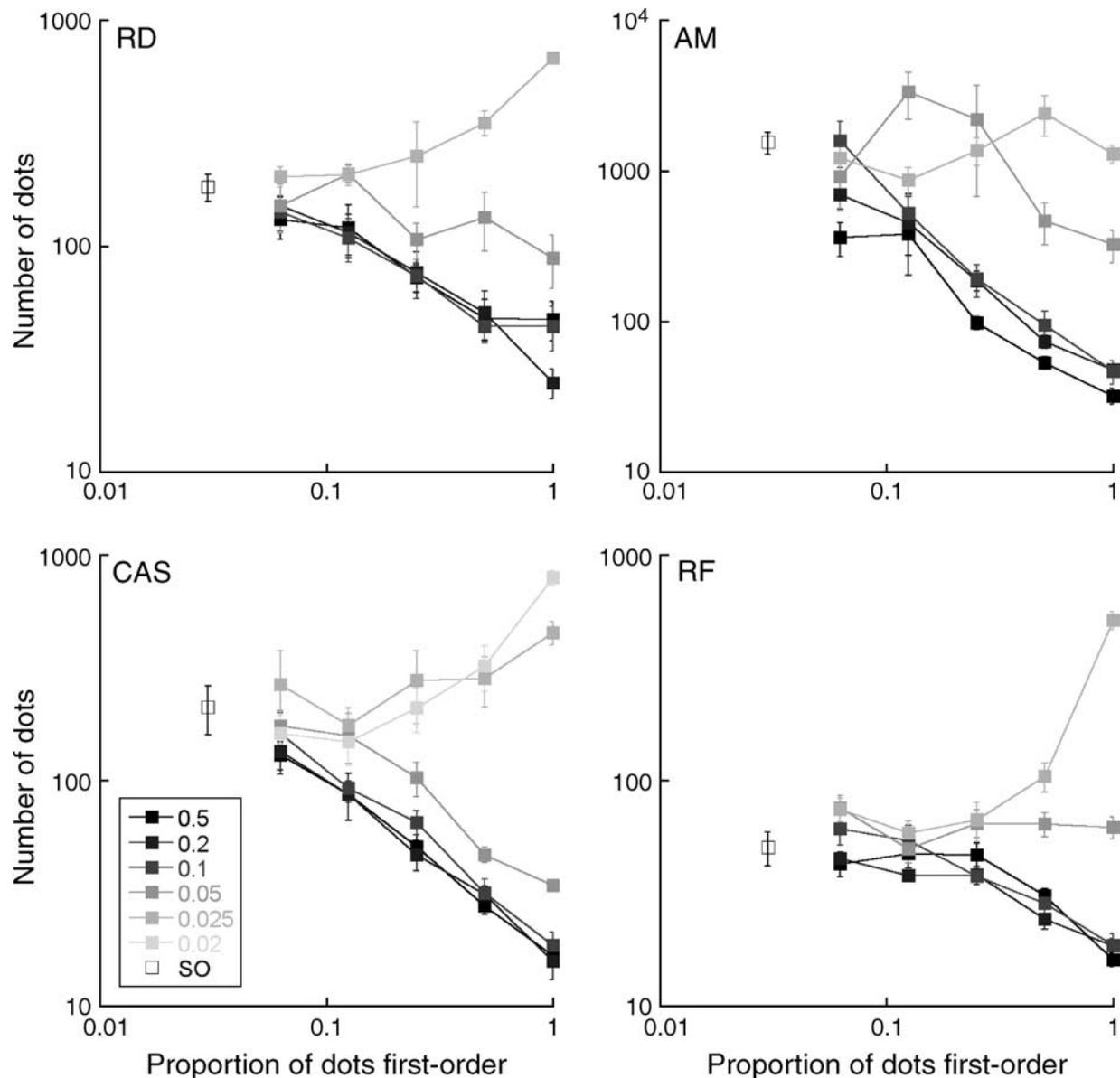


Figure 3. Data from four observers: Two authors (CAS and RF) and two observers naive to the purposes of the experiment (AM and RD). The grayscale value of the markers and lines reflects the luminance modulation of the first-order dots (given in the legend), with lighter shades reflecting smaller increments. Each data point is the mean of five threshold measurements and the error bars represent the standard error of the mean.

pathways, and the average thresholds should fall along the dotted line.

