Interaction of first- and second-order direction in motion-defined motion

Johannes M. Zanker

Department of Psychology, Royal Holloway University of London, Egham, Surrey TW20 0EX, England

Nicholas R. Burns

Department of Psychology, Adelaide University, Adelaide SA 5005, Australia

Received January 19, 2001; revised manuscript received May 16, 2001; accepted May 22, 2001

Motion-defined motion can play a special role in the discussion of whether one or two separate systems are required to process first- and second-order information because, in contrast to other second-order stimuli, such as contrast-modulated contours, motion detection cannot be explained by a simple input nonlinearity but requires preprocessing by motion detectors. Furthermore, the perceptual quality that defines an object (motion on the object surface) is identical to that which is attributed to the object as an emergent feature (motion of the object), raising the question of how these two object properties are linked. The interaction of first- and second-order information in such stimuli has been analyzed previously in a direction-discrimination task, revealing some cooperativity. Because any comprehensive integration of these two types of motion information should be reflected in the most fundamental property of a moving object, i.e., the direction in which it moves, we now investigate how motion direction is estimated in motion-defined objects. Observers had to report the direction of moving objects that were defined by luminance contrast or in random-dot kinematograms by differences in the spatiotemporal properties between the object region and the random-noise background. When the dots were moving coherently with the object (Fourier motion), direction sensitivity resembled that for luminance-defined objects, but performance deteriorated when the dots in the object region were static (drift-balanced motion). When the dots on the object surface were moving diagonally relative to the object direction (theta motion), the general level of accuracy declined further, and the perceived direction was intermediate between the veridical object motion direction and the direction of dot motion, indicating that the first- and second-order velocity vectors are somehow pooled. The inability to separate first- and second-order directional information suggests that the two corresponding subsystems of motion processing are not producing independent percepts and provides clues for possible implementations of the two-layer motion-processing network.

OCIS codes: 330.4150, 330.4060.

1. INTRODUCTION

The input to visual systems is rarely static but is characteristically affected by (and representative of) environmental motion, by movement of the observers or their eyes or by objects moving around in the world. Thus motion information is an extremely rich source of information for living organisms about their relationship to their physical and biological surroundings. As a consequence, motion-processing mechanisms have evolved to the highest levels of sophistication (for a review, see Ref. 1). It is hardly surprising, therefore, that humans are able to exploit a great variety of properties that define moving contours or objects. Typically, stimulus properties are classified as first order or Fourier because they are determined by the luminance or color attributes from a single location in the retinal image. In contrast, second-order or non-Fourier properties such as contrast, texture, flicker, or motion are determined by the relationship between the stimulus attributes from two or more stimulus locations. In order to extract this type of information, the visual system needs to employ specific nonlinear operations that compare the input signals from at least two points in time or two locations in space.[2,3] Such operations can be comparatively simple in their implementation, like the rectification of a bandpassed image, where the bandpass (which is needed to remove the average stimulus intensity) represents the input from multiple locations and the rectification represents the essential nonlinearity.[4] Because more processing stages are required for detecting higher-order motion, the number of essential nonlinear operations grows, and, correspondingly, the order of nonlinearity of the mathematical formulations increases. However, there is no unique correspondence between the mathematical order of nonlinearity of the operator required for extracting a given motion stimulus[5] and the classification of this stimulus as first or second order. Therefore it has been suggested that the terms primary and secondary be used for characterizing stimuli so as to avoid confusion with the degree of processing nonlinearity.[6] Furthermore, it has been suggested that even more complex processing mechanisms, which use saliency maps as input for motion detectors,[7] account for the extraction of motion from some types of stimuli, giving rise to the notion of third-order motion.[8] So the current debate is driven by both the question of whether a single processing system[9] can account for the whole range of phenomena in theoretical terms and the question of how many principally different phenomena should be distinguished.

In the case of motion-defined motion, the preprocessing that is necessary to extract contours from the retinal signals to produce input to motion-detecting mechanisms is a comparison across space and time, that is, an analysis of spatiotemporal correlation. Two variants on this class of motion stimuli, labeled theta and wavy motion, have been implemented in random-dot kinematograms (RDKs): (a) Theta-motion stimuli consist of distinct objects that can be seen moving across a display only because their outline is discriminated from a dynamic-noise background on the basis of the distribution of local motion signals. The crucial feature of these stimuli is that the (random-dot) texture on the object surface, which has exactly the same spatial and temporal structure as the background, does not move in the same direction as the object itself but in an orthogonal,[2] or opposite, direction.[10] Motion on the object surface therefore cannot be used to identify the direction of object motion. (b) Wavy-motion stimuli consist of motion-defined gratings, for example, RDKs constructed from alternating columns of upward and downward dot motion. When the regions of vertical dot motion in opposite directions are shifted horizontally, human observers can perceive horizontal motion despite the fact that all the dots are moving vertically.[7,11]

A question about the ecological context in which motion-defined motion would be experienced has been raised. Because it obviously would be difficult to find such stimuli outside the psychophysical laboratory, they should be regarded as artificial stimuli for studying the ability of humans to process complex motion distributions. The theoretical significance of motion-defined motion lies in the recursive nature of these stimuli, which suggests distinct ways of processing beyond the assumption of simple input nonlinearities. The fundamental properties of both theta and wavy motion can be accounted for by a two-layer model consisting of a network of motion detectors the output of which is fed into a second network of motion detectors (at a slightly lower spatial resolution), thus extracting the movement of the motion signal.[11,12] This model is consistent with a number of psychophysical and physiological properties of human motion vision,[13–15] although so far no specific locations have been specified in the cortex that can be attributed to the consecutive processing stages.[16–18] In the following, we will use the two-layer network of motion detectors as a parsimonious system that coherently explains the perception of motion-defined motion without appealing to higher-level processes such as attention. This approach offers a simple computational model for simulating all necessary operations and predicts results in a quantitative fashion. Against this background, it will be argued that because of its recursive nature, motion-defined motion can play a special role in the ongoing dispute about possible mechanisms for processing second-order stimuli in the human visual system.

It should be obvious from the phenomenological definition of first- and second-order stimuli, and from the cursory review of some aspects of the literature presented above, that there is no simple answer to the question of whether the processing of second-order information requires mechanisms that are additional to and independent of those dealing with first-order information or whether a single processing mechanism is sufficient to account for both phenomena (e.g., Refs. 9 and 19–24). Although it is clear that second-order motion cannot be detected by a system with perfectly linear preprocessing, it is also evident that systems with nonlinear preprocessing can be devised that would detect not only second-order motion but also first-order motion.[4] It would therefore be easy to construct a single system to extract the motion of contrast-defined and luminance-defined patterns, thereby raising the question of whether an independent system with perfectly linear preprocessing is implemented in motion vision. However, the detection of motion-defined gratings cannot be explained by an implicit input nonlinearity but explicitly requires motion detectors for preprocessing. In the two-layer motion detector network that accounts for this aspect of human perception, first- and second-order motion would be represented at the consecutive processing stages[12]; the simple question as to whether a single system or two separate systems are required to process first- and second-order information would find a simple answer: The two kinds of information are represented at two consecutive stages of a processing sequence.

