PERGAMON

Vision Research 41 (2001) 705–710

Vision Research

www.elsevier.com/locate/visres

Role of motion integration in contour perception

Pawan Sinha *

E25-229, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 45 Carleton Street, Cambridge, MA 02142, USA

Received 4 July 2000; received in revised form 27 November 2000

Abstract

Contours are believed to play a key role in the visual analysis of scenes by the primate brain. In dynamic scenes, the presence of contours is often signaled by discontinuities in motion fields. However, it is unclear whether the motion fields over which the visual system extracts discontinuities, correspond to the local optic-flow or the pattern motion fields obtained by integrating local estimates. A resolution of this issue would provide important clues about the organization of visual motion and form analysis processes. In this paper, we present experimental evidence which suggests that the perception of motion defined contours is strongly dependent on motion integration — an operation that is believed to take place relatively late in the visual stream. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Contour perception; Motion integraiton; Optic-flow

---

1. Introduction

One of the key tasks of visual processes in the primate brain is the extraction of image contours. Since contours are often not well-defined along their extent, the visual system needs to be able to infer their presence from partial/indirect evidence, such as discontinuities in oriented elements (Kanizsa, 1976; von der Heydt, Peterhans and Baumgartner, 1984; Petry & Meyer, 1987; Grosof, Shapley & Hawken, 1993; Soriano, Spillman & Hawken, 1996). Such filling-in of missing contour information leads to percepts of subjective/illusory contours. Several studies suggest that, in static displays, local analysis of orientation in the vicinity of the discontinuities usually suffices to predict characteristics of the contour percept (von der Heydt & Peterhans, 1989; Dresp & Bonnet, 1991, 1993). Dynamic displays, however, present an interesting open question. Unlike local estimates of static element orientation, local estimates of motion trajectory orientation (or more simply, motion direction) are typically ambiguous due to the well-known aperture problem (Hildreth, 1984; Movshon, Adelson, Gizzi, & Newsome, 1985; Nakayama & Silverman, 1988). Motion integration processes are needed to determine true pattern motion direction. It is unclear whether the visual system determines the presence of contours based on motion discontinuities extracted over the locally estimated optic-flow or, rather, the pattern motion field.

This issue is illustrated in Fig. 1(a), which shows a few frames from a simple motion sequence. In this movie, a translating diamond slides underneath an occluder with invisible boundaries (by virtue of the occluder having the same brightness as the background). Observers report perceiving a vivid subjective contour at the interface of the diamond and the occluder. In trying to understand how this percept arises, we confront the question alluded to above. The motion of the diamond segments in the vicinity of the diamond–occluder interface is ambiguous due to the aperture problem (Fig. 1b). What we are interested in finding out is whether the brain determines the presence of an occluding contour before or after motion disambiguation. Resolving this issue is important for understanding the functional organization of visual form and motion perception processes.

Random-dot displays, as shown in Fig. 2, are typically used to study the perception of dynamically defined contours (Anstis, 1970; Julesz, 1971; Braddick, 1973, 1974; Nakayama & Tyler, 1981; Lamme, Van Dijk & Spekreijse, 1993). In such a display, no contour

---

* Tel.: + 1-617-2531434; fax: + 1-617-2532964.
E-mail address: sinha@ai.mit.edu (P. Sinha).

PII: S0042-6989(00)00306-0

is perceived when the two textured planes are at rest. However, as soon as they begin moving in different directions and/or with different speeds, a vivid contour becomes evident at their interface. Though they serve as striking demonstrations of motion-defined contours, such displays prove inadequate to address the issue of whether the contour percepts are based on discontinuities computed over local flow or pattern motion fields. This is because the unambiguously moving punctuate features in such displays render them immune to the aperture problem. Consequently, the locally computed flow is identical to the pattern motion estimate and there is no way to distinguish between the two alternatives.

We have devised simple displays that allow us to dissociate the contributions of local and pattern motion estimates. The basic idea behind the design of our displays is to have multiple linear segments whose locally estimated component motions can be integrated in different ways to yield different overall pattern motion fields. We then assess whether contour percepts change across displays having the same component motions, but different pattern motion fields.