The average thresholds suggest that some co-operation is occurring between first- and second-order mechanisms in the analysis of SFM. The number of FO dots necessary to obtain a threshold is always less than that predicted by the number of FO dots in the stimulus. The data fall inside the “no summation” box. Most of the time, thresholds fall somewhere between no summation and linear summation, but the data are unfortunately very noisy.

Most of this noise appears to come from one observer, AM, whose data can be seen in Figure 3 to be much noisier, and whose thresholds for low luminance modulations and

for purely second-order stimuli are much higher, than the other observers. This observer reported having greater difficulty perceiving depth in the stimulus at low luminance and for purely second-order stimuli. It may therefore be that this observer was attempting to complete the task in these situations using an inefficient cognitive strategy such as tracking and correlating the motion of several individual dots. This observer’s normalized data are shown in Figure 5 (left) and do not conform to linear summation, but neither do they conform to a “no summation” prediction.

When observer AM’s data are excluded from the group average (Figure 5, right), the remaining observers’ data conform very well to linear summation. The slopes of the

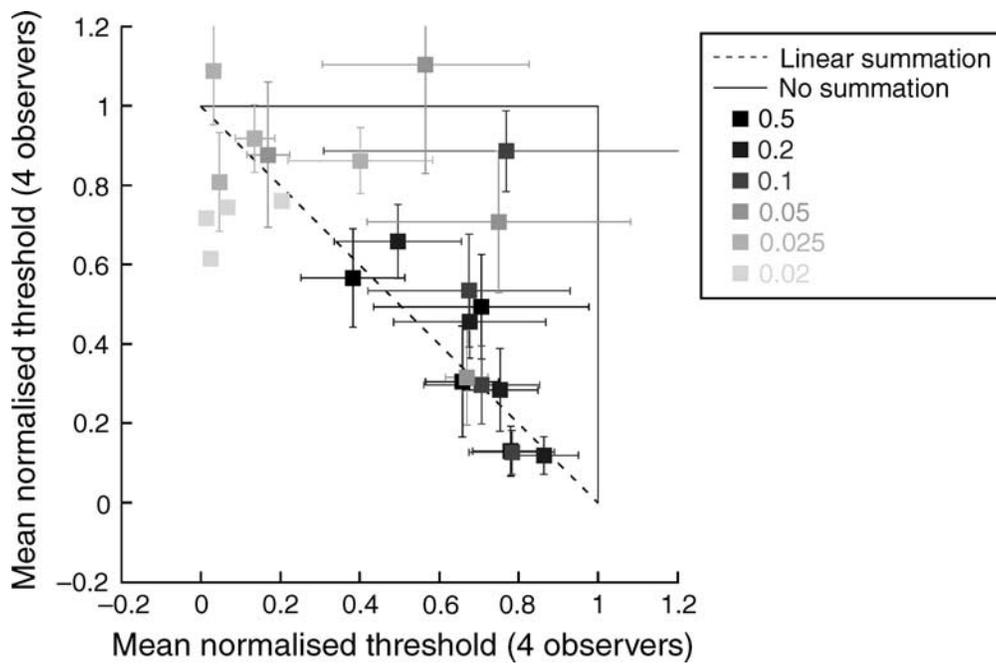


Figure 4. The average normalized thresholds for all four subjects. Each data point is the mean across observers and error bars show the standard error of the mean. The grayscale value of the markers and lines reflects the luminance modulation of the first-order dots (given in the legend), with lighter shades reflecting smaller increments. The dotted line shows the prediction from linear summation, and the solid line shows the prediction from zero summation.