To gain further insight into the processing structure of motion information, we want to look at another aspect, which goes beyond the basic segmentation problem[25]: the integration of first- and second-order motion information, or the interaction between the stages of the two-layer model. To this end, we exploit the crucial peculiarity of motion-defined motion: The perceptual quality that defines an object (motion of the texture on the object surface) and that is attributed to the object as an emergent feature (motion of the object itself) are the same. The interaction of first- and second-order information has previously been analyzed in some detail in a direction-discrimination task that assessed sensitivity to theta motion and revealed cooperativity over a limited range of directions.[26] If these two types of motion information are integrated in the human visual system, one would expect that this should be reflected in the most fundamental property of a moving object, i.e., the direction in which it is seen to move. We wanted to investigate whether the perceived direction of a moving object is determined by a combination of first-and second-order motion information.

In the study presented here, we examined how motion information is integrated across the different layers by asking observers to estimate the absolute direction of object motion. For this purpose we compared direction sensitivity in a simple psychophysical paradigm that minimized position cues for four different classes of object motion:

1. Luminance-defined motion: A bright object is moving in front of a dark background; this stimulus was used as a reference, being the most fundamental apparent-motion stimulus that can be seen by everyone without further explanation (because the object is visible even in the single, still frame).
2. Fourier motion: In an RDK a group of dots is displaced coherently in front of a dynamic-noise background (classical RDK, as in Ref. 27); this stimulus has the same spatial frequency composition as 2. and 3.
3. Drift-balanced motion: The RDK consists of a dynamic-noise background in which a region is embedded

where all dots are static; this region itself is moving (comparable to the stimuli used in Refs. 28 and 29).

4. Theta motion: In an RDK a region of moving dots is moving in front of a dynamic-noise background; the dots within this region are moving not in the same direction as the object but with the same speed at a fixed angle relative to the object (similar to the stimuli used in Ref. 26).

It will be shown that not only is the accuracy of direction estimation affected by the particular class of motion stimulus but that we also misperceive the direction of object motion when the texture on the object is not moving along with the object.

2. METHODS

A. Subjects and General Procedures

Six observers (aged between 30 and 43 years; three male and three female) with normal or corrected-to-normal eyesight participated in our experiments. Two of them were the authors of this paper, but they did not show any systematic difference from the other four, completely naïve, observers. The experiments presented here are the result of an interhemispheric collaboration: Two subjects were tested in Adelaide and four in London on two identical sets of the experimental apparatus. In a dimly lit room, subjects were seated in front of a monitor at a viewing distance of 0.75 m with the head stabilized by a chin rest. Inexperienced participants familiarized themselves with the experimental setup and the task by running through as many practice trials as they felt appropriate (approximately 20 trials were sufficient). Subjects could then run the experiment by themselves in a self-paced manner guided by short instructions appearing on the stimulus monitor. A new trial was triggered by entering the response to the previous stimulus. Subjects needed about one hour to complete all four experiments.

B. Apparatus

Stimuli were generated on a Cambridge Research Systems Visual Stimulus Generator (VSG 2/3) and were displayed on a 20-in., high-resolution monitor (EIZO T662), with a display area of 800×600 pixels. The stimulus sequence for each trial, consisting of 32 frames, was prepared in the memory of the graphics board before each presentation and was copied to the display screen at a frame rate of 80 Hz during the trial. A square-shaped stimulus field of 256×256 pixels was positioned in the center of the screen. At a viewing distance of 0.75 m the stimulus field subtended 8.8 deg of visual angle in each dimension. The luminance of the entire screen background was set to 22 cd/m^2. For luminance-defined motion, the dark stimulus field was set to 5 cd/ m^2, and the bright object was set to 53 cd/m^2. For the other classes of motion, the individual dots in the RDKs were 1×1 pixel in size (2.1 arc min in each dimension), and they were set randomly with equal probability to be dark (5 cd/m^2) or bright (53 cd/m^2) dots, leading to a contrast of 75%.

For all classes of motion stimuli, square-shaped objects of 32×32 pixels size (1.1 deg in each dimension) were displaced by 2 pixels between consecutive frames, leading to a speed of 5.5 deg/s. Because the object could move in any direction, the position of the object (i.e., the memory location at which the object region was plotted) was calculated for each frame, and intermediate locations were rounded to the nearest integer. This procedure was adequate to create motion stimuli that appeared as perfectly smooth to our observers. Motion directions were chosen from eight sectors in random order by assigning a displacement angle that was combined from an integer multiple of 45 deg and a random number between $+15$ and -15 deg (directional jitter). Thus the data could be analyzed in categories corresponding to cardinal or diagonal directions, but the stimuli had no visible preference for any particular motion direction. So that subjects were prevented from judging motion direction from the location at which the object disappeared at the end of a stimulus, the trajectory did not cross the center of the stimulus field. Rather, the start position was randomized with an independent positional jitter of 64 pixels in both vertical and horizontal directions. This positional uncertainty, being of the same order as the complete trajectory, was large enough to make subjects focus on motion direction instead of object position.

C. Stimuli

Before each motion stimulus sequence an empty background frame (5 cd/m^2) was presented for 1 s. The actual motion sequence (32 frames displayed in 400 ms) was preceded by a short beep and followed by another 1-s presentation of the empty background frame. The subject was then cued to report the perceived direction of motion (see below) by another short beep. The four classes of motion stimuli were tested in blocks of trials, always in the following order:

1. Luminance-defined motion: Each stimulus frame was generated by copying the bright object region into the empty background field, as described above, leading to a bright square moving in front of a dark background (see inset in Fig. 1 below). A random direction in each of the 8 sectors was tested in a single block, and subjects completed 80 trials in 10 successive blocks.

2. Fourier motion: For each stimulus frame the background was generated by copying a 256×256 pixel region from a random position in a large random-dot pattern. This procedure leads to random changes of luminance for all background pixels. A region, randomly chosen (once for each trial) from another large random-dot pattern, was copied into the background field for each of the frames, with the object position determined as described above. In consequence, a square region of coherently moving dots is traveling in front of a dynamic noise background (see inset in Fig. 2 below). Subjects completed 80 trials in 10 successive blocks.