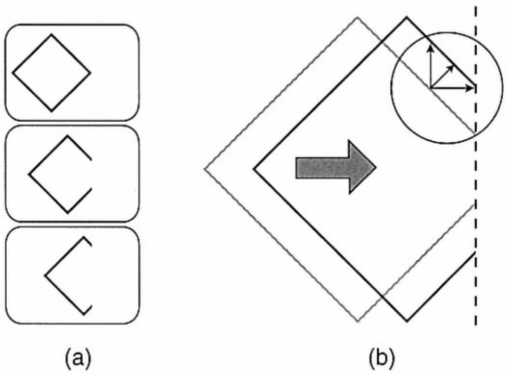

Fig. 1. (a) A simple motion sequence that provides a vivid percept of a subjective occluding contour. The sequence shows a translating diamond sliding underneath an occluder with invisible boundaries. (b) The aperture problem renders the motion information at the interface of the diamond and the occluder ambiguous.

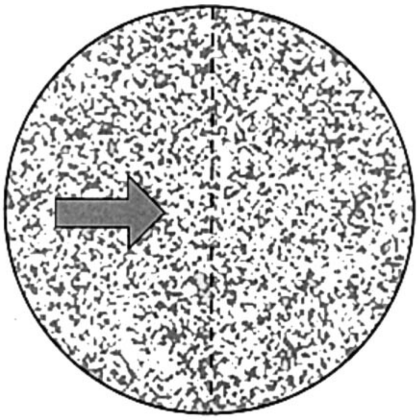

Fig. 2. A typical dynamic display used in studies of motion defined contour perception so far. Because of the presence of unambiguously moving point features, such a display is immune to the aperture problem.

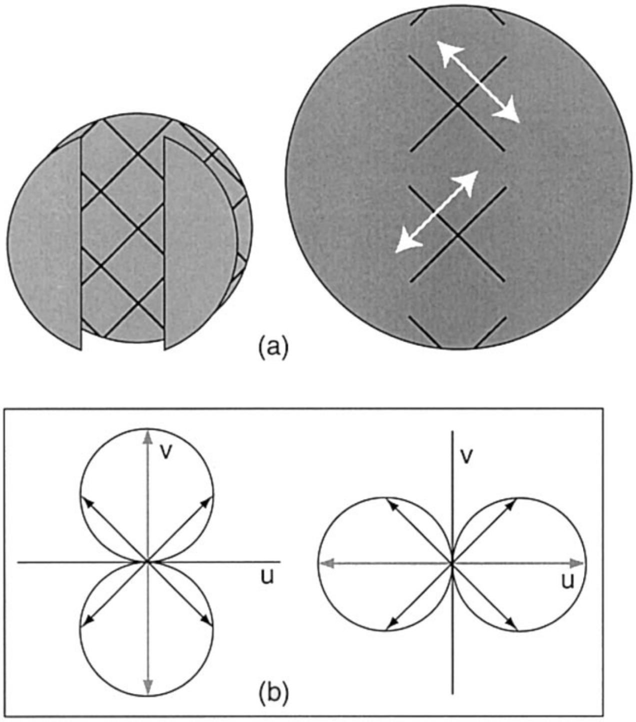

Fig. 3. (a) In our experiments, we used a partially occluded plaid pattern. The left panel indicates the structure of the display, while the right panel shows its appearance as evident to our experimental subjects. The occluders had the same luminance as the background. The constituent gratings of the plaid had identical speeds and spatial frequencies (sf) were mutually orthogonal and oriented at ±45° to the horizontal. (b) By changing the relative phase of the two grating motions, we were able to generate two different pattern motion directions (vertical and horizontal, indicated by gray arrows) using the same set of component motions.

2. Methods and results

2.1. Experiment 1

Our display comprised a partially occluded plaid pattern (Fig. 3a). The occluders created a vertical aperture and had exactly the same brightness as the plaid's background and their boundaries were, therefore, rendered invisible. The two obliquely oriented gratings that made up the plaid were mutually orthogonal. The gratings oscillated perpendicular to their orientations with identical speeds. The identical grating speeds and spatial frequencies induced strong coherence (Movshon et al., 1985).