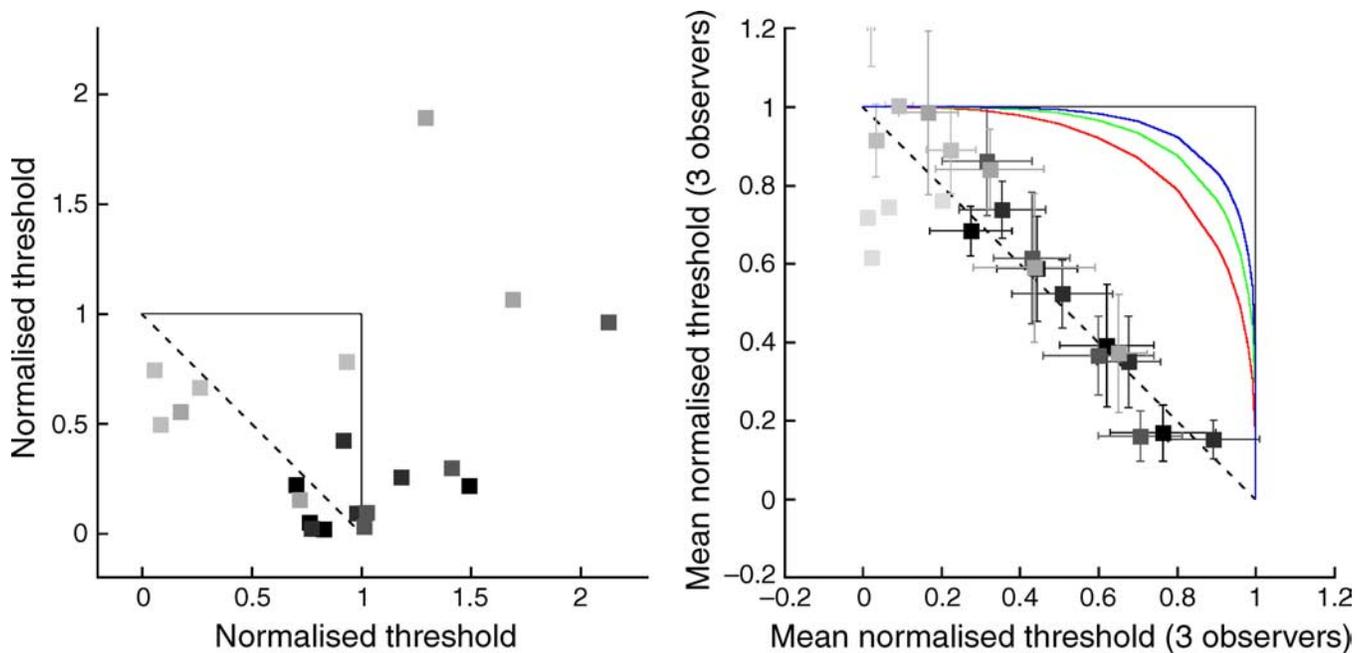


Figure 5. The left-hand plot shows the normalized thresholds for observer AM, whose data were particularly noisy and whose thresholds were noticeably higher than the other observers. His data do not conform to linear summation, but neither do they conform to no summation. The three colored lines show predictions from probability summation (Quick, 1974) with exponents of 3 (red), 4 (green), and 5 (blue). The right-hand plot shows the average normalized thresholds for the other 3 observers (CAS, RD, and RF). Each data point is the mean across these 3 observers and error bars show standard error of the mean. The grayscale value of the markers and lines reflects the luminance modulation of the first-order dots (given in the legend), with lighter shades reflecting smaller increments.

psychometric functions were extracted from the data and a mean slope of 4.3408 ( $SD$ : 1.6372) was obtained. Figure 5 shows the prediction from probability summation (Quick, 1974) for slopes of 3–5. It is clear that the data of these three observers do not correspond to probability summation of independent mechanisms.

## Discussion

The results reported here are consistent with one previous report (Prazdny, 1986), but inconsistent with several previous studies (Doshier et al., 1989; Hess & Ziegler, 2000; Landy et al., 1991; Mather, 1989) that have shown a distinct advantage for the first-order pathway in the analysis of structure from motion.