3. Drift-balanced motion: For each stimulus frame the dynamic-noise background was generated as described in 2. from a large random-dot pattern. A region, randomly chosen (once for each trial) from another large random-dot pattern, was copied into the background field for each of the frames. The copy and the paste position of this region were incremented identically to shift this region across the stimulus field. This leads to a region of

static dots moving within a field of dynamic noise [see inset in Fig. (3)]. Subjects were tested with 80 trials in 10 successive blocks.

4. Theta motion: For each stimulus frame the dynamic-noise background was generated as described in 2. from a large random-dot pattern. A region randomly chosen (once for each trial) from another large random-dot pattern was copied into the background field for each of the frames. In this case, the copy base position and the paste position were incremented as in all other cases (see above), but the copy position was incremented by the same amount as the paste position plus an additional displacement determined by the motion of the dots on the object surface relative to the object itself (for more detailed methods, see Ref. 12). The motion of surface dots was either +45 or −45 deg relative to the direction of object motion and was always at the same speed as the object. As a result, an object region is moving within a dynamic-noise background and the dots within this region are moving diagonally relative to the object itself [see inset in Fig. (4)]. A random direction in each of the 8 sectors was tested twice in a single block in random order, one with each relative motion direction of the surface dots, and subjects were tested with 160 trials in 10 successive blocks.

D. Task

After each trial, a short beep sounded and then a gray ring (inner diameter 280 pixels, outer diameter 360 pixels) with a small bright disk (8 pixels in diameter) in its center appeared on the screen along with a green cursor cross that could be moved around by means of the mouse. The subjects were asked to position the cursor on the circle at the location that corresponded to the direction of the object motion from the central bright spot. Before the experiment, the subjects were given clear instructions to report the direction of the object motion, no matter how these objects might appear to them, without considering what might happen on the object surface. They were also reminded to keep their gaze still on the center of the screen and not to follow the movement of the object. (No fixation target was presented in the main experiments because we did not want to provide a static reference cue in the region of object motion.) Furthermore, it was explicitly pointed out to the subjects that they should not indicate where the object had disappeared nor try to use position cues but to report, in this polar coordinate system, the motion direction of an object that could move at variable locations within the stimulus field. Once they felt comfortable with the direction setting, the subjects pressed a mouse button, and the cursor turned red for a short period before the next trial was initiated.

3. RESULTS AND DISCUSSION

The first experiment used luminance-defined objects to provide a reference data set that described direction sensitivity for the simplest possible apparent-motion stimuli. The systematic error and the accuracy achieved for this stimulus reflect the performance limitations determined by the experimental conditions used here.

The results of this experiment are illustrated in Fig. 1. A general impression of the response accuracy is given by the scatter diagram [Fig. 1(a)], in which the motion direction reported by a typical subject (MBH) for each of the 80 trials is plotted as a function of the actual direction of ob-

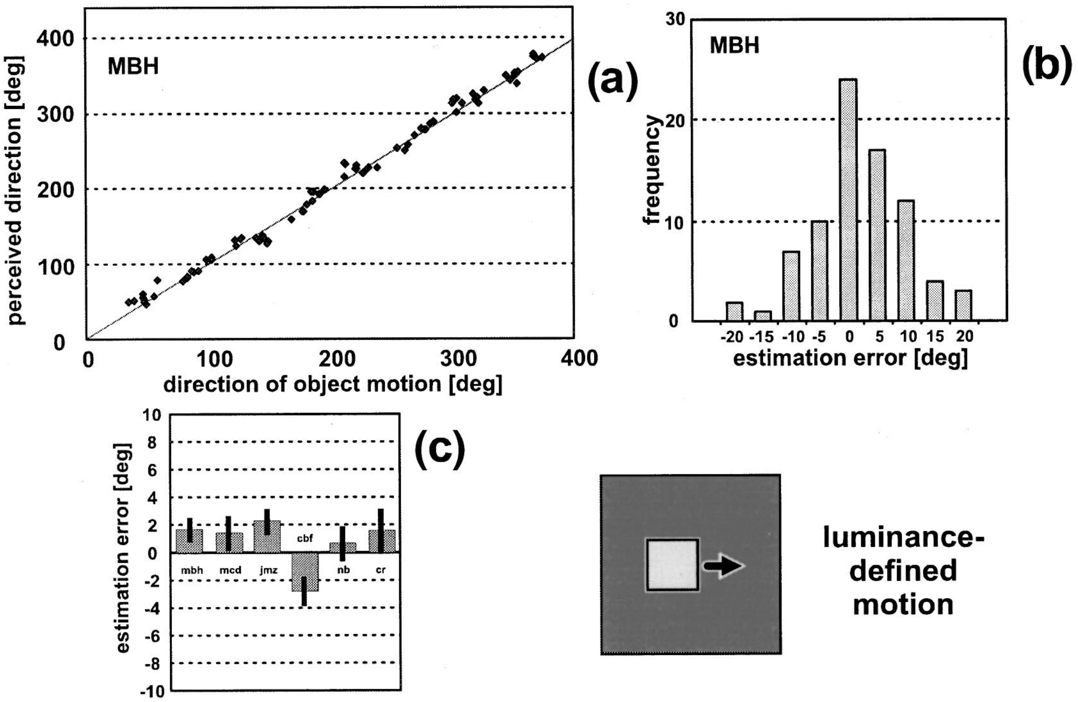

Fig. 1. Experimental results for luminance-defined objects (sketched as inset). (a) Scatter diagram of perceived motion direction (ordinate) plotted for each of 80 trials as function of the actual direction of object motion (abscissa) for subject MBH; the diagonal line indicates the veridical direction. (b) Histogram of direction-estimation errors (the difference between perceived and veridical motion direction) for subject MBH. (c) Means and SEMs ($n = 80$) of the estimation errors for each of six subjects (indicated by initials).

ject motion. The data points are rather close to the diagonal line that indicates the veridical direction. The scatter diagram shows a trend that the estimation errors, i.e., the deviations from the veridical direction, tend to be smaller for stimuli around 90, 180, 270, and 360 deg, that is along the cardinal directions, as compared with stimuli along oblique directions, i.e., 45, 135, 225, and 315 deg. This behavior resembles the oblique effect described for direction discrimination with use of other motion stimuli.[30] However, the current result could also be partly due to our stimulus configuration, since the square-shaped stimulus field and object provided strong reference signals along the cardinal directions. Because such oblique effects were not our primary interest and because we are using the direction sensitivity for luminance-defined objects as a reference point for the other classes of motion stimuli that were tested under exactly the same conditions, we did not analyze this aspect of the data further and instead pooled data from all sectors of the stimulus domain.