As shown in Fig. 3(b), by changing only the relative phase of oscillation of the two gratings, the pattern motion direction of the plaid could be rendered either horizontal or vertical. We thus had two displays with identical local motions of the grating bars, but very different pattern motion fields. We reasoned that if contour estimation was based on kinetic discontinuities computed over the local flow, then the two displays would afford identical percepts. If, on the other hand, contour perception was based on discontinuities computed over the pattern motion fields, then the display with the horizontal motion field would yield percepts of subjective occluding contours (at the statically invisible left and right aperture boundaries), while no such contours would be perceived in the display with the vertical motion field. This is because the horizontal motion vectors would have accretion and deletion discontinuities due to the vertically oriented occluders, while no such discontinuities would exist for vertical motion vectors.

The gratings we used had square-wave profiles with low (≈ 0.03) duty-cycles. We experimented with two different grating spatial frequencies (0.58 and 0.75 c deg^{-1} at a viewing distance of 60 cm) and speeds (0.6 and 1.0 deg s^{-1}). This yielded four plaids (in every plaid, the two gratings had identical spatial frequencies and speeds). Each of the four plaids was presented five times (two instances each of vertical and horizontal pattern motions and one static presentation) to every observer, for a total of 20 trials per subject. Each sequence lasted for 5 s. The separation between the two occluders, through which the plaid pattern was visible, subtended an angle of 2.5° at a viewing distance of 60 cm.

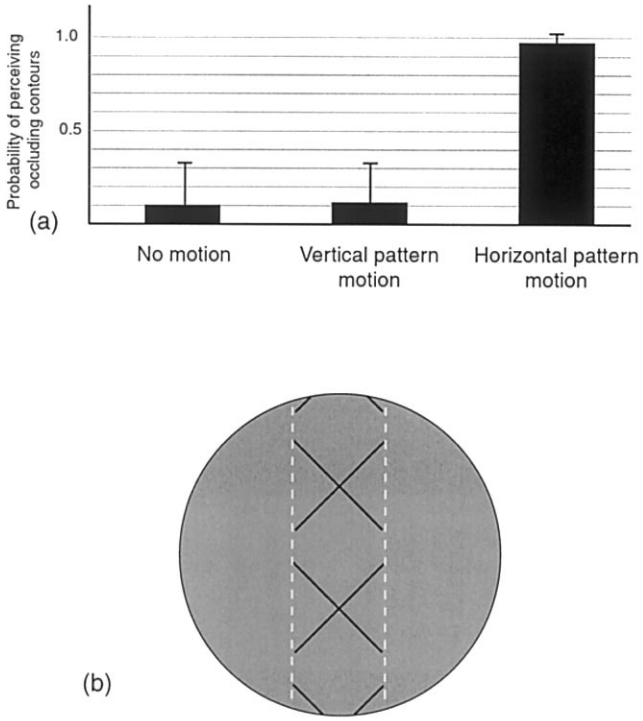

Fig. 4. Results pooled across five subjects. (a) The graph shows the probability of perceiving subjective occluding contours on the left and right margins of the plaid patterns as a function of different pattern motion directions. (b) The locations of the contours reported by the subjects.

All stimuli were presented on a 19″ Trinitron color monitor with antireflective coating connected to a 500 MHz Pentium III computer. The screen resolution was set to 1280 × 1024. Five subjects (the author and four unpaid volunteers) with normal (or corrected to normal) acuity participated in the study. The experiments were conducted individually with each subject. Before testing, subjects were shown a printed image of a Kanizsa triangle (Kanizsa, 1976) to familiarize them with the notion of subjective contours. During testing, subjects were presented with the motion sequences for 5 s each and asked to indicate if they perceived subjective contours anywhere in the display. If they answered in the affirmative, they were asked to indicate the locations of the perceived contours on a printout showing one frame from the motion sequence.

Fig. 4 shows the results pooled across five subjects. The data indicate that the probability of a subject perceiving subjective contours was nearly zero, when the pattern motion was vertical. Strong contours, however, were almost always perceived when the pattern motion was rendered horizontal. This difference is significant at the $P < 0.001$ level (two-tailed t-test for paired data). We also experimented with the entire display rotated by 90°. In this setting, vertical pattern motion yielded vivid contour percepts, while horizontal pattern motion did not. These results, we believe, support the hypothesis of contour perception being based on discontinuities detected over pattern motion fields.