We show here that, if the strength of the first-order input is reduced (lower luminance contrast), thresholds first equal and then exceed that obtained with purely second-order stimuli. When first- and second-order dots are combined in the same stimulus and performance of the task would benefit from integrating first- and second-order local motion signals, the data of all but one observer show near perfect summation between the first- and second-order pathways. This is good evidence for involvement of a cue-invariant mechanism in the extraction of SFM.

Higher order motion perception may, in general, be cue-invariant. In our previous experiment using biological motion stimuli, we used noise dots to mask the signal and showed visibility-dependent interactions between first- and second-order pathways (Aaen-Stockdale et al., 2008). While it may be the case that the second-order system is more susceptible to noise than the first-order system, sensitivity to global motion of various types (translational, rotational, and radial) increases as the luminance or the contrast modulation of the local dots is increased, and although first-order sensitivity increases faster than second-order, the two types of motion asymptote at similar levels (Aaen-Stockdale, Ledgeway, & Hess, 2007). This suggests that, given local inputs of equal strength, the system responsible for analyzing second-order motion is just as efficient at segregating noise as the first-order. Thus, a parsimonious explanation is that first- and second-order local motion signals feed into a global motion system that is cue-invariant. Recent work has shown similar levels of performance between first- and second-order motions on a biological motion task that does not have a noise component (Gurnsey & Troje, 2010) and the experiments we outline in this paper regarding extraction of SFM did not involve masking dots, yet showed clear visibility-dependent interactions between the two pathways.

An early non-linearity in the visual system has been empirically measured (Scott-Samuel & Georgeson, 1999), which can introduce first-order motion energy into second-order stimuli, resulting in “pseudo-second-order”

motion. This distortion product becomes measurable at high noise contrasts and high temporal frequencies. Despite the fact that our stimuli were rotating quite slowly, this may have been an issue and future experiments may require very different stimuli to eliminate this possibility.

## Acknowledgments

This work was supported by a Natural Sciences and Engineering Research Council of Canada Grant (No. 46528-06 to RFH).

Commercial relationships: none.

Corresponding author: Craig R. Aaen-Stockdale.

Email: c.aaen-stockdale@bradford.ac.uk.

Address: Richmond Road, Bradford, West Yorkshire, BD7 1DP, UK.

## References

- Aaen-Stockdale, C. R., Ledgeway, T., & Hess, R. F. (2007). Second-order optic flow processing. *Vision Research*, *47*, 1798–1808.
- Aaen-Stockdale, C. R., Thompson, B., Hess, R. F., & Troje, N. F. (2008). Biological motion perception is cue-invariant. *Journal of Vision*, *8*(8):6, 1–11, <http://www.journalofvision.org/content/8/8/6>, doi:10.1167/8.8.6. [PubMed] [Article]
- Andersen, R. A., & Bradley, D. C. (1998). Perception of three-dimensional structure-from-motion. *Trends in Cognitive Sciences*, *2*, 222–228.
- Badcock, D. R., & Khuu, S. K. (2001). Independent first- and second-order motion energy analyses of optic flow. *Psychological Research*, *65*, 50–56.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4*, 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*, 1986–2007.
- Clocksink, W. F. (1980). Perception of surface slant and edge labels from optical flow: A computational approach. *Perception*, *9*, 253–269.
- Derrington, A. M., Badcock, D. R., & Henning, G. B. (1993). Discriminating the direction of second-order motion at short stimulus durations. *Vision Research*, *33*, 1785–1794.