In the next step of the analysis we determined the distribution of direction-estimation errors, that is, the difference between perceived and actual motion direction. The resulting histograms [shown for subject MBH in Fig. 1(b)] show moderately wide (range larger than 30 deg) and approximately normal distributions. Thus we calculated the means and variances for the estimation errors, to measure any systematic deviation of perceived direction and the accuracy of the perceptual decision as reflected by the width of the histogram, respectively. The means for each of our six subjects and their standard errors (SEMs) are displayed in Fig. 1(c) for luminance-defined objects. All individual misjudgments are rather small, and the variation between subjects is small as well; there is no indication of any systematic trend in the estimation errors. The average of the direction estimation error for our six subjects was 0.76 deg, and the average standard deviation (reflecting the accuracy of the direction judgment) was 10.72 deg. As another performance measure we calculated the means of the unsigned estimation errors (so that errors of opposite sign do not cancel each other), leading to an average value of 8.19 deg for our six subjects (6.87 deg standard deviation). These results suggest that there was no strong systematic estimation error and reasonable reliability for the luminance-defined stimuli.

The corresponding set of direction sensitivity data for Fourier motion stimuli is shown in Fig. 2. The scatter diagram [Fig. 2(a)] and the histogram of estimation errors, that is, the differences between perceived and actual motion direction [Fig. 2(b)], for the typical subject MBH are very similar to those for luminance-defined objects. This impression is confirmed by the means and SEMs of the estimation errors for all six subjects shown in [Fig. 2(c)]. The average mean and standard deviation of the estimation error for all six subjects were -0.01 deg and 9.94 deg, respectively. The average of the mean unsigned estimation errors was 7.73 deg for our six subjects (7.67 deg standard deviation). Comparing Fig. 2(c) with Fig. 1(c) indicates that there is no consistent pattern of deviations for either of the two stimulus classes, or between them, but there seems to be a trend that the systematic deviations from zero estimation error and the size of the error bars are smaller than for luminance-defined motion. However, this trend is not confirmed by statistical tests; the differences between the average standard deviations of the estimation errors (10.72 versus 9.94 deg) and between the average unsigned estimation errors (8.19 versus 7.73 deg) were not significant (we computed the test statistic for comparing the variances of two related samples suggested by Pitman[31] and the usual paired *t*-test for comparing means). These results suggest that there was no systematic difference in estimation error for Fourier motion direction as compared with luminance-defined motion. Furthermore, accuracy is at least as good as that for luminance-defined motion-direction estimation. This outcome is not surprising when one takes into consideration that the dots moving together with the object provide a rich source of local motion signals, which may be at least as efficient as the low-spatial-frequency signals arising from the luminance-defined object. In the present context, the crucial finding is that our Fourier motion stimulus can be regarded as close to optimal (i.e., leading to a performance like that for the most salient motion stimulus) and will thus be used as the reference stimulus for this study.

When the RDK contains a flicker-defined object, as a simple case of a second-order motion stimulus, the situation changes quantitatively. The set of direction-sensitivity data for drift-balanced motion is shown in Fig. 3 in the same format as the corresponding data for first-order stimuli. The scatter diagram [Fig. 3(a)] for the typical subject MBH suggests larger deviations of the perceived motion directions from the veridical ones than for either luminance-defined or Fourier objects and generated a correspondingly broader histogram of estimation errors [Fig. 3(b)]. The means and SEMs of the estimation errors for all six subjects are plotted in Fig. 3(c) and show that, as with the other two stimulus classes, there was no consistent pattern of biased direction misjudgments. The average mean and standard deviation of the estimation error for all six subjects were 1.31 and 15.74 deg, respectively. The difference in the standard deviations for drift-balanced and Fourier motion stimuli, 15.74 and 9.94, respectively, while not statistically significant at the $p = 0.05$ level, is clearly substantial ($F_{\max} = 2.51$), suggesting a difference between the distribution widths of estimation errors and hence in the accuracy of the responses. This result is confirmed by the average of the mean unsigned estimation errors which was 11.82 deg for our six subjects (11.99 deg standard deviation); this is significantly different from the value of 7.73 deg for Fourier motion [$t(5) = 3.97$, $p < 0.01$, paired *t*-test, two tailed]. These results confirm the initial impression, gained from looking at the data for subject MBH, that accuracy is impaired when the texture on the object surface is not moving together with the object. This outcome is consistent with the general view that the extraction of the second-order object requires additional processing that would therefore increase the internal level of uncertainty about the current object location. It is also consistent with independent experimental data: Coherence thresholds for direction discrimination tend to be higher for second-order motion than for comparable first-order motion stimuli.[26,32]

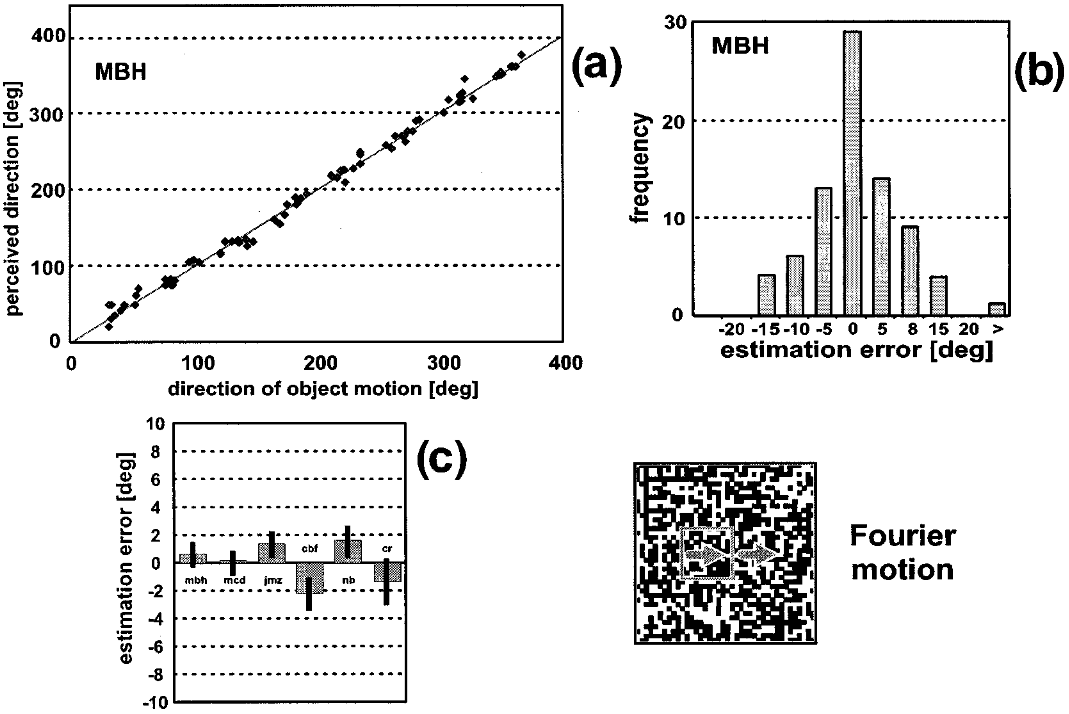

Fig. 2. Experimental results for Fourier motion objects (sketched as inset). Other particulars are the same as for Fig. 1.