2.2. Experiment 2: effect of keeping local phase relationships unchanged

To test whether the differences in subjects' contour percepts in the two displays might be due simply to the difference in the local motions' relative phase rather than the pattern motion fields, we performed additional experiments using the sequences shown in Fig. 5(a,b). The most important characteristic of these two sequences for our purposes was that they had identical local patterns of motion (speed and relative phase) along their left and right boundaries (zones 'A' and 'B' of the frames shown in Fig. 5a,b). Both sequences depicted four oblique lines in motion. The four-line display subtended an angle of 4° horizontally at a viewing distance of 60 cm. The speed of the line segments was 1.2 deg s^{-1}. By changing the motion of the line terminators near the center of the display, we could induce different pattern motion percepts. In the first, the lines appeared to translate vertically, independently of each other (Fig. 5c). The second sequence allowed a pair-wise grouping of the moving lines, in order to yield a percept of horizontal motion (Fig. 5d).

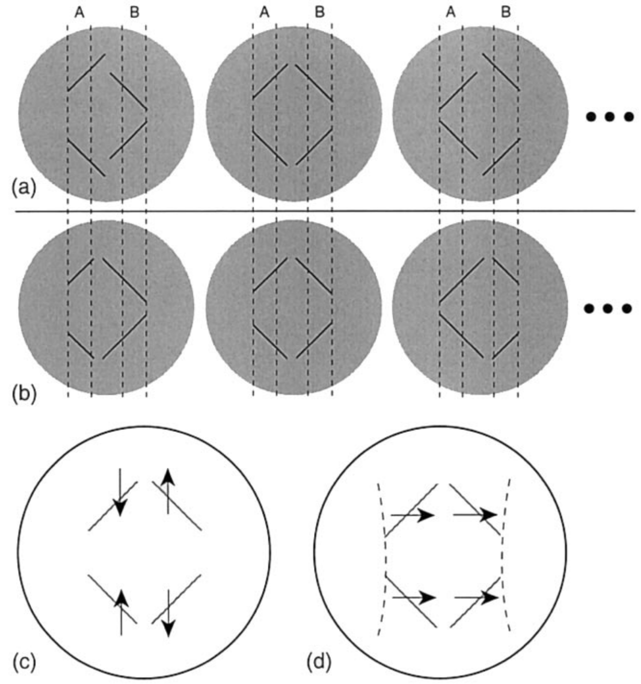

Fig. 5. (a,b) Two sequences showing four lines in motion. The first shows them translating vertically, while the second shows them moving together horizontally. Conceptually, the displays can be divided into three equal-width vertical regions — a middle region flanked by zones ‘A’ and ‘B’. The local patterns of motion in zones ‘A’ and ‘B’ for the two sequences are identical (the amplitude of oscillation of the line segments was set in order to ensure that this held true throughout the sequence). The terminators in the center zone oscillate vertically in the first case and horizontally in the second. The terminator speed for a given line-segment in both cases was 1/cos (45°) times the orthogonal speed of the line segment. (c,d) Reported percepts with sequences (a) and (b), respectively. No contours are perceived in the first case. Strong contours (indicated by dashed lines) are seen delimiting the left and right margins of the moving figure in the second sequence.

Notwithstanding the identical local patterns of motion along their left and right margins, the contour percepts yielded by the two sequences were very different. Subjects did not perceive any subjective contours when the lines were perceived as translating vertically. However, when the lines were seen moving horizontally, strong contours were reported at the locations shown in Fig. 5(d). The lines appeared to undergo accretion and deletion at these boundaries. Fig. 6 summarizes the results. It is also evident from these results that the differences in these percepts were maintained even for small viewing distances, which effectively increased the size of the neighborhood around the occluding contours, where the two motion sequences were identical. These results reinforce the hypothesis that motion-defined contour perception may be based on discontinuities computed over the pattern motion field, rather than the locally estimated flow.

Why does the incidence of contour perception with horizontal motion fields register a small decrease at small viewing distances (Fig. 6)? We believe that this is due to the greater difficulty in obtaining a coherent motion percept when the display subtends a large angle. Subjects reported being unable to perceive occluding contours whenever they failed to get a coherent percept of horizontal pattern motion. Interestingly, some subjects were able to volitionally control the percept of coherent motion. For such subjects, the contour percepts were temporally precisely correlated with the perception of a horizontal motion field. Differential eye movements are unlikely to be able to account for these results, since the percepts were maintained even when the subjects fixated on a point in the center of the display. We find it intriguing that elementary contour percepts are under such intimate control of high-level volition.