- Dosher, B. A., Landy, M. S., & Sperling, G. (1989). Kinetic depth effect and optic flow: I. 3D shape from Fourier motion. *Vision Research*, *29*, 1789–1813.
- 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., & Nishida, S. (2004). Contrast-reversing global-motion stimuli reveal local interactions between first- and second-order motion signals. *Vision Research*, *44*, 1941–1950.
- Farivar, R. (2009). Dorsal-ventral integration in object recognition. *Brain Research Reviews*, *61*, 144–153.
- Farivar, R., Blanke, O., & Chaudhuri, A. (2009). Dorsal-ventral integration in the recognition of motion-defined unfamiliar faces. *Journal of Neuroscience*, *29*, 5336–5342.
- Grzywacz, N. M., & Hildreth, E. C. (1987). Incremental rigidity scheme for recovering structure from motion: Position-based versus velocity-based formulations. *Journal of the Optical Society of America A*, *4*, 503–518.
- Gurnsey, R., & Troje, N. F. (2010). Peripheral sensitivity to biological motion conveyed by first and second-order signals. *Vision Research*, *50*, 127–135.
- Hess, R. F., & Ziegler, L. R. (2000). What limits the contribution of second-order motion to the perception of surface shape? *Vision Research*, *40*, 2125–2133.
- Hildreth, E. C., Ando, H., Andersen, R. A., & Treue, S. (1995). Recovering three-dimensional structure from motion with surface reconstruction. *Vision Research*, *35*, 117–137.
- Husain, M., Treue, S., & Andersen, R. A. (1989). Surface interpolation in three-dimensional structure-from-motion perception. *Neural Computation*, *1*, 324–333.
- Koenderink, J. J., & van Doorn, A. J. (1986). Depth and shape from differential perspective in the presence of bending deformations. *Journal of the Optical Society of America A*, *3*, 242–249.
- Landy, M. S., Dosher, B. A., Sperling, G., & Perkins, M. E. (1991). The kinetic depth effect and optic flow: II. First- and second-order motion. *Vision Research*, *31*, 859–876.
- Lappin, J. S., Doner, J. F., & Kottas, B. L. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, *209*, 717–719.
- Ledgeway, T., & Hess, R. F. (2002). Failure of direction identification for briefly presented second-order motion stimuli: Evidence for weak direction selectivity of the mechanisms encoding motion. *Vision Research*, *42*, 1739–1758.
- Ledgeway, T., Hess, R. F., & McGraw, P. V. (2002). Masking effects between local first-order and second-order motions in the extraction of global-motion direction depend critically on stimulus visibility. *Perception*, *31*(ECP Abstract Supplement), 99.
- 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.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B: Biological Sciences*, *208*, 385–397.
- Lu, Z. L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*, 2697–2722.
- Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, *18*, 2331–2370.
- Mather, G. (1989). Early motion process and the kinetic depth effect. *Quarterly Journal of Experimental Psychology*, *41*, 183–198.
- Nishida, S., Ledgeway, T., & Edwards, M. (1997). Dual multiple-scale processing for motion in the human visual system. *Vision Research*, *37*, 2685–2698.
- Pavan, A., Campana, G., Guerreschi, M., Manassi, M., & Casco, C. (2009). Separate motion-detecting mechanisms for first- and second-order patterns revealed by rapid forms of visual motion priming and motion aftereffect. *Journal of Vision*, *9*(11):27, 1–16, <http://www.journalofvision.org/content/9/11/27>, doi:10.1167/9.11.27. [PubMed] [Article]
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Prazdny, K. (1986). Three-dimensional structure from long-range apparent motion. *Perception*, *15*, 619–625.
- Quick, R. F., Jr. (1974). A vector-magnitude model of contrast detection. *Kybernetik*, *16*, 65–67.
- Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research*, *39*, 2853–2865.
- Shariat, H., & Price, K. E. (1990). Motion estimation with more than 2 frames. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*, 417–434.
- Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: Fact or artifact? *Vision Research*, *37*, 45–62.
- Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1

- neurons to transparent motion. *Journal of Neuroscience*, *11*, 2768–2785.
- Treue, S., Andersen, R. A., Ando, H., & Hildreth, E. C. (1995). Structure-from-motion: Perceptual evidence for surface interpolation. *Vision Research*, *35*, 139–148.
- Ullman, S. (1984). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, *13*, 255–274.
- Watson, A. B., & Pelli, D. G. (1983). QUEST: Q Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*, 113–120.
- Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, *9*, 79–97.