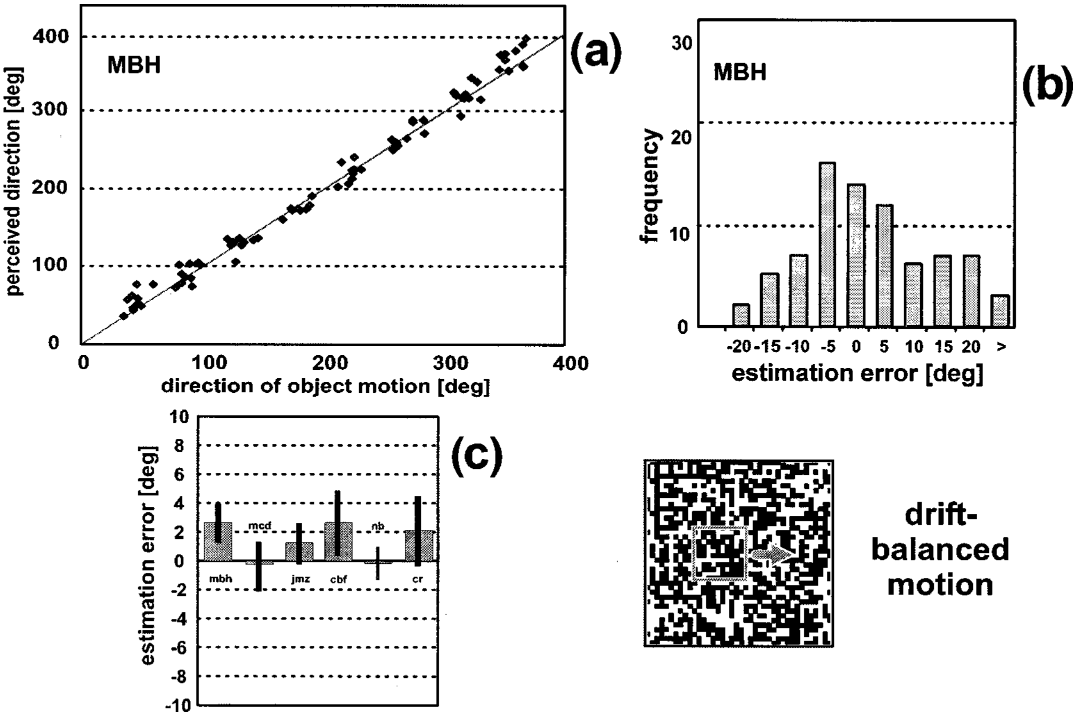

Fig. 3. Experimental results for drift-balanced objects (sketched as inset). Other particulars are the same as for Fig. 1.

In the case of theta motion we have to compare two sets of data: For half of the trials the dots on the object surface were moving diagonally leftward (+45 deg) relative to the direction of object motion, and for the other half they were moving diagonally rightward (−45 deg) relative to the object trajectory. Accordingly, to present this set of direction-sensitivity data, we changed the format of Fig. 4 from that of the previous figures. The scatter diagram [Fig. 4(a)] for the typical subject MBH immediately demonstrates that the responses fall into two populations, indicated by two symbols. The perceived directions for the +45 deg stimuli scatter above the diagonal of veridical direction estimates and those for the −45 deg stimuli scatter below the diagonal. As a consequence, the two histograms of estimation errors for the same subject MBH [Fig. 4(b)]; please note the change of abscissa scale compared with previous figures) clearly separate into two distinct distributions. The +45 deg stimulus histogram peaks in the +30 and +40 deg bins, and the −45 deg stimulus histogram peaks in the −30 deg bin; and there is little overlap of the distributions. Apart from these systematic deviations from zero of the estimation error, the widths of the two histograms, as well as the spread of data in the scatter diagram, suggest a further increase of the uncertainty in this subject's responses compared with all previous stimulus classes. The means and SEMs of the estimation errors for the two conditions for all six subjects are plotted in Fig. 4(c). The figure shows a clear and consistent pattern of deviations from zero in the direction of motion of the object surface texture. The average mean and standard deviation of the estimation error for all six subjects are 32.1 and 17.61 deg for the leftward (+45 deg) texture motion and 32.6 and 17.76 deg for the rightward (−45 deg) texture motion, respectively. Whereas the average mean estimation errors and the average unsigned means (which are influenced by the systematic error and thus no longer can be used as measure of accuracy) are significantly different from those of all other stimulus conditions ($p < 0.01$ in all cases, paired t-test, two tailed), differences in average standard deviations were not significant for the comparison with drift-balanced stimuli (+45 deg theta: 17.61 versus 15.74, $F_{\max} = 1.24$; −45 deg theta: 17.76 versus 15.74, $F_{\max} = 1.27$) or for the comparison with Fourier stimuli (+45 deg theta: 17.61 versus 9.94, $F_{\max} = 3.14$; −45 deg theta: 17.76 versus 9.94, $F_{\max} = 3.19$); for the latter comparisons, however, the differences in standard deviations are clearly substantial. This pattern of results confirms the initial impression from the typical subject that observers generate substantial systematic estimation errors but does not support the view that accuracy deteriorates further when the texture on the object surface is moving in a direction independent from that of the object itself.

The most interesting result of this experiment, however, is the systematic error produced by all our observers, who reported a direction of object motion that was intermediate between the veridical direction of the object and the direction in which the dots within the object region moved. It is important to note that this is not a problem with the instructions given to the participants, because through the succession of motion stimuli tested they were trained to focus on the direction of object motion. Furthermore, within the short inspection period they did not even notice that for the last experimental condition the surface texture was moving in a direction different from that of the object. In this context it is also noteworthy that the two authors, the only subjects who knew about the stimulus design, could hardly describe the surface texture motion in an individual trial and certainly did not produce the largest effects. Since naïve observers, who did not even know about the dot motion, were clearly most strongly affected by it, the systematic deviation from veridical object direction cannot be attributed to expectation of such an effect. Consequently, it is clear that the direction perceived as object motion is determined by both the direction of the moving texture and of the object itself. This means that there is a strong