3. Discussion

The perception of motion-defined contours has traditionally been assumed to be mediated by early visual processes (Bishop, Coombs, & Henry, 1971; McKee, 1981; Lamme et al., 1993), while the estimation of

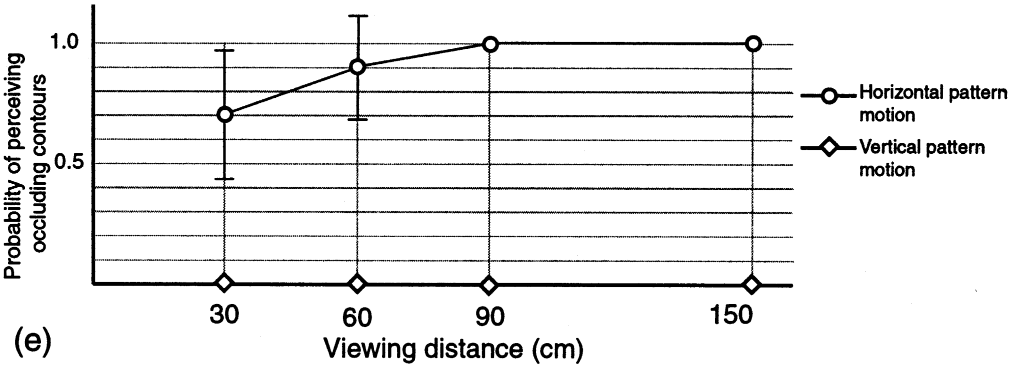

Fig. 6. A summary of results across five subjects using the experimental displays shown in Fig. 5. The probability of perceiving occluding contours is plotted as a function of viewing distance. Notice that even at a viewing distance of 30 cm, when the width of each of the ‘A’ and ‘B’ zones is > 2.5°, the difference in contour percepts is still maintained.

pattern motion is believed to happen at a later stage in the visual pathway (Movshon et al., 1985; Rodman & Albright, 1989; Born & Tootell, 1992). Our results showing that the former might rely on the latter are, therefore, somewhat counter-intuitive. Ecologically, however, this strategy seems justified. It allows for the notion of boundaries to be tied to the notion of perceptually significant entities, such as surfaces and objects (Nakayama, He, & Shimojo, 1995; Shipley & Kellman, 1997; Anderson and Sinha, 1997). It is also worth noting that the idea that contour percepts may rely on relatively sophisticated motion processes serves to complement studies showing influences of figural processing on motion estimation (Shimojo, Silverman & Nakayama, 1989; Vallortigara & Bressan, 1991; Tommasi and Vallortigara, 1999).

It is important to note that our results do not rule out the existence of mechanisms that detect differences in locally estimated flow patterns (Nakayama & Loomis, 1974; Clocksin, 1980; Thompson, Mutch, & Berzins, 1985; Yuille & Ullman, 1990; Spoerri & Ullman, 1991). Indeed, such mechanisms may well underlie the perception of motion-defined forms in displays with dense flow-fields, such as that shown in Fig. 2. However, object surfaces in the real world are often not densely textured and the motion of their bounding contours is typically locally ambiguous. What our results point towards is a mechanism that may be responsible for determining the existence of contours in such settings. This ability to extract extended contours from sparse ambiguous motion information can prove valuable for analyzing dynamic natural imagery.

An interesting issue that arises in the context of our experiments is whether the results should be accounted for in terms of pattern motion directions or occlusion. In other words, instead of saying that contour percepts arise when the pattern motion direction is perpendicular to the boundary, could we not have alternatively said that the percepts arise whenever the display suggests occlusion? I believe that both kinds of accounts are admissible, but the one based on global motion directions is more parsimonious. Terms like 'occlusion' and 'shear' are interpretations based on the underlying motion field (Richards & Lieberman, 1982). If we were to frame the explanation for our results in terms of occlusion and shear, we would have left open the issue that is central to the paper — is the occlusion or shear being determined over the local optical flow or the pattern motion field? What our results demonstrate is that contour percepts arise whenever the pattern motion field is such that it indicates occlusion. The two kinds of accounts, one based on global motion and the other on occlusion are, therefore, not in conflict with each other. We favor the former because it is more specific and enables us to formulate more precise hypotheses about the nature of the underlying processes.