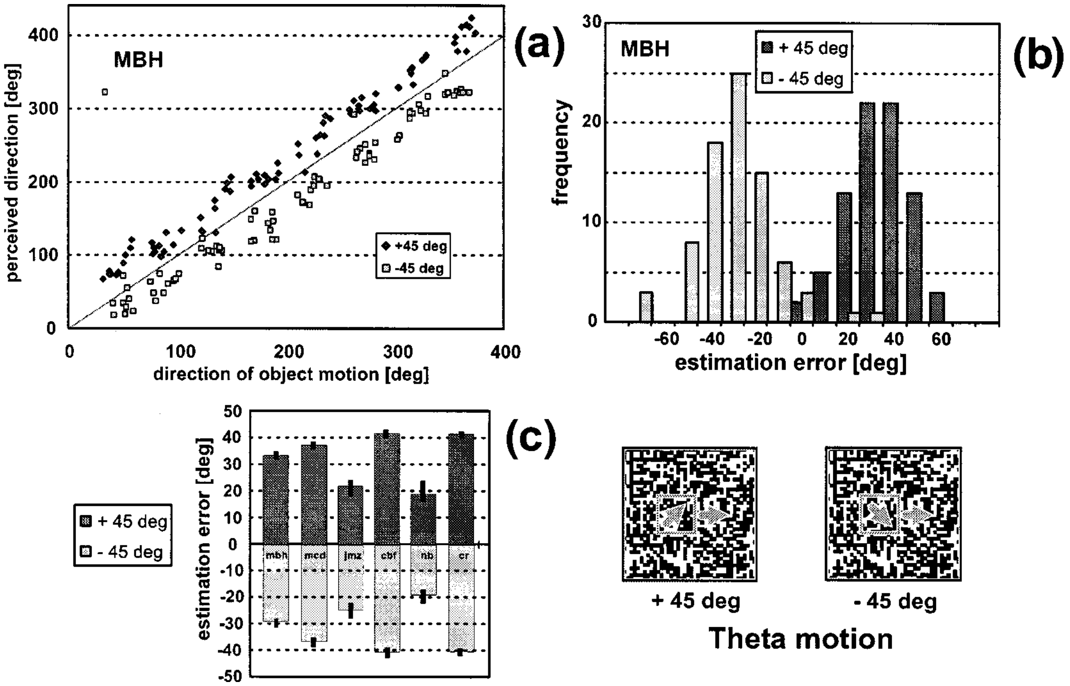

Fig. 4. Experimental results for theta motion objects (sketched as inset). (a) Scatter diagram of perceived motion direction (ordinate) for leftward (+45 deg, diamonds) and rightward (−45 deg, squares) relative texture motion, respectively, plotted as function of the actual direction of object motion (abscissa); the diagonal line indicates the veridical direction. (b) Histograms of direction-estimation errors (the difference between perceived and veridical motion direction) for the two conditions (indicated by different shading). (c) Means and SEMs ($n = 80$) of the estimation errors for each of six subjects for the two conditions (indicated by different shading).

Table 1. Mean Estimation Errors for Fourier and Theta (+45 and −45 deg Texture Motion) Object Motion for Four Subjects for Three Stimulus Configurations

	Standard[a]			Fixation[b]			Short[c]		
		Theta			Theta			Theta	
Subject	Fourier	+45 deg	−45 deg	Fourier	+45 deg	−45 deg	Fourier	+45 deg	−45 deg
MBH	0.64	33.24	−30.16	2.78	39.15	−37.34	1.35	44.87	−42.30
JMZ	1.28	21.17	−25.40	0.34	20.43	−22.97	1.97	35.35	−35.80
NB	1.62	18.90	−19.96	2.63	27.30	−22.45	2.17	44.27	−41.13
CR	−1.40	41.10	−41.08	−0.06	41.12	−41.77	2.54	43.57	−40.95

[a] No fixation target, 400-ms duration, data as in Figs. 2 and 4.

[b] Central fixation target, 400-ms duration.

[c] No fixation target, 200-ms duration.

misjudgment of object motion direction that suggests that the two layers of motion processing interact in direction estimation.

However, with a stimulus duration of 400 ms, and in the absence of a fixation target (which we did not use in our main experiments because it would have provided a clear position reference), it must be realized that it is rather difficult for the observers not to move their eyes.[33] Therefore, because the possibility cannot be excluded that our results could have been affected by involuntary eye movements, we performed two control experiments with four of our subjects: (1) We provided a fixation target (a white spot with a diameter of 6 pixels) at the center of the stimulus display, which should help the subjects to keep their gaze stable at the center of the screen. Although this fixation target provides an additional positional cue in the stimulus, which might have been used to improve the judgment of motion direction, there is no effect. As can be seen from Table 1, perceived directions were virtually identical with and without the fixation target. This clearly speaks against an explanation in terms of involuntary eye movements. (2) In an attempt to minimize the possible elicitation of any eye movements without changing the spatial layout of our stimulus by introducing a fixation target, we reduced stimulus duration to 200 ms, which is too short a time to allow for any substantial eye movements.[34] As can be seen in Table 1, this also did not reduce the misjudgment of object direction but seems to increase even further the systematic deviations from veridical object-motion direction. Again, one has to conclude that estimation errors cannot be ascribed to eye movements. This observation, however, points to the possibility of temporal changes in the interaction between the first- and second-order motion information. It should be noted that the dominance of first-order motion during the initial phase of the stimulus (as demonstrated by our second control experiment) is consistent both with reports that second-order motion needs longer to be reliably detected than does first-order motion and with the hierarchical processing models outlined in Section 1. In this context, it is also of interest that longer trajectories can lead to the perception of ambiguous, curved paths and paradoxical separations of perceived motion direction and object location. At present it is not known whether and to what extent eye movements contribute to this effect, and it is clear that this issue needs further investigation. The phenomenon as such, however, only supports the notion that there is a substantial misjudgment of object motion direction that makes the object appear in unexpected locations.

4. CONCLUSIONS

Interactions between first- and second-order motion signals have been described for the control of eye movements.[35,36] In the present study we investigated directional sensitivity, the main finding being that when subjects are asked to estimate object motion direction, the motion signals on the surface of a moving object interact with the motion signals from the object itself. Whereas we found only a small improvement of direction sensitivity, if any, when the dots on the object surface moved coherently with the object, accuracy deteriorated when surface dots did not move at all. Most important, we found systematic misjudgments of object motion in the direction of dot motion when objects and surface dots moved in different directions. This clearly indicates that the motion signals within the object region do interact with the information about object displacement (which is derived from these signals), at least for our experimental conditions. The lack of independence between these two signal groups is also demonstrated by the phenomenal coherence of a unitary (though strange) object moving during the short stimulus speeds used here. Taking together direction and speed misjudgments for motion-defined objects, it is apparent that the first- and second-order velocity vectors are somehow pooled together.