To summarize, a model for motion induced contour perception that relied solely on optic-flow signals estimated in the local neighborhood of the contour locations is unlikely to be able to account for the differences in percepts reported above. Pattern motion estimation (involving spatial integration of motion signals across differently oriented segments and/or over an extended area) seems necessary to explain these results. In terms of the known physiology of the primate visual pathway, the motion-integration based contour perception might be accomplished in at least two, not mutually exclusive, ways. In the conventional hierarchical framework, the contour from motion mechanisms could derive their input from cortical areas, such as the medial–temporal, which are believed to play a role in pattern motion perception (Movshon et al., 1985; Rodman & Albright, 1989; Stoner & Albright, 1992). Alternatively, integration may be achieved via lateral connections (Blasdel, Lund, & Fitzpatrick, 1985; Gray, König, Engel, & Singer, 1989; Sompolinsky, Golomb, & Kleinfeld, 1990; Stemmler, Usher, & Niebur, 1995; Li, 1998) between units in the primary visual cortex. Our inability to conclusively pick a particular hypothesis is a consequence of the lack of a clear consensus regarding how the visual system estimates pattern motion. However, given our current understanding of visual processing, the first hypothesis seems to be the more likely one. This is because it is currently unclear precisely how one might account for pattern motion recovery, which seems necessary for contour perception in our experiments, with the second hypothesis.

To study how the processing constraints suggested by our psychophysical results are actually implemented in the brain, an important next step would be to use the stimuli described here for directly probing neural responses. Three questions deserve particular attention. First, which cortical areas are involved in motion-defined contour perception? A few studies have already yielded some candidate answers to this question (Peterhans & von der Heydt, 1991; Lamme et al., 1993). Potential sites so far include cortical areas V1 and V2. Second, are the contour responses of units in these areas sensitive to pattern motion fields? This question can be addressed using the stimuli described in this paper. For instance, using the display shown in Fig. 3, one can test whether the response of a unit, whose receptive-field straddles the plaid–occluder interface, changes in concert with changes in the plaid pattern motion direction. Especially interesting (but, perhaps experimentally challenging) in this regard would be an investigation of high-level volitional control (alluded to in Section 2) on the responses of such units. And, third, what are the units' response latencies relative to those in the putative pattern motion areas? The existence of a consistent temporal lag in the responses of the contour units relative to the pattern motion units may indicate an involvement of the latter in determining the responses of the former. Answers to these questions, coupled with the psychophysical results of the kind reported here, can aid in furthering our understanding of the architecture of the networks responsible for motion-based contour perception.

Readers can obtain copies of the motion sequences used in the experiments reported above by writing to the author.

Acknowledgements

The author wishes to thank Dr Gadi Geiger, Dr Jitendra Sharma, Professor Randolph Blake, Professor Daniel Kersten, Professor Tomaso Poggio and two anonymous reviewers for useful discussions and comments regarding this work.

References

Anderson, B. L., & Sinha, P. (1997). Reciprocal interactions between occlusion and motion computations. *Proceedings of the Natural Academy of Science USA*, *94*, 3477–3480.

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research*, *10*, 1411–1430.

Bishop, P. O., Coombs, J. S., & Henry, G. H. (1971). Responses to visual contours: spatiotemporal aspects of excitation in the receptive fields of simple striate neurones. *Journal of Physiology*, *London*, *219*, 625–657.

Blasdel, G. G., Lund, J. S., & Fitzpatrick, D. (1985). Intrinsic connections of macaque striate cortex: Axonal projections of cells outside lamina 4C. *Journal of Neuroscience*, *5*, 3350–3369.

Born, R. T., & Tootell, R. B. H. (1992). Segregation of global vs. local motion processing in primate visual area MT. *Investigative Ophthalmology and Visual Science*, *33*, S1315.

Braddick, O. J. (1973). The masking of apparent motion in random-dot patterns. *Vision Research*, *13*, 355–369.

Braddick, O. J. (1974). A short-range process in apparent motion. *Vision Research*, *14*, 519–527.

Clocksin, W. F. (1980). Perception of surface slant and edge labels from optical flow: a computational approach. *Perception*, *9*, 253–269.