Position misjudgments have been observed for motion-defined targets in a number of experimental conditions.[37–40] Such perceptual errors can be regarded as a phenomenon related to the shifts in directional tuning for motion-defined objects reported here, because they indicate a rather similar interaction between the perceived location of an object and the motion on its surface. The misjudgment of object motion direction can be seen as a logical extension of position misjudgment, with systematic errors in the internal representation of the instantaneous object location automatically leading to changes in perceived direction. However, our results are not easily explained just as a position effect, because we minimized absolute position cues in our experimental design. Moreover, a simple comparison of object positions would be difficult during the short stimulus period used here, because these positions did not give the observer any clear and unambiguous indication as to where the object was moving. Therefore the most immediate way to discuss our observations is as an inability to separate first- and second-order directional information in the context of the two-layer model of motion processing. This model, it should be noted, also accounts for the perception of motion-defined motion as such. In doing this, we are not denying that there are alternative models that can deal adequately with these stimuli, such as the third-order system proposed by Lu and Sperling.[7] The crucial point here is that two pieces of information, namely, the motion of the texture on the object surface and the motion of the object itself, interact in judgments of motion direction.

The two-layer motion-detector model as originally proposed[12] is sketched in Fig. 5(a). The output of a fine-grain array of elementary motion detectors (EMDs) is spatially integrated and then used as input for a second, coarse-grain array of EMDs that detects the movement of motion-defined contours. It should be noted that the models sketched in Fig. 5 are one-dimensional simplifications and that motion signals in a wide range of directions, not only along the same axis, can be used to detect theta motion. Whereas the range of possible texture motion directions has been analyzed in detail for direction discrimination,[26] the experiments presented here indicate the effects on direction sensitivity only for oblique motion of the texture relative to the motion-defined object, and the full input tuning curve requires further experimental study. With the only difference being in spatial resolution, the same (correlation-type) mechanism is used for the two processing stages in the two-layer model. This suggestion is strongly supported by the specific tuning of the human visual system to the temporal frequency of motion-defined gratings.[11] In order to account for the major result of our present experiments, that the perceived direction of a motion-defined object is determined by an interaction of primary and secondary motion information, one would need to assume that the output signals from both layers are somehow pooled to estimate the two-dimensional direction of object motion.

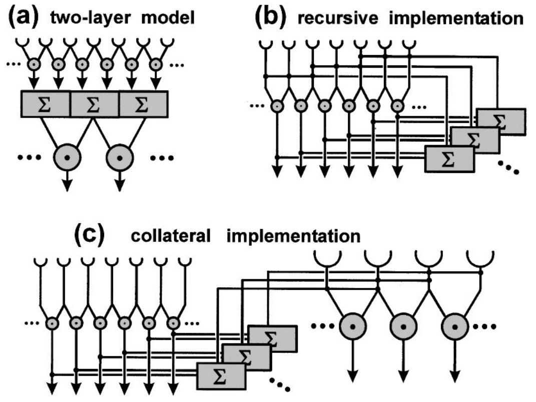

Fig. 5. Models for the detection of theta motion. (a) Two-layer motion-detector model as originally proposed: The output of a fine-grain array of EMDs is spatially integrated and then used as input for a second, coarse-grain array of EMDs. (b) Recursive implementation: The output of the motion-detector network is fed back after appropriate spatial pooling into the same network. (c) Network of collateral connections: The output of a primary, fine-grain network of EMDs is processed after spatial integration in a secondary, coarse-grain network of EMDs.

It has been mentioned in a previous publication[11] that a variant of the two-layer network could be implemented in terms of a recursive architecture. In such a model, the output of the motion-detector network, after appropriate spatial pooling, is fed back into the same network [one-dimensional sketch in Fig. 5(b)]. Such a system would have some nontrivial properties as far as dynamics and stability are concerned, and therefore it was not pursued further. The fundamental ability to extract motion-defined motion, however, would be conserved through the recursive structure of this network, which allows differences in local motion signal strength (in various directions) to be used as the input signal for motion detectors. An interesting property emerging from the recursive implementation is that the two-dimensional direction of motion for motion-defined objects would be intermediate between the primary (texture motion) signal and the secondary signal (extracted from feedback), because both types of information are represented by the same array of EMDs. Although these theoretical considerations result in a clear expectation of an intermediate direction estimate, further experimental and computational analyses are required to assess which factors determine the exact balance between texture and object motion and to determine the stability of such a model variant.

The clear difference between the spatial resolution for primary and secondary motion information[11] hints at another variation on the same theme of sequential processing in motion-detector networks, but now with the use of collateral connectivity in a massively parallel architecture. The analysis of the visual input in parallel channels tuned to different spatial frequencies is a fundamental property of the human visual system that is also realized in sets of motion detectors with different receptive field dimensions.[41,42] In Fig. 5(c), two sets of EMDs with different receptive field sizes are sketched schematically to show how a set lateral connections between such channels can generate all the features necessary to account for our experimental results. Through the network of collateral connections, the output of a primary, fine-grain network of EMDs is processed after spatial integration in a secondary, coarse-grain network of EMDs, which is thus able to extract motion-defined motion. At the same time, the pooling of motion information across the set of spatial channels will immediately generate intermediate two-dimensional motion directions, as reported in our present study. It should be noted that this model structure can in some sense be seen as related to a two-stage model suggested for the perception of moving plaids.[20] This model also assumes the integration of two channels tuned to motion at different spatial frequencies. The differences between the two model structures are (a) that our proposal does not include an input nonlinearity in the coarse channel to detect contrast modulations (which are not considered here) and (b) that we add a lateral connection between the two channels. This lateral connection (see also the model sketch in Ref. 11), which is necessary for detecting motion-defined motion, is the crucial addition that distinguishes our proposal from the model of Wilson *et al.*[20] and generates the additional features that explain the directional misjudgments reported here.

In summary, in the present paper we demonstrate a strong influence of surface-texture motion on the perceived direction of motion-defined objects that produces systematic and substantial misjudgments. Eventually, these misjudgments can even lead to the effect that a moving object may appear perceptually at an unexpected location. Generating such illusions, this class of stimuli and the experimental and computational analyses of directional sensitivity for motion-defined objects obviously have great potential for the study of the motion-processing architecture in the human visual system.

The authors' e-mail addresses are as follows: J.Zanker@rhul.ac.uk and nicholas.burns@adelaide.edu.au.

REFERENCES

1. J. M. Zanker and J. Zeil, *Motion Vision: Computational, Neural and Ecological Constraints* (Springer, Berlin, 2001).
2. P. Cavanagh and G. Mather, "Motion: the long and short of it," Spatial Vision **4**, 103–129 (1989).
3. G. Sperling, "Three stages and two systems of visual processing," Spatial Vision **4**, 183–207 (1989).
4. C. Chubb and G. Sperling, "Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception," J. Opt. Soc. Am. A **5**, 1986–2006 (1988).
5. T. Poggio and W. Reichardt, "On the representation of multi-input systems: computational properties of polynomial algorithms," Biol. Cybern. **37**, 167–186 (1980).
6. J. M. Zanker, "Of models and men: mechanisms of human motion perception," in *Early Vision and Beyond*, T. V. Papathomas, C. Chubb, A. Gorea, and E. Kowler, eds. (MIT Press, Cambridge, Mass., 1995), pp. 156–165.
7. Z.-L. Lu and G. Sperling, "Attention-generation apparent motion," Nature **377**, 237–239 (1995).
8. G. Sperling and Z.-L. Lu, "A systems analysis of visual motion perception," in *High-Level Motion Processing*, T. Watanabe, ed. (MIT Press, Cambridge, Mass., 1998), pp. 153–183.
9. A. Johnston, P. W. McOwan, and H. Buxton, "A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells," Proc. R. Soc. London Ser. B **250**, 297–306 (1992).
10. J. M. Zanker, "Theta motion: a new psychophysical paradigm indicating two levels of visual motion perception," Naturwissenschaften **77**, 243–246 (1990).
11. J. M. Zanker, "On the elementary mechanism underlying secondary motion processing," Proc. R. Soc. London Ser. B **351**, 1725–1736 (1996).
12. J. M. Zanker, "Theta motion: a paradoxical stimulus to explore higher order motion extraction," Vision Res. **33**, 553–569 (1993).
13. D. R. Patzwahl, J. M. Zanker, and E. Altenmüller, "Cortical potentials reflecting motion processing in humans," Visual Neurosci. **11**, 1135–1147 (1994).
14. D. Braun, D. Petersen, P. Schönle, and M. Fahle, "Deficits and recovery of first-order and second-order motion perception in patients with unilateral cortical lesions," Eur. J. Neurosci. **10**, 2117–2128 (1998).
15. D. R. Patzwahl and J. M. Zanker, "Mechanisms for human motion perception: combining evidence from evoked potentials, behavioural performance and computational modelling," Eur. J. Neurosci. **12**, 273–282 (2000).
16. O. J. Braddick, "Motion may be seen but not used," Curr. Biol. **2**, 587–589 (1992).
17. D. R. Patzwahl, J. M. Zanker, and E. Altenmüller, "Cortical potentials in humans reflecting the direction of object motion," NeuroReport **4**, 379–882 (1993).
18. D. R. Patzwahl, T. Elbert, J. M. Zanker, and E. Altenmüller, "The cortical representation of object motion in man is interindividually variable," NeuroReport **7**, 469–472 (1996).
19. A. M. Derrington and D. R. Badcock, "Separate detectors for simple and complex grating patterns?" Vision Res. **25**, 1869–1878 (1985).
20. H. R. Wilson, V. P. Ferrera, and C. Yo, "A psychophysically motivated model for two-dimensional motion perception," Visual Neurosci. **9**, 79–97 (1992).
21. Y.-X. Zhou and C. L. Baker, Jr., "A processing stream in mammalian visual cortex neurons for non-Fourier responses," Science **261**, 98–101 (1993).
22. J. C. Boulton and C. L. Baker, Jr., "Dependence on stimulus onset asynchrony in apparent motion: evidence for two mechanisms," Vision Res. **33**, 2013–2019 (1993).
23. A. T. Smith and T. Ledgeway, "Separate detection of moving luminance and contrast modulations: fact or artifact?," Vision Res. **37**, 45–62 (1997).
24. N. E. Scott-Samuel and M. A. Georgeson, "Does early nonlinearity account for second-order motion?" Vision Res. **39**, 2853–2865 (1999).
25. O. J. Braddick, "Moving on the surface," Curr. Biol. **4**, 534–536 (1994).
26. J. M. Zanker and I. S. Hüpgens, "Interaction between primary and secondary mechanisms in human motion perception," Vision Res. **34**, 1255–1266 (1994).
27. O. J. Braddick, "A short-range process in apparent motion," Vision Res. **14**, 519–527 (1974).
28. A. M. M. Lelkens and J. J. Koenderink, "Illusory motion in visual displays," Vision Res. **24**, 1083–1090 (1984).
29. C. Chubb and G. Sperling, "Two motion perception mechanisms revealed through distance-driven reversal of apparent motion," Proc. Natl. Acad. Sci. USA **86**, 2985–2989 (1989).
30. B. L. Gros, R. Blake, and E. Hiris, "Anisotropies in visual motion perception: a fresh look," J. Opt. Soc. Am. A **15**, 2003–2011 (1998).
31. E. G. J. Pitman, "A note on normal correlation," Biometrika **31**, 9–12 (1939).
32. J. M. Zanker, "Perceptual learning in primary and secondary motion vision," Vision Res. **39**, 1293–1304 (1999).
33. B. J. Murphy, E. Kowler, and R. M. Steinman, "Slow oculomotor control in the presence of moving backgrounds," Vision Res. **15**, 1263–1268 (1975).
34. P. E. Hallett, "Eye movements," in *Handbook of Perception and Human Performance I*. K. R. Boff, L. Kaufman, and J. P. Thomas, eds. (Wiley, New York, 1986), pp. 10-1–10-112.
35. L. R. Harris and A. T. Smith, "Interactions between first- and second-order motion revealed by optokinetic nystagmus," Exp. Brain Res. **130**, 67–72 (2000).
36. A. Lindner and U. J. Ilg, "Initiation of smooth-pursuit eye movements to first-order and second-order motion stimuli," Exp. Brain Res. **133**, 450–456 (2000).
37. R. L. DeValois and K. K. DeValois, "Vernier acuity with stationary moving Gabors," Vision Res. **31**, 1619–1626 (1991).
38. T. Banton and D. M. Levi, "Spatial localization of motion-defined and luminance-defined contours," Vision Res. **33**, 2225–2237 (1993).
39. A. Hayes, "Apparent position governs contour-element binding by the visual system," Proc. R. Soc. London B **267**, 1341–1345 (2000).
40. S. Nishida and A. Johnston, "Influence of motion signals on the perceived position of spatial pattern," Nature **397**, 610–612 (1999).
41. W. A. Van de Grind, J. J. Koenderink, and A. J. Van Doorn, "The distribution of human motion detector properties in the monocular visual field," Vision Res. **26**, 797–810 (1986).
42. S. J. Anderson and D. C. Burr, "Spatial summation properties of directionally selective mechanisms in human vision," J. Opt. Soc. Am. A **8**, 1330–1339 (1991).