Dresp, B., & Bonnet, C. (1991). Psychophysical evidence for low-level processing of illusory contours and surfaces in the Kanizsa square. *Vision Research*, *31*(10), 1813–1817.

Dresp, B., & Bonnet, C. (1993). Psychophysical measures of illusory form perception: Further evidence for local mechanisms. *Vision Research*, *33*(5/6), 759–766.

Gray, C., König, P., Engel, A., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, *338*, 334–337.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque V1 neurons can signal illusory contours. *Nature*, *365*, 550–552.

Hildreth, E. (1984). *The measurement of visual motion*. Cambridge, MA: MIT Press.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.

Kanizsa, G. (1976). Subjective contours. *Scientific American*, *234*, 48–52.

Lamme, V. A. F., Van Dijk, B. W., & Spekreijse, H. (1993). Contour from motion processing in primary visual cortex. *Nature*, *363*, 541–543.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, *10*, 903–940.

McKee, S. P. (1981). A local mechanism for differential velocity detection. *Vision Research*, *21*, 491–500.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross, *Pattern recognition mechanisms* (pp. 117–151). Vatican City: Pontifical Academy of Science.

Nakayama, K., & Loomis, J. M. (1974). Optical velocity patterns, velocity sensitive neurons, and space perception: a hypothesis. *Perception*, *3*, 63–80.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research*, *21*, 427–433.

Nakayama, K., & Silverman, G. (1988). The aperture problem — II. Spatial integration of velocity information along contours. *Vision Research*, *28*, 747–753.

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: a critical link between lower-level and higher-level vision. In S. M. Kosslyn, *Vision: an invitation to cognitive science* (pp. 1–70). Cambridge, MA: MIT Press.

Peterhans, E., & von der Heydt, R. (1991). Elements of form perception in monkey prestriate cortex. In A. Gorea, Y. Fregnac, Z. Kapoula, & J. Findlay, *Representations of vision: trends and tacit assumptions in vision research* (pp. 111–124). Cambridge: Cambridge University Press.

Petry, S., & Meyer, G. E. (1987). *The perception of illusory contours*. New York: Springer-Verlag.

Richards, W., & Lieberman, H. R. (1982). Velocity blindness during shearing motion. *Vision Research*, *22*, 97–100.

Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, *75*, 53–64.

Shipley, T. F., & Kellman, P. J. (1997). Spatio-temporal boundary formation: the role of local motion signals in boundary perception. *Vision Research*, *37*, 1281–1293.

Shimojo, J., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619–626.

Sompolinsky, H., Golomb, D., & Kleinfeld, D. (1990). Global processing of visual stimuli in a neural network of coupled oscillators. *Proceedings of the Natural Academy of Science USA*, *87*, 7200–7204.

Soriano, M., Spillman, L., & Bach, M. (1996). The abutting grating illusion. *Vision Research*, *36*, 109–116.

Spoerri & Ullman, S. (1991). Early detection of motion discontinuities. MIT AI Technical Report, No. 1275.

Stemmler, M., Usher, M., & Niebur, E. (1995). Lateral interactions in primary visual cortex: a model bridging physiology to psychophysics. *Science*, *269*, 1877–1880.

Stoner, G. R., & Albright, T. D. (1992). Neural correlates of perceptual motion coherence. *Nature*, *358*, 412–414.

Thompson, W. B., Mutch, K. M., & Berzins, V. A. (1985). Dynamic occlusion analysis in optical flow fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *7*, 374–383 PAMI.

Tommasi, L., & Vallortigara, G. (1999). Figure ground segregation modulates perceived direction of ambiguous moving gratings and plaids. *Vision Research*, *39*, 777–787.

Vallortigara, G., & Bressan, P. (1991). Occlusion and the perception of coherent motion. *Vision Research*, *31*, 1967–1978.

von der Heydt, R., & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience*, *9*, 1731–1748.

von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). *Illusory contours and cortical neuron responses*, vol. 224 (pp. 1260–1262). New York: Science.

Yuille, A. L., & Ullman, S. (1990). Computational theories of low-level vision. In D. N. Osherson, S. M. Kosslyn, & J. M. Hollerbach, *Visual cognition and action*. Cambridge, MA: MIT Press.